UNIVERSITAT DE BARCELONA

A proposal for the inclusion of accessibility criteria in the authoring workflow of images for scientific articles

Bruno Splendiani

UNIVERSITAT DE BARCELONA

A proposal for the inclusion of accessibility criteria in the authoring workflow of images for scientific articles

Bruno Splendiani

UNIVERSITAT DE BARCELONA

Library and Information Science Department

Faculty of Library and Information Science

Doctoral program in Information and Documentation in the Knowledge Society (Información y Documentación en la Sociedad del Conocimiento)

A proposal for the inclusion of accessibility criteria in the authoring workflow of images for scientific articles

PhD Candidate

Bruno Splendiani

PhD Supervisors

Mireia Ribera Turró, PhD

Roberto García González PhD

Barcelona, Spain 2015

Acknowledgments

To my supervisor Mireia Ribera, who taught me what hard work really means.

To my supervisor Roberto García, for his recommendations and support.

To my friends and PhD students of the Library and Information Science faculty – Anita Locher, Marta Platero, Miguel Navas and Tahere Dehdarirad - for sharing with me all the moments of enthusiasm and discouragement on the long and winding road to a PhD.

To all staff of the Adaptabit Group and the department of Library and Information Science, for supporting me in all aspects, moral, scientific and administrative. Special thanks go to Marina Salse, Miquel Térmens, Miquel Centelles, Inma Rodríguez, Constança Espelt, Ernest Abadal, Mario Pérez-Montoro and Alberto Martinez.

To Josep Marco, for his technical advises about medical imaging.

To Lluis Jover, for his valuable recommendations about academic publishing.

To David Fuster and the staff of the Centro de Diagnóstico por la Imagen Clínic (CDIC) in Barcelona, for helping me in exploring the practices of academic researchers in medical imaging.

To the staff of the GRIHO group, for receiving me during my research internship in Lleida.

To Simon Harper, Caroline Jay, Giorgio Brajnik and all the staff of the Web Ergonomics Lab of the University of Manchester, for their willingness, kindness and consideration of my work.

To my mother Adriana and to my father Giuseppe, I know he would be proud of me if he were here.

To my wife Titika, the “gasoline” of my life, for her admirable patience, ceaseless support and unconditional love. Thank you mami.

To my beloved son Oreste, a true masterpiece.

“Littera gesta docet,

Quod credas allegoria,

Moralia quod agas,

Quo tendas anagogia”

Augustine of Dacia

Contents

Summary of the thesis in Spanish

Esta tesis estudia la cuestión de cómo proporcionar imágenes accesibles en los artículos académicos de Ciencia, Tecnología, Ingeniería y Matemáticas (disciplinas "STEM" en inglés) y en particular en biomedicina. A menudo las figuras son una fuente de información esencial para la comprensión de un artículo científico. Las personas con discapacidades visuales u otras discapacidades se enfrentan a barreras que les impiden acceder a la información proporcionada por las figuras. Por lo tanto, cuando los artículos académicos no incluyen imagen accesibles para todos, la habilidad de percibir y comprender la información científica se reduce drásticamente en lectores ciegos, lectores con discapacidades visuales y cognitivas o lectores que acceden al contenido digital en contextos y condiciones específicas (por ejemplo, utilizando una pantalla reducida y monocromo). Factores éticos, legales y comerciales impulsan a los actores responsables del proceso de publicación a garantizar el acceso universal a los artículos académicos, los principales productos de difusión de información científica.

El estudio propone una serie de directrices dirigidas a los principales actores del proceso editorial para mejorar las capacidades, las oportunidades y las motivaciones en la publicación de imágenes accesibles en artículos científicos de disciplinas STEM. La propuesta se basa en un acercamiento holístico a la accesibilidad de imágenes, consideradas como un complejo objeto de estudio dentro del flujo de trabajo de las publicaciones y que implican la intervención de varios actores con diferentes roles y necesidades informacionales. Por ese motivo, la tesis investiga las políticas y las prácticas actuales de los editores y las prácticas y las opiniones de los autores en el proceso de autoría y envío de imágenes científicas, con el objetivo de optimizar la integración de requerimientos de accesibilidad en el proceso de creación y sumisión de figuras en artículos científicos.

En particular, la propuesta de la tesis incluye una serie de recomendaciones sobre cómo crear imágenes accesibles basadas en las prácticas actuales de edición de imágenes, con el objetivo de minimizar el esfuerzo extra requerido a los autores para adaptar las imágenes a las necesidades específicas de lectores con discapacidad. Además, se propone el uso de textos comúnmente relacionados con figuras en artículos científicos – los pies de figura y las menciones de figuras en el texto principal de los artículos – como textos alternativos, equivalentes textuales de las imágenes que pueden remplazan y describir las figuras sin que se pierda información esencial para su comprensión. El objetivo de esta propuesta es el de ahorrar a los autores (y a los editores) la tarea de crear textos alternativos desde cero.

El estudio se ha estructurado según los pasos siguientes:

- En primer lugar, se ha realizado una revisión preliminar de la literatura dentro de las disciplinas de accesibilidad, biblioteconomía y documentación, visualización de la información y web semántica con el objetivo de documentar las propuestas corrientes sobre el acceso y descripción de información visual. Las diferentes disciplinas contribuyen a definir un modelo para el análisis y descripción de las imágenes y proporcionan las soluciones técnicas para mejorar su accesibilidad.

- A continuación, se ha efectuado una auditoría de las prácticas, políticas y directrices editoriales con respeto a la accesibilidad de las imágenes en revistas académicas en las disciplinas de biomedicina, informática y matemáticas. Esta reseña ha definido el nivel actual de accesibilidad de las imágenes y las características y las funciones de la información visual que normalmente se utiliza en las publicaciones académicas STEM.
- Después, se ha realizado una encuesta y una serie de entrevistas con investigadores en medicina y biología para determinar sus prácticas y opiniones en los procesos de autoría, sumisión y publicación de imágenes en artículos científicos.
- Finalmente, se ha propuesto un conjunto de recomendaciones para mejorar la accesibilidad de las figuras basadas en las prácticas corrientes de autores y editores.

Esta investigación evidencia una falta de accesibilidad de las imágenes y de inclusión de pautas de accesibilidad en el actual flujo de trabajo de autoría de imágenes y propone recomendaciones y directrices para mejorar la situación.

Los resultados de la tesis además evidencian que la adopción de un modelo general de descripción textual para todos los tipos de imágenes científicas presenta limitaciones para la de creación de alternativas textuales eficaces, debido a la variedad de características, funciones y contextos de las imágenes mismas. Por ese motivo, la tesis propone la adopción de un modelo de descripción de imágenes basado en el análisis de cada tipología de imagen y adaptado a cada una de ella, de forma más específica posible.

La contribución principal de la propuesta es la de ser adaptada al contexto específico de la publicación de artículos académicos STEM y de ser conectada a las prácticas corrientes de autoría, sumisión y publicación de imágenes. Esta característica la diferencia de las actuales directrices de accesibilidad, que suelen ser generales y desconectadas de las directrices y prácticas propias del contexto en el que se deberían aplicar. Por consiguiente, la propuesta apunta a limitar los cambios en el actual flujo de trabajo editorial para la implementación de directrices de accesibilidad, aumentando las oportunidades de una verdadera aceptación de principios de accesibilidad.

Overview

This thesis investigates the problem of how to provide accessible images in academic articles in the research fields of Science, Technology, Engineering and Mathematics (STEM) and in particular in biomedicine. Currently, graphics in scientific articles are a critical information source and often provide essential information for a thorough understanding of scientific articles. People with visual and other impairments experience specific barriers that prevent them from accessing the information conveyed by figures. Therefore, when academic articles do not include universally accessible images, the ability to perceive and comprehend scientific information is dramatically reduced in blind readers, readers with visual impairments and other impairments (such as cognitive impairments) or readers accessing digital content under specific context and conditions (e.g., using a small and monochrome display).

This study proposes a set of specific interventions addressed to the main actors involved in the academic publishing workflow in order to improve their capability, opportunity and motivation for publishing accessible images in STEM articles. The proposal is based on a holistic approach to image accessibility, which considers the image as a complex object of study included in the publishing workflow, where many actors with different roles and information needs are involved. Therefore, the thesis investigates the current publishers' policies and practices and the authors' practices and attitudes in the process of image authoring and submission, in order to optimize the integration of accessibility requirements in the image creation and submission process. In particular, the proposal includes recommendations on how to make accessible images based on current practices in image editing, in order to minimize the authors the extra work for adapting images to special needs and suggests the use of suitable, effective and meaningful captions and mentions as texts that can replace and describe the visual information, in order to spare authors (and publishers) the task of creating alternative texts.

The study can be divided into the following steps:

- Firstly, an initial multidisciplinary literature review from the research fields of accessibility, library and information science, information visualization and semantic web is performed to assess the current proposals on the image access and description. The different disciplines contribute to define a model for the analysis and description of images and to provide technical solutions for improving their accessibility.
- Secondly, an audit of the current publishing practices, policies and submission guidelines related to the accessibility of images in academic journals in the fields of biomedicine, computer science and mathematics is performed. This review assesses the current accessibility level of images and the characteristics and functions of the visual information commonly used in academic publications.
- Thirdly, a survey and interviews with academic researchers in medicine and biology are performed for assessing practices and attitudes in the process of image authoring, submission and publishing.

- Finally, a set of recommendations for improving the accessibility of images is proposed. It is based on the current practices of authors and publishers.

This research demonstrates a deficient accessibility of images, a lack of inclusion of accessibility issues into the current authoring workflow and proposes recommendations and guidelines on how to improve these conditions.

The results of the thesis also emphasize that the adoption of a general model of textual descriptions covering all types of images has limited results for the creation of effective text alternatives, due to the high heterogeneity of the characteristics, functions and context of the images. Hence, the thesis proposes the adoption of a model for the textual description of images based on the analysis of each type of image and adapted to each of them.

The main contribution of the proposal is that it is tailored to the specific context of STEM academic publishing and is linked to the current practices in image submission and publishing. This is the main difference of the proposal compared to the current accessibility guidelines, which are seldom general and disconnected from the policies and practices of the specific context where they should be applied. Therefore, it aims at limiting the changes demanded for the implementation of accessibility requirements in the current image publishing workflow, increasing the opportunities of an actual application of accessibility principles.

Chapter 1

Introduction

1 Introduction

1.1 The accessibility of scientific images

The amount of digital information published around the world has increased dramatically over the last decades and its access has been universally widespread thanks to the development of the World Wide Web. An inherent characteristic of the Web is its universality and, according to the intentions of its inventor, Tim Berners- Lee, it is essential that it can be accessed by everyone regardless of disability. Equal access to information and communications technologies, including the Web, and equal opportunity to people with disabilities is a basic human right recognized by the United Nations (United Nations, 2006).

Visual information conveyed by pictures, graphics and all type of images has a pivotal role in complementing the text information. In particular, most of the academic works published in digital format and in article form includes visual information. Images are a critical source in the communication of science concepts (Jones & Broadwell, 2008), since scientific information requires to be “visible” for being understandable (Costa, 1998). In specific research fields such as medicine, a growing number of clinicians, educators, researchers, and other professionals use digital images in their work (Hersh, Müller, & Kalpathy-Cramer, 2009) and figures are a very common type of information used in full-text biomedical journal articles (Apostolova et al., 2013) (Yu et al., 2009). Images provide essential information in biomedical publications (Antani et al., 2008).

People with print impairments - any visual, cognitive, or physical disability that hinders the ability to read print – face barriers in accessing visual information that limit their ability of capturing the whole content of academic research articles. The specific needs of this kind of readers are addressed by digital accessibility, a discipline that offers methods and techniques for providing access to digital resources to all users, including users with various types and degrees of disabilities. Concerning visual resources, the aim of digital accessibility is to make visual content perceptible, operable and comprehensible to all readers, beyond specific disabilities (sensory, cognitive and physical), the technology employed and the context of access.

A typical solution proposed by accessibility guidelines for enabling people with print impairments to access the content of visual information is to provide a text alternative to the image. This solution especially benefits blind people and people with visual impairments, which can read the texts associated to the image through a screen reader or Braille display. People with cognitive disabilities can experience difficulties in understanding the information presented in images (Murphy, 2005) and the audio presentation of the text alternative can also be beneficial to them (Cunningham, 2012). Furthermore, textual descriptions of images allow visual information to be adapted to different access modes, such as listening to an audio version of the article or reading it on the small screen of mobile devices, a media rapidly expanding in specific research fields such as medicine (Perera, 2012). Another parallel advantage of text alternatives, and in particular of the *alt* attribute in the *img* element of

HyperText Markup Language (HTML) documents, is to be indexed by Web crawlers and improve the retrieval of images it refers to (Srinivasarao, Pingali, & Varma, 2011).

Besides text alternatives, other possible alternatives rely on the conversion of information normally acquired by one sense to a modality suitable for another sense. For example, visual information can be converted into a tactile graphics, which enables the access to the graphics by haptic sense. Another alternative is the conversion of visual information into an abstract audio signal, called sonification, which allows graphics being accessed in aural modality. However, these solutions have currently low feasibility and they are not considered as effective as the option of text alternatives in conveying the content of graphics (Keane & Laverentz, 2014).

The need of providing text alternatives to images is acknowledged by widespread accessibility guidelines, such as the Web Content Accessibility Guidelines (WCAG) (W3C, 2008), which suggest mechanisms for associating text alternatives to images. Examples of these mechanisms for images in HTML documents are *alt* attribute, a brief text alternative of the image, and *longdesc* attribute, a long and detailed textual description for complex images. In Portable Document Format (PDF) documents, text alternatives can be associated to images by the use of *alternate text* and *actual text* attributes in the <figure> tag. Nevertheless, the implementation of text alternatives to images on the web is far from being perfect and lots of images have missing, incorrect or poor alternative text (Petrie, Harrison, & Dev, 2005). The lack of descriptive text alternatives has also been recognized as one of the most common accessibility barriers in research journals (Coonin, 2002) and digitized documents in academic libraries (Southwell & Slater, 2012). Concerning scientific literature, it has been emphasized that only a very small part of the research literature is fully understandable because mathematical formulas and critical graphical information is presently not directly accessible by any non-visual means (Gardner, Bulatov, & Kelly, 2009).

The lack of application of text alternatives and other design principles for making images universally accessible hinder users with any type and degree of visual impairments to access visual information. Those users are the main beneficiaries of the solutions proposed for overcoming the barriers to visual content and are the main users to whom the study is addressed. People with cognitive impairments are secondary beneficiaries of these solutions and, in general, all users with a contextual or "situational disability", which depend on the specific context of access and the technology used to access information.

1.2 Justification

The problem of the full access to information in academic publishing to people with print disabilities is neither an insubstantial problem nor applicable only to a tiny number of people (Brown, Jay, & Harper, 2010). Scientific publications are required to guarantee an equal access to information for individuals with disabilities for social and legal reasons. Making digital publications accessible is also financially beneficial for publishers, as it increases the audience of their products.

About 15 percent of world population has a disability (WHO, 2011), 285 million people are visually impaired and age-related visual impairment is increasing along with the percentage of the elderly population in many countries (WHO, 2014). Only in Spain, there are 3, 8 million people with some disability. Among them, almost a million people older than six years have some severe visual impairment (Observatorio estatal de la discapacidad, 2013). The national organization of blind people in Spain (Organización Nacional Ciegos de España (ONCE)) includes more than 71.000 members (ONCE, 2013). Computer use amongst blind people, although below the levels of the population as a whole, is widespread. In the US the estimated visually impaired computer users are approximately 1.5 million (Brown, Jay & Harper, 2010). A survey commissioned by Microsoft on 2003 estimated that 20% of working-age adults in United States have a cognitive difficulty or impairment, which were reduced to 16% when considering only computer users (Microsoft, 2003).

Equal access to academic journals shall respect the principles of emancipatory research, recommended by some publishers (Clark, 2002) (Barnes, 2002). The need for a full inclusion of researches with disabilities into the scientific community has been emphasized by Ribera (2010a).

Traditionally, the right to access scientific publications has been defined from the perspective of the Open Access movement, which considers the peer-reviewed journal literature “accessible” when it is free from economic barriers (BOAI, 2002). According to proposal of Ribera (Ribera, 2010b) requirements for accessible scientific publications should consider not only economic barriers, but also barriers to use. Therefore, academic literature would be “universally accessible” not only when it is available and reachable beyond the financial capability of the reader, but also when it does not exclude any user due to his/her disability.

Many legal mandates over the world require public and private organizations to provide accessible services and products. Many countries, including the European Union, have ratified the UN Convention on the Rights of Persons with Disabilities (United Nations, 2006). Some countries have addressed the access issue through the technology procurement process. For example in the United States Section 508 Amendment to the Rehabilitation Act establishes accessibility requirements for any electronic and information technology developed, maintained, procured, or used by the federal agencies (US Section 508, 2002a). Recently, the European standard EN 301 549 (ETSI, 2014) specifies the functional accessibility requirements applicable to ICT products and services suitable for use in public procurement within Europe. This standard could drive to a scenario in which public authorities, such as universities and hospitals, will be compelled to acquire digital academic journals subject to compliance with accessibility requirements. In the case of Spain, the Real Decreto Legislativo 1/2013 (Real Decreto Legislativo, 2013) grants to people with disability the right to access information and communications technology (ICT) products and services provided by public institutions and private companies. According to these laws and standards, websites of online academic journals and scholarly articles in digital formats should be accessible to all users.

The creation of accessible products has also positive effects on the business (Bias & Mayhew, 2005) for several reasons:

- Accessible design encourages software development practices that follow standards and allow for design for flexible presentation and maximum platform independence. The maximization of the compatibility with user agents enables the possibility of reaching a broader market and reverts in increased sales. This is especially true on some services and products whose target users may have disabilities or are elderly users, such as medical publications.
- The positive market perception (social return on investment) is improved and may lead to better customer retention.
- Information retrieval is also improved when the web content is created to be accessible for people with disabilities, including visual content. As emphasized by Pemberton (2003), Web search engines work similarly to Web users who are blind and accessible Web pages have more chances to be indexed by a Web crawler. In particular, internet search engines assign high importance to the terms present in the *alt* text attribute associated to images (Srinivasarao, Pingali, & Varma, 2011) and their correct application can therefore foster the retrieval of academic articles.
- A text-based alternative description of images could be especially useful in a mobile context of access, for instance in the case in which users billed for download volume turn off images to save costs (W3C, 2013).
- In general, products and services designed for being universal accessible are more usable for everyone. Clearly written and detailed descriptions of images benefit not only readers with disabilities, but also sighted readers (Sutton, 2002).

The need for providing accessible images in academic publications is particularly evident in specific research fields belonging to Science, Technology, Engineering and Mathematics (STEM). In particular, publications in biomedicine, computer science and mathematics have characteristics that justify this need. These characteristics concern the importance of these publications in the current outlook of academic research (both in Spain and internationally), the weight of visual information provided in these publications and their impact on readers with impairments.

It has been estimated that the scientific, technical, and medical (STM) segment of the English-language academic journal publishing industry generated about $9.4 in revenue in 2011 (Ware & Mabe, 2012). Medicine, followed by biochemical, genetics and molecular biology are the research fields with most scientific production in Spain (FECYT, 2013).

Images are a critical source in the communication of science concepts (Jones & Broadwell, 2008), many STEM subjects rely heavily on visual resources (Cryer, 2013) and mathematical formulae, diagrams and other two-dimensional structures are a critical information source in STEM academic literature (Sojka et al., 2013). Figures also provide essential information in biomedical publications (Antani et al., 2008).

Medical information is currently not only searched and consulted by specialists, but also by patients that access it through the web (Iezzoni & O'Day, 2005) and, more specifically, using online medical journals (Smith, 2006), since it enhances their insight and control over their health and health care. The access to medical information by blind people, partially sighted people or other people with disabilities is an issue that has been stressed by some authors (Iezzoni & O'Day, 2005). There is evidence that users with disabilities regularly engage in searching for online health information (Fox, 2007) (Liang, Xue, & Chase, 2011) and, compared with other e-patients, their medical treatment decisions have been influenced more significantly by the information they retrieve online (Fox, 2008).

Several barriers discourage the access to Computer Science career by blind and visually impaired students and this community is significantly underrepresented in computer science (Dunn et al., 2012) (Moon, N.W., 2011). They have often to overcome significant technical and practical barriers (Stefik, Hundhausen, & Smith, 2011), mainly course projects and materials with inaccessible visual resources (Cryer, 2013).

"Math education poses a serious roadblock in entering technical disciplines" (Karshmer, Gupta, & Pontelli, 2007) even for many sighted students. The output of mathematics literature can create even greater barriers to visually impaired students (Peňáz & Ondra, 2012) (Smeureanu & Isaila, 2011) and students with learning disabilities (Lewis, Noble, & Soiffer, 2010), due to the technical notations they include, the large number of visual resources used (such as diagrams, graphs and charts) and the inclusion of visual concepts, such as spatial concepts.

To sum up, the need for making images accessible in academic publications is evident in STEM research fields. In particular in biomedicine, computer science and mathematics publications there are many reasons that justify such a need: the importance of these fields in the current academic research landscape (in Spain and worldwide), the relevance of the visual information in these publications and the impact that visual information can have towards readers with disabilities.

1.3 Accessibility in the publishing process

Traditional approaches to accessibility have focused on the conformance with purely technical criteria defined in accessibility guidelines adopted as legislation. The paradigm of this approach are the WCAG, which are the de facto standard for measuring technical web content accessibility (Sloan & Kelly, 2011) and have been adopted in the legislation of several countries, including Spain. However, the focus on guideline conformance can fall short on providing direct requirements to consider the context of use and the user experience of people with disabilities (Horton & Sloan, 2014). As accessibility efforts have had historically a strong failure rate despite investments made on it (Cerf, 2012) (Ballesteros et al., 2014), new approaches focus on the overall process rather than on the outputs and consider accessibility a process rather than product (EUAIN, 2008). According to this approach, the best way for designing accessible products is by addressing accessibility issues from the beginning of the product lifecycle.

A good example of this new approach is the BS 8878: 2010 Web Accessibility Code of Practice, developed by the British Standard Institute. It provides a framework that aims to embed accessibility into the organization's business ("Organizational Web Accessibility Policy") and to standardize the accessibility in the production workflow following a user-centered production process for every product ("Web Product Accessibility Policy"). It allows the definition – and measurement – of the process undertaken by organizations to procure an optimally accessible web site (Sloan & Kelly, 2011). This standard is process-oriented and complements other web accessibility standards (such as WCAG), which focus on the technical, design or testing elements of accessibility. It provides a code of practice that gives guidance rather than detailed technical specifications on the accessible design of any content for the web ("web products") and is applicable to all types of organizations.

In the last few years, this approach has been proposed for its adoption within the publishing industry, as a novel point of view inside the accessible publishing vision. Accessible publishing aims to publish books and other texts available in alternative formats designed to aid the reading process for people with print impairments. Alternative formats are designed to fulfill the needs of different readers depending on their capabilities, skills and preferences. Some readers prefer a large print, whereas others require formats as a DAISY Digital Talking Books file, an EPUB 3 file or an HTML based e-book, which provide a fully navigable structured content that they can use with text to speech (TTS) software. Other readers prefer reading a document in Braille, which can be accessed in the standard embossed edition or in electronic version through a refreshable Braille display.

Historically, accessible publishing focused primarily on providing material in Braille and other tactile formats and recorded audio versions of books (audiobooks or "talking books") on different supports, such as vinyl records and cassette. Documents in special versions, such as large print or hard copies of braille documents, had the limitation to be substantially larger than standard print documents and often had great costs of production and a considerable delay in distribution. Advances in digital technologies and standards have offered the opportunity to overcome these limitations. The growing availability of digital versions of books (eBooks) has allowed people with print impairment to become customers for mainstream published products. A fully accessible eBook provides a fully accessible reading experience to people with print impairments. They are able to read an eBook independently in synthetic audio (via synthetic or human speech audio), electronic braille (via a connected refreshable braille display) and electronic large print (by increasing the font size on a display screen) (Creaser, 2013). In 2012, 84% of the most popular 1,000 books in the UK were available in the three formats of braille, audio and large print using fully accessible eBooks (Creaser, 2013). However, over 90% of all published materials are still inaccessible by blind or low vision people («WBU Priorities and Goals», s. f.). This scarcity of published works in accessible formats has been defined as "global book famine"(Garrish, 2014, p. 3).

The effort of improving the availability and quality of mainstream publications to people with print disabilities has been addressed since the beginning of the "digital revolution" by organizations that serve communities of people with disabilities. For example, the DAISY Consortium (DAISY Consortium, n. d.), an international consortium that led the transition from analog to digital talking books, includes among its members several organizations that provide library services to blind people (such as the National Library Service for the Blind and Physically Handicapped (NLS), Learning Ally and the National Federation of the Blind (NFB)). The consortium has developed the Digital Accessible Information System (DAISY) format for talking books for people with print impairments and promoted its adoption as a standard in the publishing industry. Daisy format supports the rendering of digital talking books with enabled navigation within a sequential and hierarchical structure consisting of (marked-up) text synchronized with audio. Currently DAISY consortium has advocated for the use of EPUB as the most promising format for accessible publishing, replacing the DAISY format.

Advances in accessible publishing have been driven by the joined efforts of disability rights organizations together with large tech companies and publishers. Currently, several organizations working on accessible academic publishing addressed the need of delivering every native digital content ("born-digital") in an accessible way ("born accessible"). The nonprofit social enterprise organization Benetech, which operates the largest library of accessible books in the world (Bookshare), recommend the "born accessible" approach as a digital imperative for publishers and content creators (Benetech, 2014).

The recently created Accessible Books Consortium (ABC) («Accessible Books Consortium», 2014) promotes inclusive publishing by encouraging publishers to deliver digital publications for sighted audiences that are equally accessible to the print disabled and to adopt an industry-wide accessibility standard. Currently it represents the biggest effort for bringing together all the stakeholders of the accessible publishing. It comprises the World Intellectual Property Organization (WIPO); the DAISY Consortium; organizations representing people with print disabilities, such as the World Blind Union (WBU); authors, International Authors Forum (IAF) and library and information services, International Federation of Library Associations and Institutions (IFLA); publishers associations International Publishers Association (IPA), and individual publishers such as Elsevier, the world-leading provider of scientific, technical and medical information products and services.

The EDItEUR group, an organization developing standards and guidelines within the publishing world, together with the WIPO and the DAISY Consortium, has addressed the inclusion of accessibility policies in the publishing workflow at the organizational level by proposing guidelines to publishers (Hilderley, 2013). The guidelines emphasize that the accessibility features incorporated in publications allow the customization of the reading experience to all readers, whether they have a print impairment or not. A main point of these recommendations is that they are addressed to different roles in the publishing chain. According to them:

- The senior executive should put in place a company policy committed with accessible publishing, by making sure that this is advertised in company literature and on the website, appointing a person or team responsible for the accessibility agenda and personally understanding the issues and benefits of the accessibility in digital environment.
- The internal accessibility lead should be responsible for documenting, communicating and providing the implementation of the accessibility policies. He should have a proper understanding of the company's capabilities and attitudes, workflow and digital files used and stored. The ideal goal of this profile is ensuring that the commercial products will come to have accessibility built in.
- The editorial and design responsible should be aware of issues surrounding accessibility and prepare files not suitable only for printers, but for different digital platforms. This can be achieved following a "XML-first" workflow style, where the production of a master XML file allows the separation of structure, content and style, enabling the creation of any number of different delivery formats. Into this workflow, the preparation and editing of images should follow specific guidelines, especially in the creation of alternative descriptions of the image.
- The production and IT responsible should be able to provide different formats, also according to specific requests for accessible material. As an alternative, this action can be performed by an external supplier. In any case, text alternatives should be attached to each illustration.

The guidelines recommend capturing and retaining the structure of the documents, such as metadata, a hierarchy of headings, footnotes, index, text alternatives of figures, through the publishing process. In ideal circumstances all the structure, including text alternatives of figures, should be included at the authoring stage. In this case, guidelines recommend publishers to provide instructions to authors on how to add structure to the submission file.

The International Digital Publishing Forum (IDPF) has defined and promoted the adoption of the EPUB3 format, which is considered the successor of the DAISY standard and has been vigorously advocated by George Kerscher, the General Secretary of the DAISY Consortium and the president of the IDPF since 2009. EPUB3 includes accessibility features, inherited from the DAISY standard, for improving the reading by persons with print disabilities (International Digital Publishing Forum, 2014b).

An initiative specifically addressed to promote the commitment of scientific research in publishing accessible material ("Accessibility-aware research") is the recommendation plan proposed by the "Research and Development Working Group" (RDWG) of the World Wide Web Consortium (W3C) (W3C Research and Development Working Group, 2015). The proposal includes the creation of guidelines addressed to scientific researchers on how to incorporate accessibility considerations into their research designs and how to report their results in an inclusive way through common channels of academic communication, such as articles in PDF and HTML format or conference presentations.

In conclusion, current digital era advances in publishing offer a unique opportunity to produce and distribute resources in electronic accessible formats from the beginning of the production chain. Given the current requirements and publishing workflows, accessibility can be "built in" from the beginning of the publishing process and, in many cases, it implies just some minimal changes in the established approach (Hilderley, 2013).

In the perspective of the thesis, these new approaches and guidelines are the basis for the proposal of an academic authoring and publishing workflow whose outputs are academic articles with accessible images.

1.4 Publishing workflow of images in academic articles

The publishing chain has been thoroughly modeled for books, and can be defined as "the network of organization and activities involved in the publication, sale and distribution of books" (Thompson, 2013). It includes activities that "facilitate the creation, production, circulation, delivery and consumption of information-based products" (CEN, 2008). The digital publishing chain addresses all the different components between the creation of content and the delivery to the reader. Figure 1 shows a generic illustration of the process, including accessibility requirements along the chain.

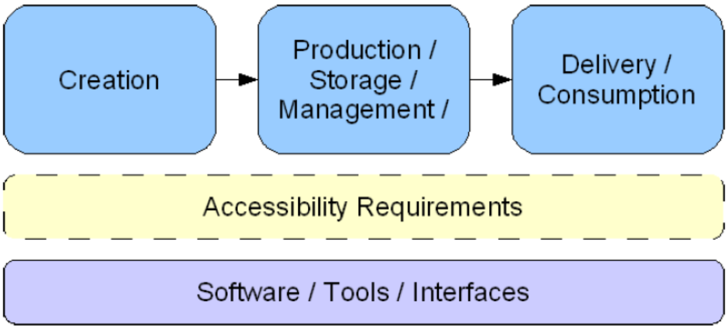


Figure 1 A generic publishing process including accessibility requirements. Extracted from: (CEN, 2008).

For academic journals, the chain includes additional steps concerning the peer-review process. A canonical generic peer-review process, according to the simplified version in Figure 2 (Ware, 2013), includes the following key stages: submission, initial screening and editorial review, selection of reviewers, receipt and consideration of reviewer reports, first decision made and communicated to authors.

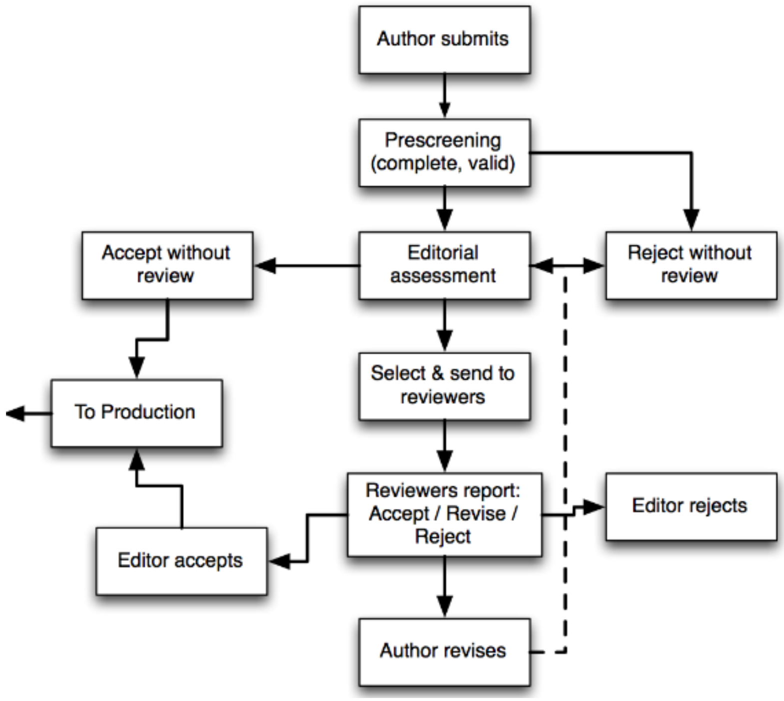


Figure 2 The peer-review process. Extracted from: (Ware, 2013).

This process starts when the author submits the initial manuscript and finishes when the production version of article is published. The published article is the result of the interaction among author(s), publisher(s), editor(s) and reviewer(s).

Image submission is integrated into the electronic submission and publication process of academic articles and many journals require authors to submit their figures as electronic files as part of a totally electronic workflow (Hames, 2007). Submission guidelines specific for images are generally provided in the journal policies. Submission and reviewing processes of the majority of journals in STEM disciplines are managed by an online web-based submission and tracking system (Ware, 2013). This system allows the management of the submitted manuscript and its associated metadata and figures. It can also provide additional features, such as automated production quality checks (such as, for example, testing the resolution of submitted images). Rolandi (Rolandi, Cheng, & Pérez-Kriz, 2011) and Lemke (Lemke, 1998) have emphasized the relevant position of visual information in the peer-review process, since the figures are often the first part of a scientific paper that is reviewed by the editor and examined by reviewers.

However, a complete panoramic vision of the process of image publishing should also consider the steps before the submission. These previous steps are part of the “authoring process”, in which the image is created, generated or acquired from a hardware (equipment) and software (computer programs), edited and archived.

An example of this process can be described in the following scenario: a researcher of a laboratory of nuclear medicine has acquired images of a patient lung from a Positron Emission Tomography (PET) scan. He is studying a specific form of lung cancer and he has found an image in which the pathology can be clearly observed among the scans acquired. Thus, he decides to include the selected image in the academic article he is writing for a submission to a highly cited medical journal. The scan has automatically added some (intrinsic) metadata describing technical and structural features of the image. He reviews the journal submission guidelines and edits the image according to the requirements of the publisher. Through the use of image manipulation software, he highlights the most interesting parts of the figure by drawing arrows and other symbols on it, modifies the dimensions and resolution of the image and changes its format. Additionally, he writes a caption describing the content of the image and assigns a specific title name to the image. He also reviews intrinsic metadata and manually completes them by adding (extrinsic) metadata. Finally the author submits the figure together with the manuscript for its editorial revision.

This scenario presents at a glance the complexity of the "image authoring chain". The complete workflow of images in academic publications is illustrated in Figure 3, where the publishing chain showed in Figure 1 is combined with the authoring steps commented previously.

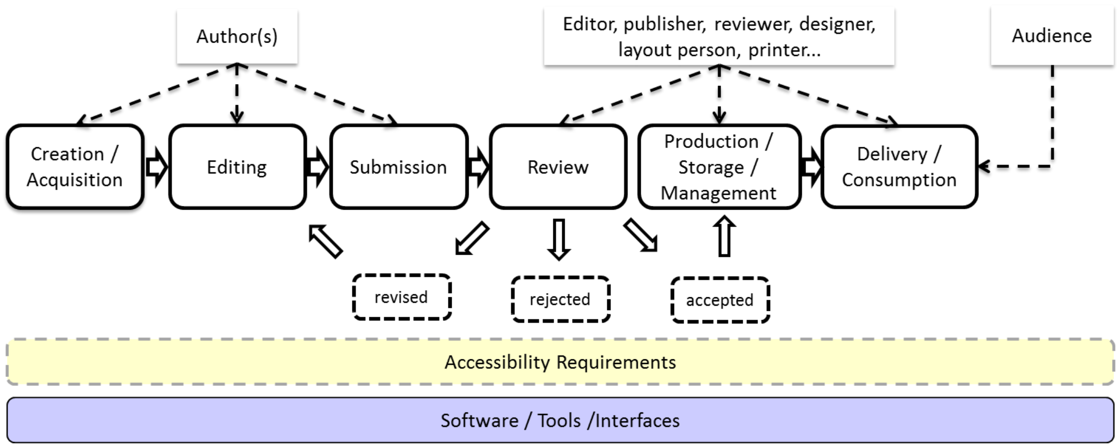


Figure 3 Image publishing workflow including accessibility issues. Figure created by the author.

Figure 3 shows the steps of the workflow and the actors involved in the process. Authors are the protagonists of the first three steps, in which the image is created/acquired, edited and submitted together with an academic article according to the instructions and requirements stated in the journal's "Instructions for Authors". In case of revision, the figure might be required to be modified and in this case the author should edit the image again and repeat the submission step. On each of these steps, accessibility requirements need to be analyzed; thus accessibility should be introduced along the content processing chain.

The rest of the steps are managed by the editorial team, which is responsible for the article editing, process, management and production. With the use of electronic files, the preparation of the manuscript and its related elements (figures, datasets, etc.) starts long before its acceptance (Morris et al., 2013). Editorial assistants check if the submitted manuscript meets the requirements for the peer-review system and for production. Requirements for peer-review affect all elements of the manuscript, including figure files. The production requirements may ask for high-quality figure file types for both online and print publication. Once the article is accepted, it moves into the traditional production phase, during which the manuscript is prepared for composition (i.e. formatting and laying out pages) and publishing online and print (Morris et al., 2013). Articles are published on the Web inside online journals. Online platforms provide the servers, operating systems and software to manage, organize and present the journal content. They can be created by the publishers themselves (e.g., Elsevier's Science Direct or Cambridge University Press's Cambridge Journal Online) or contracted with a hosting platform provider (e.g., Atypon or HighWire).

Online articles are usually accessed by readers starting from an HTML abstract page, which in many cases offers the option to view the full text of the article in HTML or as a PDF, currently the most popular formats for online articles (Morris et al., 2013). Readers with print impairments rely on specific hardware and software for accessing information in general and scholarly information in particular on the screen. Blind people usually make use of screen readers, which allow the textual information on the screen to be read aloud or converted into Braille. When an equivalent text alternative of the content of the image is provided, it can be read aloud by the screen reader. Readers with low vision usually use a screen magnifier software to adapt the content displayed on the monitor (including images) to their specific needs by reducing screen resolution or contrast level, increasing font-size, etc. and seldom to get support from a speech rendering of the information on the screen. Readers with dyslexia sometimes use text-to-speech systems that convert text in digital documents to synthesized speech to support multimodal reading. People with severe dyslexia can use specific screen readers, which are able to read the *alt* attribute and help them in reading images of text (text that has been rendered in image form rather than in selectable characters).

According to the presented framework, the accessible publishing workflow can be defined as a complex process in which many actors are involved. In particular, the authors have a primary role in the process according to the following reasons:

- As noted before, if publishers build accessible images into eBooks and web content from the very beginning of the publishing process, the cost of production is small and the benefits to the end user are large. If not, all the publishing stakeholders (schools, libraries, teachers, students and parents) incur on greater costs down the line when they must retrofit materials to make them accessible (JISC Digital Media, s. f.). In the workflow proposed, the publishing process includes the editing step, where authors prepares manuscript and figure in electronic files according to the publishing guidelines that define how to prepare them (Morris et al., 2013). In this way, authors carry out part of the editing work made by the publisher at the production step.
- The task of creating text alternatives of images should be ideally performed by the authors who submit the figures at the authorial stage, since they have a good understanding of the content of the figures (Hilderley, 2013). Allowing people other than authors to interpret the meaning of an image in a research paper "could undermine the research process" (Kelly et al., 2009).

The way the actors involved in the process take decisions and perform their tasks depends on several factors: common and personal habits, thoughts and motivations; the availability of technical accessibility standards they have to follow; the availability of tools enabling the production of accessible content and their general awareness of the requirements and resources for accessibility (Brewer, 2013).

Considering that the current application of solutions for making images accessible is considered very low and inadequate (as commented in section 1.1), a more holistic approach to the problem should consider the factors previously cited and propose specific changes in policies and practices in the publishing workflow. The definition of such interventions should be based on the knowledge of the current behavior patterns of the actors in the image publishing chain.

The Behaviour Change Wheel (BCW) (Michie, Stralen, & West, 2011) is a method for the definition of behavior change interventions supported by evidences in the targeted behavior. Although this method is addressed primarily to design interventions for changing practices in clinical medicine and public health, it is a general synthesis of existing frameworks covering different fields of intervention and it has the advantage to integrate elements of individual rational and contextual influence (Niedderer et al., 2014). According to this method, the process of applying effective behavior change interventions should follow the steps below:

- Identify the target behavior/s and understand them in context: identify key specific behaviors of the different actors that may have an impact on the accessibility of images (who, what, when, how). Understand the causes of such key behaviors ("why behaviors are as they are?") and what is needed to change for the desired behaviors to occur. The "Capability/Opportunity/Motivation - Behavior (COM-B)" model can be used as a reference to answer these questions. Such a model considers the behavior as a part of an interactive system involving three other components: capability (the psychological and physical ability to engage in the activity concerned), opportunity (physical and social environment that enables the behavior) and motivation (reflective and automatic brain processes that activate or inhibit the activity concerned).
  Concerning the activity of making images accessible for publishing by authors, the COM-B analysis can be developed as an evaluation of the three components as follows:

1. Capability: knowledge, understanding and skills of the authors on how to make images accessible
2. Opportunity: Appropriate tools, standards and social context prompting authors to create accessible images.
3. Motivation: Appropriate context that stimulate and persuade authors on reasoning about the ethical and legal consequences of making accessible images.

- Consider full range of possible activities designed to change behaviors (intervention functions) and identify specific behavior change techniques. Explore the full range of options of intervention according to nine intervention functions (Table I), which are related to the sources of behavior (capability, opportunity, motivation). Finally, define the specific behavior change techniques.

Table I Intervention functions of the Behavior Change Wheel (BCW) model

Intervention function	Definition	Component
Education	Increasing knowledge or understanding	Capability, Motivation
Persuasion	Using communication to induce positive or negative feelings or stimulate action	Motivation
Incentivisation	Creating prospect of reward	Motivation
Coercion	Creating expectation of punishment or cost	Motivation
Training	Facilitating the development of skills	Capability
Restriction	Introducing measures that prohibit or limit undesirable behaviors	Opportunity
Environmental restructuring	Changing the physical or social context	Motivation, Opportunity
Modelling	Providing an example for people to aspire to or imitate	Motivation
Enablement	Increasing means/reducing barriers to increase capability or opportunity	Capability, Opportunity

To implement each of these interventions, individuals can invoke a variety of policies, decisions made by authorities that enable or support interventions, selected according to seven categories (Table II).

Table II Example of policy categories for the implementation of interventions defined in the BCW model

Policy category	Example
Communication/ marketing	Widespread communication of best practices using print, electronic, telephonic or broadcast media
Guidelines	Creating documents that recommend or mandate practice
Fiscal	Using the tax system to reduce or increase the financial cost
Regulation	Establishing rules or principles of behavior or practice
Legislation	Making or changing laws
Environmental/ social planning	Designing and/or controlling the physical or social environment
Service provision	Delivering a submission system

The Behaviour Change Wheel approach provides a systematic method to build a proposal to link possible interventions for the effective inclusion of accessible images in STEM publishing to an analysis of the behaviors of the actors of the "image publishing workflow" defined in Figure 3.

Therefore, this thesis investigates the "image publishing workflow" and especially the current practices and attitudes of the authors in the "image authoring chain" (creation, editing and submission of images) as the starting point for proposing solutions of improvement of the accessibility with a limited impact on the current image workflow, according to the Behaviour Change Wheel approach. The study is structured in three steps (plus an introductory step) addressing different stages and actors in the "image publishing workflow", as showed in Figure 4 and detailed below:

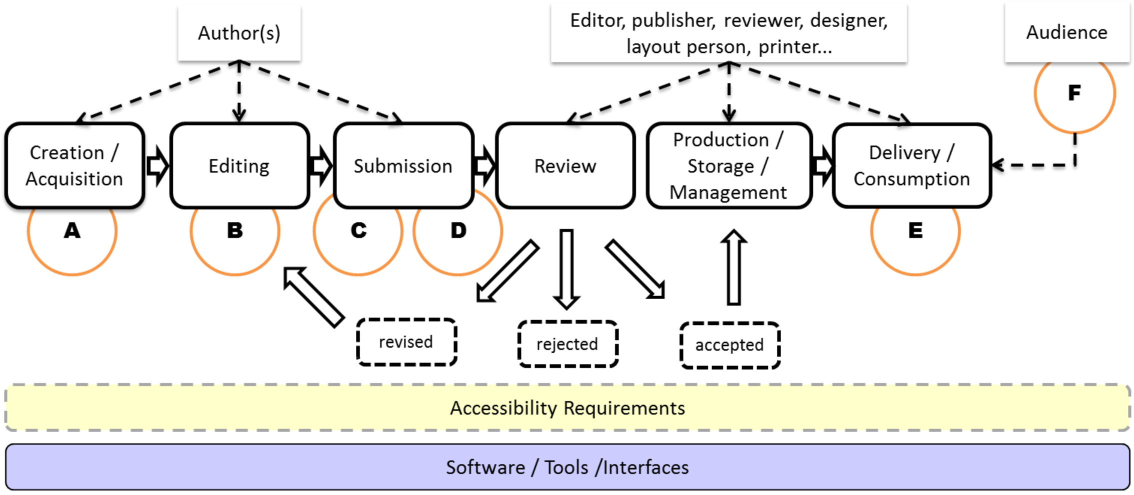


Figure 4 The image publishing workflow with the main point analyzed in the thesis: creation/acquisition of the image by the author (point A); editing of the image by the author (point B); submission of the image by the author (point C); submission guidelines of the publishers (point D); delivery of the image by the publishers (point E); profile of the audience (point F).

Step 0. A review of digital documents and graphics formats used in academic digital publishing and their accessibility support is done, in order to introduce key concepts used in the study.

Step 1. The profile of the target audience of accessible images (point F in Figure 4) is assessed and the current solutions for making images accessible in STEM academic publishing are reviewed.

The profile of the audience is defined in order to tailor the interventions of the behavior change to the specific needs of readers with impairments and the barriers they face when accessing images. In particular, the interventions concerning the education of the authors about the needs of readers with disabilities should be based on this analysis.

Benefits and limitations of the current guidelines and techniques for making images accessible are analyzed, paying particular attention to the recommendations and mechanisms for creating and adding text alternatives to images from an interdisciplinary perspective. The review of the state of the art of these solutions will help in identifying the most appropriate techniques for the academic publishing context. The intervention concerning the training of authors on making images accessible should be based on these techniques.

Step 2. An analysis of the actual accessibility of images published in STEM articles at the moment of their delivery (point E in Figure 4) is performed. This analysis should confirm the assumptions that images are currently not accessible in academic articles in STEM fields and they do not have text alternatives associated to them.

A review of the image submission guidelines for authors provided by publishers of academic journals (point D in Figure 4), and especially those guidelines related to accessibility, is also performed. The aim of this review is to assess the image accessibility requirements currently included by the publishers in the submission guidelines. Based on this, interventions aimed at restricting the requirement for the image submission according to accessible design principles may be proposed.

Step 3. The authoring workflow (the “creation/acquisition”, “editing” and “submission” steps; point A, B and C in Figure 4 respectively) is analyzed to define the current practices and attitudes of the authors in the figure authoring process in biomedical academic articles. Biomedicine is the field selected among other STEM fields for this analysis due to its importance in the scientific production in Spain (FECYT, 2013). More in detail, this analysis aims at assessing:

- The function and use of image and image-related texts (captions, mentions, etc.) in academic articles according to the opinions of their authors;
- the technical standards and tools for image editing and submission currently used among them;
- the author’s awareness of accessibility requirements and resources.

The information collected with this analysis will be_used as a basis for tailoring interventions addressed to authors for improving the process of authoring and submitting accessible images.

1.5 Purpose of the thesis

The main aim of the thesis is to propose realistic and affordable solutions for including accessibility criteria into the authoring workflow of images within academic articles in STEM domains – and in particular in biomedicine - with a minimal impact on the current publishing workflow.

The secondary objectives of the thesis related to the main aim are:

1. To define the profile of the target reader of accessible images.
2. To review the current solutions for making accessible images offered by the accessibility discipline.
3. To assess whether the figures in STEM academic articles are currently accessible. This aim depends on the starting assumption that images are currently not accessible in academic articles in STEM fields and, in particular, images do not have text alternatives associated to them.
4. To assess whether the figures in STEM academic articles are relevant.
5. To assess whether the publishers of STEM journals currently offer effective instructions for authors to make their images accessible.
6. To explore to potential use of current image-related texts currently used in STEM and in particular in biomedical academic articles as alternative descriptions of the images. This aim was defined according to the findings of the literature review. However, it is anticipated in this comprehensive list for facilitating the comprehension of the readers of the thesis.
7. To assess how and why authors currently create and submit images in biomedical academic articles. This objective is focused on biomedicine as it is considered the research field among STEM fields with most impact on academic research at a national and international level.
8. To propose recommendations on how to foster the adoption of accessibility principles concerning the images within the biomedical academic publishing. Recommendations for guiding authors in the application of effective text alternatives are also proposed. The guidelines for the formulation of text alternatives are limited to a specific type of medical image. The thesis does not aim to propose a standard alternative text covering all the possible types of graphic content available in academic publications.

The study answers the following eight questions defined according to the aims:

Q1. Which are the profiles of the target readers with disabilities of STEM academic articles?

Q2.Which are the current solutions for making visual content in STEM academic articles accessible to people with disabilities?

Q3. Are figures in STEM academic articles and in particular in biomedical articles currently accessible?

Q4. Are images in STEM academic articles and in particular in biomedical articles a relevant source of information?

Q5. Are there effective instructions for authors to make their images accessible in STEM and in particular in biomedical academic articles?

Q6. Are current image-related texts appropriate as alternative descriptions of the images in STEM and in particular in biomedical academic articles?

Q7. How do authors create and submit images in biomedical academic articles?

Q8. How to foster the adoption of accessibility principles concerning images within the STEM publishing and in particular in biomedical publishing?

1.6 Thesis outline

The thesis is organized as follows:

Chapter 2 (“Methodology”) presents the methodology used in this thesis.

Chapter 3 (“Definitions and background”) presents the conceptual framework of the thesis, basic terminology about images, the characteristics of digital images and the most common formats of image and document currently used in academic publications.

Chapter 4 (“Literature review“) presents a literature review of theories, models, guidelines and best practices about the analysis and description of images in digital Accessibility, Information Visualization (InfoVis), Library and Information Science (LIS) and Semantic Web.

Chapter 5 (“Analysis of published articles and authors’ practices and attitudes”) describes the analysis of the current policies and practices of inclusion of figures in articles in biomedicine, computer science and mathematics journals, with particular emphasis on assessing their current accessibility. The analysis of authors’ behaviors and attitudes concerning the use and function of images and image-related texts in academic articles in medicine and biology is also presented. Results are evaluated and discussed according to the research questions.

Chapter 6 (“Proposal”) presents a proposal of guidance addressed to the different actors in the authoring workflow for the production and publication of accessible images in biomedical academic articles, according to the Behavior Change Wheel model. In particular, recommendations for guiding authors in the application of effective text alternatives are proposed. Specific recommendations in the formulation of text alternative are limited to a specific type of medical image. The thesis does not aim to propose a model of text alternative covering all types of images currently included in academic publications.

Chapter 7 (“Conclusions”) presents the conclusions of the thesis.

Chapter 8 (“Contributions”) presents a lists of publications derived from the studies developed in the thesis and disseminate in articles in academic journals and in presentations and poster sessions in international conferences and workshops.

Chapter 2

Methodology

2 Methodology

The study is executed according to three main steps, which answer to the questions defined in section 1.5 and investigate the specific key points of the "image publishing chain" defined in Figure 4 in section 1.4. Multiple research methodologies and techniques are adopted and quantitative and qualitative methodologies are combined, as follows:

- A review of the bibliography about the image accessibility and alternative descriptions from a multidisciplinary viewpoint.
- An audit of the current policies and practices in the publishing of STEM articles and in particular of biomedical articles.
- An analysis of the practices, needs and attitudes of authors of biomedical academic articles collected through the techniques of survey and interview.

2.1 Literature review

A bibliographic review was performed with the aim to define the profile and needs of readers who benefit from making STEM academic articles accessible (point *F* in the image publishing chain showed in Figure 5) and to explore the current solutions on how to make images accessible according to the target audience profile.

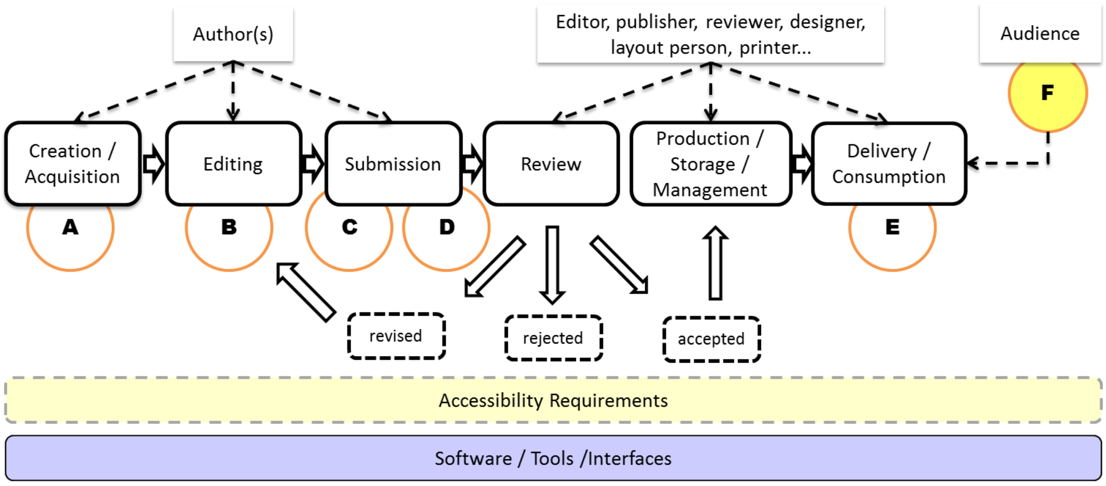


Figure 5 The image publishing workflow" with the main point analyzed in the literature review (audience) emphasized.

Its objective was to answer the following questions:

Q1. Which are the profiles of the target readers with disabilities of STEM academic articles?

Q2.Which are the current solutions for making visual content in STEM academic articles accessible to people with disabilities?

Firstly, the most adopted guidelines for the creation of accessible content (WCAG 2.0) were reviewed in order identify and describe the design principles and recommendations about accessible images, the profile of the users to whom these guidelines are addressed and the most common barriers in accessing images.

Secondly, methods and techniques for making images in scientific articles accessible were reviewed, focusing on the inclusion of text alternatives in the images included in the most common document formats used in academic publishing (HTML and PDF). The specific recommendations on how to create text alternatives were thoroughly reviewed. Besides the WCAG, other sources consulted in this review were:

- Academic researchers and internationally renewed institutions and organizations in the accessibility area, such as the National Center for Accessible Media (NCAM), the DIAGRAM Center, the UK Association for Accessible Formats (UKAAF), the Royal National Institute for Blind People (RNIB) and WebAIM.
- Researchers and organizations in the academic digital publishing, such as the International Digital Publishing Forum (IDPF).
- Other organizations addressing specific issues in the publishing workflow, such as the Metadata Working Group.

Thirdly, relevant research projects proposing solutions for the adaptation of images to readers with disabilities were identified and described. This research was extended to include the contributions from the Semantic Web research field.

Fourthly, the specific issue of which information should be included in text alternatives was addressed by reviewing the sources in accessibility fields previously commented. Due to the complexity of the visual information conveyed by scientific figures and the task of expressing it by words, theoretical contributions from the main models for textually describing images offered by the Library and Information Science discipline were also reviewed.

2.2 Audit of policies and practices in STEM articles publishing

An audit of the policy and implementation of visual content in highly-cited articles in biomedicine, computer science and mathematics journals was performed.

This analysis focused on the submission and delivery steps in the image publishing chain (point D and E in Figure 6, respectively).

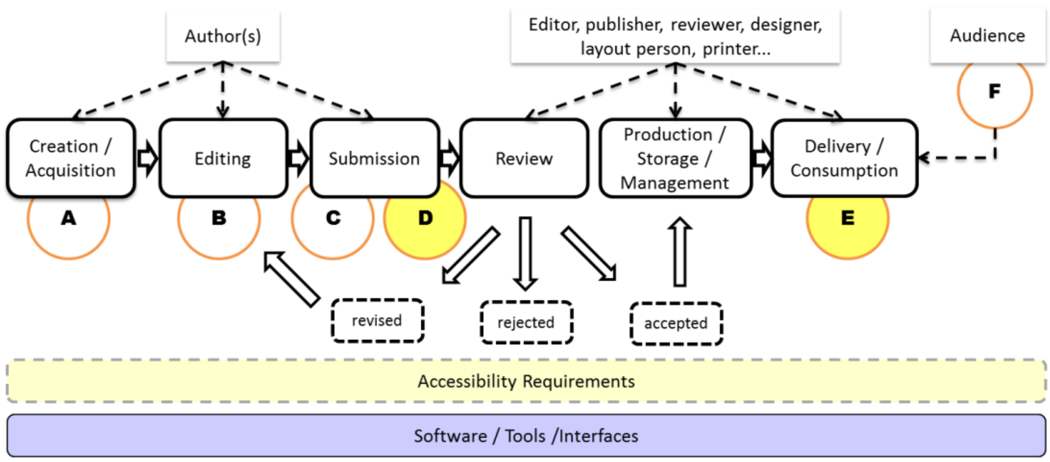


Figure 6 The image publishing workflow with the main points addressed in the audit of policies and practices (submission and delivery/consumption) emphasized.

Its aim is to verify the assumption that images in STEM academic publications are not accessible and, in particular, images do not have text alternatives associated to them, and to answer the following questions:

Q3. Are figures in STEM academic articles and in particular in biomedical articles currently accessible?

Q4. Are images in STEM academic articles and in particular in biomedical articles a relevant source of information?

Q5. Are there effective instructions for authors to make their images accessible in STEM and in particular in biomedical academic articles?

Q6. Are current image-related texts appropriate as alternative descriptions of the images in STEM and in particular in biomedical academic articles?

According to the questions, the analysis was structured as follows:

1. Analysis of the accessibility of images published in the articles in biomedicine, computer science and mathematics (point E in the image publishing chain). In particular, to verify the assumption that images in STEM academic publications do not have text alternatives associated to them, the application of text alternatives to images was assessed. The correspondence between the submission practices in the application of textual image descriptions and the accessibility policy provided by the journal was also analyzed. This analysis aimed to answer question Q3 ("Are figures in STEM academic articles and in particular in biomedical articles currently accessible?").
2. Analysis of the use of figures in biomedical articles, computer science and mathematics, in order to answer question Q4 ("Are images in STEM academic articles and in particular in biomedical articles a relevant source of information?).
3. Audit of the submission policies for authors provided by the publishers of the journals where the STEM articles were published (point *D* in the image publishing chain). This analysis aimed to answer question Q5 ("Are there effective instructions for authors to make their images accessible in STEM and in particular in biomedical academic articles?").
4. An additional analysis was performed to assess the function and use of image-related texts (captions, mentions, text in the figures and metadata) currently used in biomedical academic articles, with the aim of evaluating the extent to which they could be potentially appropriate as alternative descriptions of the images. This analysis aimed to answer question Q6 ("Are current image-related texts appropriate as alternative descriptions of the images in STEM and in particular in biomedical academic articles?").

2.2.1 Sources of information

This study was focused on the analysis of academic articles in online journals in the research fields of biomedicine, computer science and mathematics. These journals were reasonably considered representative of the STEM academic journals. In particular biomedicine was most thoroughly reviewed as the major publishing discipline in STEM.

32 academic journals with the highest impact factor were chosen for the study. The impact factor is a cumulative sum of the citations of a particular journal and a numerical indicator of its repercussion on other works. The Journal Citation Reports (JCR), an annual publication of the Science and Scholarly Research division of Thomson Reuters, ranks academic journals according to this factor.

Twelve journals in the field of biomedicine with the highest impact factor according to the JCR Science Edition 2011 (accessed 26 September 2012) were identified. The selection included two multidisciplinary journals that publish articles in the field of biomedicine (*Nature* and *Science*) because of their high impact factor and their relevance in the research community. The JCR subjects of the selected journals were Oncology, Medicine, General and Internal, Immunology, Cell Biology, Genetics and Heredity, Immunology Biochemistry and Molecular Biology and Multidisciplinary science. For each journal, the first research article published in 2012, containing figures and available online was selected.

Ten journals in computer science with the highest impact factor according to the ranking of the JCR Science Edition on 2011 (accessed 23 April 2013) were identified. The *International Journal of Neural Systems* (impact factor 4.260) was initially selected. However, due to subscription limitations that didn't allow accessing a digital copy of its articles, this journal was replaced by the *International Journal of Computer Vision* (impact factor 3.741), the eleventh in the JCR ranking. The JCR subjects of the selected journals in Computer Science were: Computer Science and Information Systems, Artificial Intelligence, Interdisciplinary Applications and Theory & Methods. For each journal, the last research article containing figures, available online and published not later than April 2013 was selected.

Ten journals in mathematics with the highest impact factor according to the ranking of the JCR Science Edition (accessed 01 October 2013) were identified. The JCR subjects of the selected journals in mathematics were Mathematics and Applied Mathematics. For each journal, an article containing figures, available online and published not later than October 2013 was selected.

2.2.2 Requirements about figures

The concept of "figure" is elusive. Thus, images that illustrate, explain or complete the text content of the article were primarily considered for the analysis. Then other types of images, such as the photograph of the author and the logo of the journal, were also included. Non-informative images as advertisements and icons were excluded. Neither tables nor mathematical formula were considered, as they can be offered by more effective solutions than figures with an alternative text, such as data tables and MathML code (see section 3.4). The condition of exclusion was also applied to the case of a table defined as "image" in the article CS5 (Figure 7).

Algorithms based on characters comparison		
Ahmed-Kaykobad-Chowdhury	Ahmed et al. [2003]	Sec. 4.1.1
Fast-Search	Cantone and Faro [2003a]	Sec. 4.1.1
Forward-Fast-Search	Cantone and Faro [2005]	Sec. 4.1.1
Backward-Fast-Search, Fast-Boyer-Moore	Cantone and Faro [2005]	Sec. 4.1.1
Sheik-Sumit-Anindya-Balakrishnan-Sekar	Sheik et al. [2004]	Sec. 4.1.1
Thathoo-Virmani-Sai-Balakrishnan-Sekar	Thathoo et al. [2006]	Sec. 4.1.1
Boyer-Moore-Horspool using Probabilities	Nebel [2006]	Sec. 4.1.1
Franek-Jennings-Smyth	Franek et al. [2007]	Sec. 4.1.1
2-Block Boyer-Moore	Sustik and Moore [2007]	Sec. 4.1.1
Two-Sliding-Window	Hudaib et al. [2008]	Sec. 4.1.2
Wu-Manber for Single Pattern Matching	Lecroq [2007]	Sec. 4.1.3
Boyer-Moore-Horspool with q-grams	Kalsi et al. [2008]	Sec. 4.1.3
Genomic Rapid Algo for String Pm	Deusdado and Carvalho [2009]	Sec. 4.1.3
SSEF	Külekci [2009]	Sec. 4.1.4
Algorithms based on automata		
Extended Backward Oracle Matching	Faro and Lecroq [2008]	Sec. 4.2.1
Simplified Extended Backward Oracle Matching	Fan et al. [2009]	Sec. 4.2.1
Forward Backward Oracle Matching	Faro and Lecroq [2008]	Sec. 4.2.1
Simplified Forward Backward Oracle Matching	Fan et al. [2009]	Sec. 4.2.1
Double Forward DAWG Matching	Allauzen and Raffinot [2000]	Sec. 4.2.2
Wide Window	He et al. [2005]	Sec. 4.2.3
Linear DAWG Matching	He et al. [2005]	Sec. 4.2.3
Improved Linear DAWG Matching	Liu et al. [2006]	Sec. 4.2.3
Improved Linear DAWG Matching 2	Liu et al. [2006]	Sec. 4.2.3
Algorithms based on bit-parallelism		
Simplified BNDM	Peltola and Tarhio [2003]	Sec. 4.3.1
BNDM with loop-unrolling	Holub and Durian [2005]	Sec. 4.3.1
Simplified BNDM with loop-unrolling	Holub and Durian [2005]	Sec. 4.3.1
BNDM with Horspool Shift	Holub and Durian [2005]	Sec. 4.3.1
Horspool with BNDM test	Holub and Durian [2005]	Sec. 4.3.1
Forward SBNDM	Faro and Lecroq [2008]	Sec. 4.3.1
Forward Non-deterministic DAWG Matching	Holub and Durian [2005]	Sec. 4.3.2
Two-Way Non-deterministic DAWG Matching	Peltola and Tarhio [2003]	Sec. 4.3.2
Bit parallel Wide Window	He et al. [2005]	Sec. 4.3.3
Bit-Parallel^2 Wide-Window	Cantone et al. [2010b]	Sec. 4.3.3
Bit-Parallel Wide-Window^2	Cantone et al. [2010b]	Sec. 4.3.3
Small Alphabet Bit-Parallel	Zhang et al. [2009]	Sec. 4.3.3
Shift Vector Matching	Peltola and Tarhio [2003]	Sec. 4.3.4
Fast Average Optimal Shift-Or	Fredriksson and Grabowski [2005]	Sec. 4.3.5
Average Optimal Shift-Or	Fredriksson and Grabowski [2005]	Sec. 4.3.5
Bit-Parallel Length Invariant Matcher	Külekci [2008]	Sec. 4.3.5
BNDM for Long patterns	Navarro and Raffinot [2000]	Sec. 4.3.5
Long patterns BNDM	Peltola and Tarhio [2003]	Sec. 4.3.5
Factorized BNDM	Cantone et al. [2010a]	Sec. 4.3.5
Factorized Shift-And	Cantone et al. [2010a]	Sec. 4.3.5
BNDM with q-grams	Durian et al. [2009]	Sec. 4.3.6
Simplified BNDM with q-grams	Durian et al. [2009]	Sec. 4.3.6
Shift-Or with q-grams	Durian et al. [2009]	Sec. 4.3.6
Constant-space string matching algorithms		
Tailed-Substring	Cantone and Faro [2004]	Sec. 4.4.1

Figure 7 Example of a table defined as a figure in an article. Extracted from (Faro & Lecroq, 2013).

The only exception to this condition was the case of a table with a decorative element defined as "image" in a biomedical article (Figure 8), which was analyzed as the rest of figures.

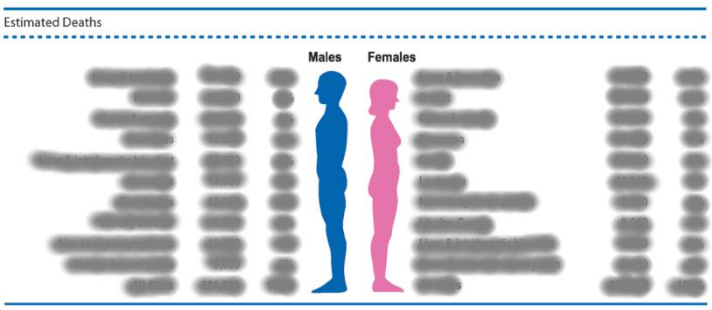


Figure 8 Example of a table with decorative elements defined as a figure in an article. Image edited by the authors. Extracted from (Siegel, Naishadham, & Jemal, 2012).

2.2.3 Limitations of the study

The analysis performed was manual. Such a method was chosen since it is more efficient for collecting precise data on graphics in academic articles and overcomes the limitations of the automatic analysis methods (Hartley et al., 2014). For example, the automatic count of graphs in publishing documents may require the support of optical recognition software (OCR). These limitations were experienced during the study. For example, several attempts were performed to automatically extract figures and figures' metadata from PDF documents with several tools, such as the QPDF library (Berkenbilt, 2014), the XPDF software (Glyph & Cog, 2014) and the LogicMighty XMP Explorer software (Logicmighty Ltd, 2014). The information provided by the automatic extraction was always considered partial and finally a manual revision was required.

While the manual method made possible to perform a very in-depth analysis, it was very time-consuming and therefore a limited set of journals and articles was selected.

2.2.4 Analysis of the journals

Each of the 32 journals was manually analyzed and documented as follows:

1. The accessibility policy of the journal, as stated in the journal policies on the publisher's website (where available) was reviewed. Usually, the accessibility policy is provided through an accessibility statement declaring the conformance of the website to the accessibility requirements and offering related information about accessibility.
2. The submission guidelines of the journal addressed to authors were reviewed analyzing the journal instructions for authors on the publisher's website and specific recommendations related to the image accessibility were retrieved (where available).

   The guidelines for authors regarding image submission in the articles, as stated on the journal's website were examined. The instructions concerning the submission of figures as stated in the general submissions guidelines for authors and other instructions provided by the publisher (if available), such as specific documents on how to prepare artwork for submission, were reviewed. The reviewed documentation included instructions concerning technical guidelines for image submission and in some cases tips or "Frequently Asked Questions" related to them. In the submission instructions, references about the specific treatment of figures targeted to users with special needs (for example, use of specific colors for color-blind readers or high contrast for low-vision readers) were identified when available.

   The technical requirements of the image were also reviewed, especially those which can have an impact on the accessibility of the images: digital image format accepted, recommended resolution, recommended dimensions, recommendations on the use of color and color contrast, preferred file size and name, preferred font size and family of image labels and comments on multipanel images.

   Finally, the instructions concerning image-related texts in the article, such as the caption, the textual references to images in the main content of the article (mentions) and metadata were reviewed.
3. For biomedical journals, the retrieval options for visual information on the journal's website were also reviewed. It was checked whether the online journals offered a specific retrieval system for images, and whether it used image descriptions. The presentation options of the images (such enlarged version, zoom option, etc.) were also reviewed.

2.2.5 Analysis of the articles

2.2.5.1 Quantitative analysis of documents and images

The analysis of the characteristics of the published STEM articles was manually performed on 30 articles: 10 articles in biomedicine, 10 in computer science and 10 in mathematics. Two biomedical articles belonging to the journals with the lowest impact factor among those selected in biomedical discipline were excluded from the analysis, in order to compare the same number of articles in each discipline.

The key topics for the analysis were selected according to the general principles of accessibility features in digital visual information identified according to the WCAG 2.0 and the conclusions of the literature review. They will be described in detail in chapter 4.

Each article was manually analyzed and documented at two levels:

1. Whole article in PDF version
   General information about the PDF document: total number of pages; figure count per document and page; inclusion of structural tags, level of encryption; version; application or software used to create/edit and convert the original document to PDF.

2. Images in the article
   In the PDF and HTML versions

   - Text alternatives. Application of the *alt* attribute in the *img* element of the HTML version of the article and the *alternate text* and *actual text* attributes in the <figure> tag of the PDF version of the article.
   - Technical features of the image, in particular those on which can depend the accessibility of the images: digital image format, resolution, recommended dimensions, use of color and color contrast, file size and name, font size and family of image labels.

   Only in the PDF version

   - General information about the figures in the PDF document: type of graphics (vector or bitmap); position in the document, page number and section.
   - Information extracted from the metadata related to the figures within the PDF document: creator, title, format, height, width, color space, bit depth, compression filter and file size.
   - Textual labels inside the image: text transcription; length in characters; position in the document, page number and section.
   - Captions and mentions in the text: text transcription; length in characters; position in the document, page number and section.

The complete dataset of the collected information is available on the website of the Adaptabit group (University of Barcelona) (Adaptabit, 2014).

2.2.5.2 Lexical analysis of captions and mentions

Besides the general analysis of captions and mentions, which measured their application in PDF articles and their length, specific quantitative and qualitative content analyses of the text included in these image-related texts were performed to analyze their lexical content and compare their functions, identify complementarities and overlaps in their use and assess the extent to which they could be potentially appropriate as alternative descriptions of the images.

Quantitative content analysis is a method mainly used to count manifest textual elements and produces numbers that can be manipulated with statistical methods. It was applied to compare the frequency of occurrences of technical vocabulary in captions and mentions, based on the Paul Nation's Vocabulary Profiler (University of Hong Kong, 2001).

However, the quantitative content analysis method is often criticized for missing syntactical and semantic information embedded in the text (Zhang & Wildemuth, 2009). For this reason, the qualitative content analysis was combined with a qualitative analysis. Qualitative content analysis explores the meanings underlying physical messages and it is mainly inductive, grounding the examination of topics and themes in the data.

Since it is intensive and time-consuming, content analysis usually involves small samples (Macnamara, 2005). Therefore, this analysis was performed on a subsample of nine articles (three for each discipline) chosen among those selected for the previous analysis. The subsample was selected in order to reflect the variety of types of images presented in the three disciplines (illustrations, diagrams, photographs (including medical images) and line art).

Text in captions and mentions was manually preprocessed in order to fix typographic errors. Special characters such as α, β and other mathematical symbols, individual letters and numbers were not considered. The text was analyzed by the TAPoR comparator tool (TAPoR Team, 2012), applying the list of English corpus stop words provided by the software. Three lists of occurrences were generated: word types that appeared only in captions, word types that appeared only in mentions and word types that appeared in both texts. Each word type was manually assigned to one of the three general categories of the theoretical model for the image analysis proposed by Jaimes & Chang (1999):

- Syntactic level: visual and perceptual attributes defining general visual characteristics of the image (type and technique), low-level perceptual features (i.e. the overall color or texture of the image), particular features within the image (such dot, line, texture) and composition (e.g. centered objects, diagonal leading line, and symmetry).
- Semantic level: visual attributes associated with the concepts conveyed by the visual content. They describe the objects and scenes showed in the image at different levels, from a generic object (e.g. a lumbar spine) and scene (e.g. oblique/rotated view of a scanned brain) to a more specific or abstract objects (e.g. gas patterns showed in an x-ray) and scene (e.g. pathology judgments in a medical context).
- External or non-visual attributes: information referring to the history and context of the image such as date, location, rights and source.

This model, which will be described in detail in section 4.9, is a conceptual framework focused on indexing different aspects of general visual information. However, it was selected as a model for the content analysis following the example of Wang (Wang et al., 2011), who applied it in the categorization of textual descriptions of medical images.

2.3 Analysis of authors' practices and attitudes

An analysis of the authors' practices, needs and attitudes concerning the use and functions of images in academic articles in medicine and biology was performed, paying particular attention to their awareness and perception of accessibility. Due to time constrains, this analysis was focused on the biomedical research field, excluding other research STEM fields such as Computer Science and Mathematics.

The use and function of image-related texts, and especially of captions, was also analyzed in depth. The analysis focused on the steps of creation/acquisition, editing and submission in the authoring workflow (steps *A*, *B* and *C* in Figure 9, respectively).

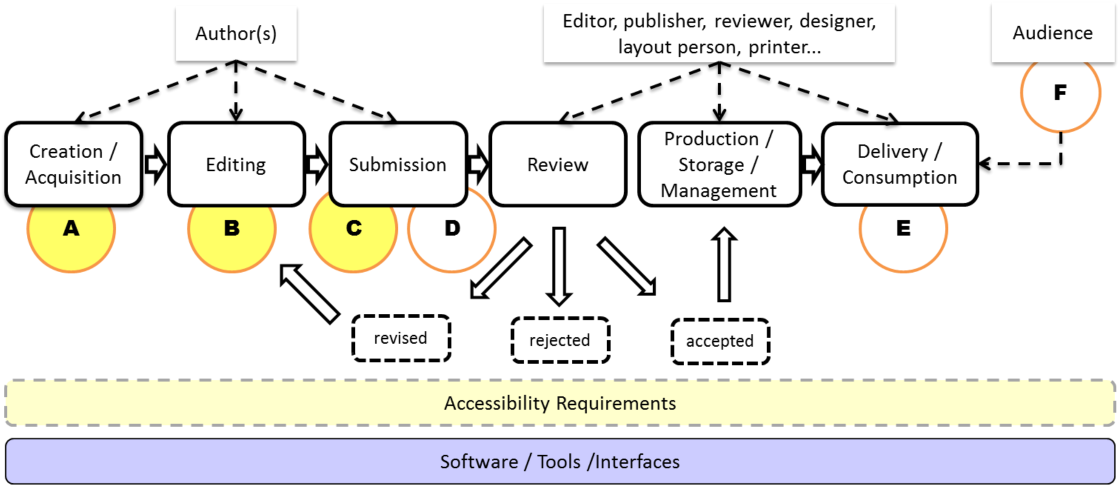


Figure 9 The image publishing workflow with the main points addressed in the analysis of the authors' practices and attitudes (creation/acquisition, Editing and submission) emphasized.

This aims of this analysis is to assess the current practices of researchers in medicine and biology in the selection and inclusion of images and descriptive text in academic articles, as a necessary foundation for defining practical recommendations for the inclusion of accessible images in academic articles and enabling the support of the authors and the widely adoption into the academic publishing.

A secondary objective of the analysis is to evaluate the possibilities of descriptive texts commonly associated to images, in particular captions and mentions, as a mechanism to give access to visual content in scientific articles aligned with current practices. Another secondary objective is assessing the relevance of images in STEM academic articles according to the opinions of their authors.

In synthesis, the aim of this analysis is to answer the following questions:

Q4. Are images in STEM academic articles and in particular in biomedical articles a relevant source of information?

Q6. Are current image-related texts appropriate as alternative descriptions of the images in STEM and in particular in biomedical academic articles?

Q7. How do authors create and submit images in biomedical academic articles?

Although many categories of researchers use images in their work, this study was limited to the fields of medicine and biology, since they are the research fields with most scientific production in Spain (FECYT, 2013).

The investigation was carried out using qualitative research methods, since they are most appropriate when the purpose of the study is to investigate human thoughts, believes and behaviors (Denzin & Lincoln, 2005) and are considered particularly useful for understanding under-researched areas, as authoring images is (Beaudoin, 2009). Qualitative research methods are also considered appropriate for studies adopting the perspective of the Library and Information science disciplines (Borrego & Frías, 2004), such as this thesis does.

The study combined multiple collection techniques in the form of a survey and a semi-structured interview.

2.3.1 Online survey

An online transverse survey was created with the aim to study the practices, thoughts and beliefs of academic researchers in medicine and biology in order to understand why and how they use, edit and describe images for publishing purposes. The survey was descriptive and consisted of a combination of 23 open-ended questions and 2 closed-ended questions (9 single selection and 16 multiple selections). The complete questionnaire (in Spanish) is available in Appendix A.

In order to define of the questions and the possible value of the answers, a literature review was previously performed identifying similar research (Beaudoin, 2009) (Graham, Perriss, & Scarsbrook, 2005) (Hersh, Müller, & Kalpathy-Cramer, 2009) (Kahn et al., 2012) (Sedghi, Sanderson, & Clough, 2008) (Sedghi, Sanderson, & Clough, 2012) (Sedghi, Sanderson, & Clough, 2011). These references were the source of inspiration for the creation of the questions, except in the case of the questions related to accessibility, which were created ad hoc.

The questions were grouped to gather the following information:

- Participant's demographics and interests (questions P1- P7).
- The use of images in the publications of the participants: type of publication, motivations, type of image and format (P7 – P11).
- The access to images: frequency, relevance of the image-related text, sources of the images, relevance of the images in academic articles (P12 – P15).
- Common practices in the textual description of images (annotations and textual descriptions) (P16-P21).
- Knowledge and awareness of image accessibility (P22 - P25).

2.3.1.1 Characteristics and selection of the participants of the survey

The inclusion criterion was empiric and the selection criteria were the following:

- To be part of the academic staff of the University of Barcelona belonging to a research group in the medicine or biology fields.

- To have published academic publications in the last few years before the survey.

The selection criteria were intentionally general, since the purpose of the survey was to explore the current motivations and practices in image submission from a panoramic point of view of the current needs and behaviors of a specific group of researchers in medicine and biology. Researchers from the University of Barcelona were selected since they belonged to the same university of the author of the study. This condition was supposed to raise their engagement to the study. The fields of medicine and biology were selected for being the research fields with most scientific production in Spain, as commented before. The scope of the study was also limited to these two specific fields according to the limited resources available and the possible workload performed by one person over a period of several months of data collection and analysis.

In order to follow the selection criteria, the participants were selected in the directory of researchers of the University of Barcelona (GREC) (GREC, 2013) on April 2013. Although the GREC database is oriented to the institutional control of the research activity of the research staff of the University of Barcelona, it was supposed that the criteria of belonging to a research group gave reliability to the sample (inclusive in the case when the information was not completely updated in the GREC database and the researcher would have belonged to a research group previously), taking into account the objectives of the investigation.

Researchers belonging to groups of the Faculty of Medicine and the Faculty of Biology involved in research projects were selected and their contact information was recollected (Figure 10).

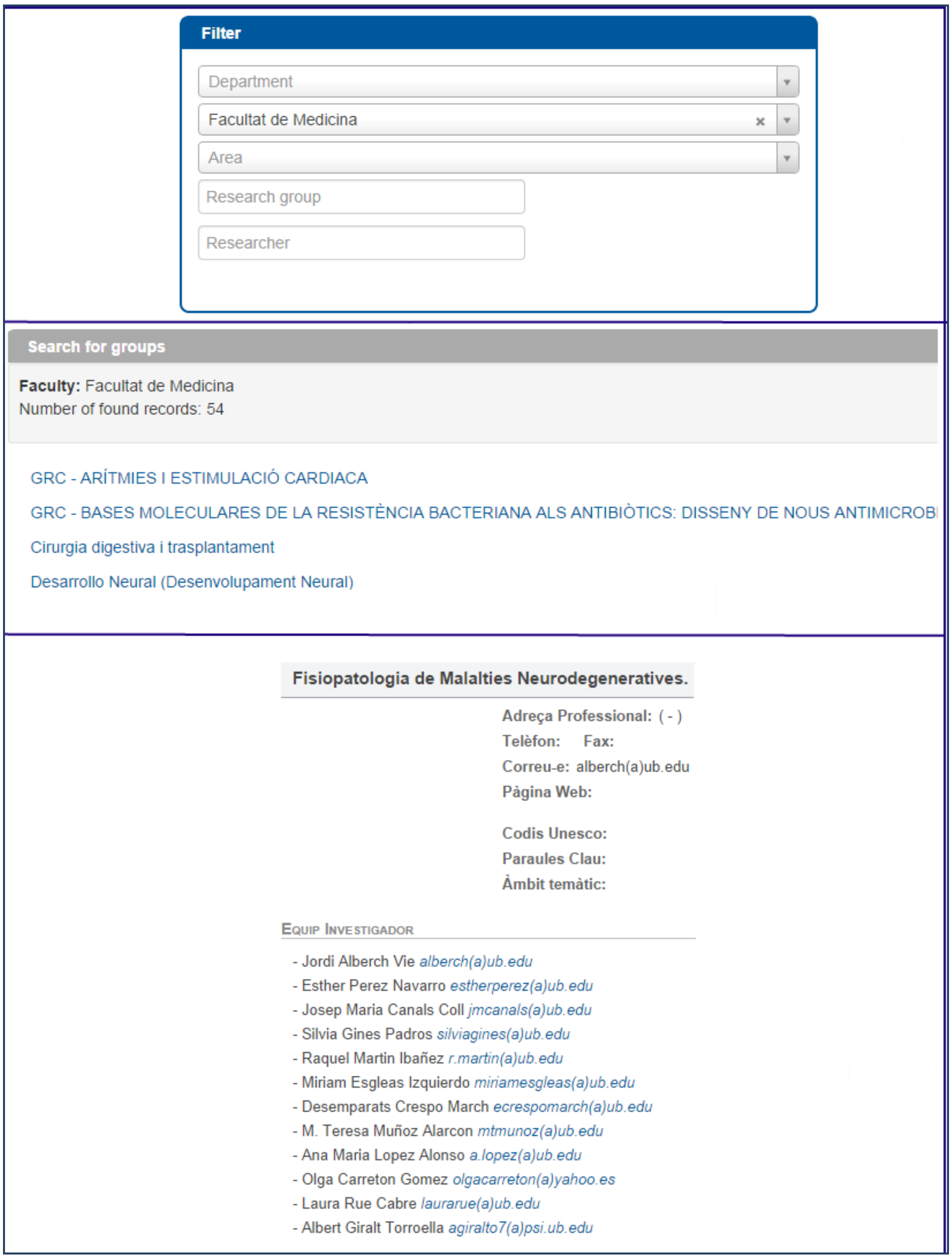


Figure 10 Screenshots of GREC interface: search filter, results of research groups in the Faculty of Medicine and example of personal information about the member of a research group in the medicine area.

2.3.1.2 Recruitment of the participants of the survey

1005 emails were sent on April 2013 to the selected participants using Limesurvey – a survey software tool - and 797 emails were delivered (79%). The email (appendix B, in Spanish) included an introductory presentation of the study and a link to the survey. Participants were also informed that the information of the survey was collected anonymously, treated as strictly confidential and used only for research purposes. After a week a remainder email (appendix B, in Spanish) was sent to those participants who had not responded to the first email. A confirmation email (appendix B, in Spanish) was sent to the participants who answered the survey.

From the 797 participants who correctly received the email, 198 participants responded to the questionnaire: 179 participants completed the entire questionnaire (23%), 19 completed the questionnaire partially (2%), while 551 participants did not answer the questionnaire (75%) (Figure 11).

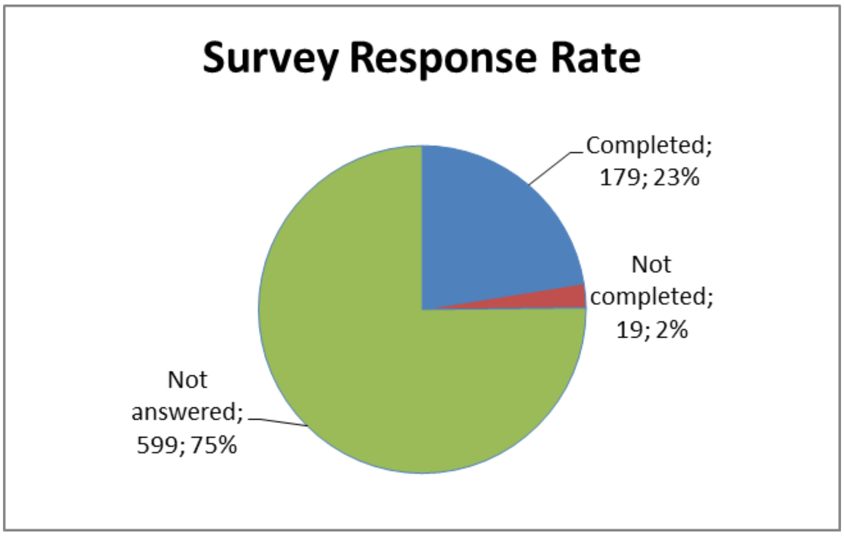


Figure 11 Response rate of the online survey

2.3.1.3 Analysis of the survey's data

The responses of the participants were extracted and analyzed. Results from responses to closed questions were automatically generated by the Limesurvey software and ordered in tables. Open-ended questions were analyzed according to the qualitative data analysis method that will be described in the next section.

2.3.1.4 Limitations of the survey

The focus of this study was on the fields of medicine and biology and the participants of the survey were selected among researchers from the University of Barcelona. Concerning the first limitation, it is important to emphasize that the conclusions derived from the results of the study apply only to the specific fields analyzed, and further studies are required in order to compare the findings about visual information to other research fields and reinforce the conclusions. Concerning the second limitation, since the participants belonged to the same institution in Spain, further studies should be performed in order to extrapolate the results of the analysis to other countries or inclusive to other academic institutions in Spain.

2.3.2 Interviews

In order to examine more in deep some interesting aspects emerged from the survey's answers and complete the information collected from the survey and to gain a more qualitative perspective (Beaudoin, 2009) of the opinions and practices of the researchers concerning the use and function of the images in academic publications, a semi-structured face-to-face interview was prepared. It was considered opportune to perform the interview in the same workplace of the interviewees, since the presence of artifacts frequently enables the examination of the workflow, resources and tools (Marshall, 2009).

The interview addressed the following topics:

- Image needs: Why do academic researchers in medicine and biology select and use images for publication purposes?
- Image use: How do academic researchers in medicine and biology select and use images for publication purposes?
- Related texts needs: Why do academic researchers in medicine and biology select and use textual complementary description of images in academic papers?
- Related texts use: How do academic researchers in medicine and biology select and use textual description of images in academic papers?
- Reading: Why and how do academic researchers in medicine and biology look at images and their related text during the reading of an academic paper?

2.3.2.1 Characteristics and selection of the participants of the interview

The inclusion criteria for selecting participants for the interview were to be a researcher in the medicine and biology fields and to regularly use images in his/her works. Participant sampling was also based on the purpose of selecting information-rich cases representing the diversity of profiles involved in the daily use of images for research purposes in medicine and biology.

2.3.2.2 Recruitment of the participants of the interview

The participants' recruitment process involved two steps: at the beginning, people who answered the survey were contacted by email (appendix B, in Spanish) and were asked to take part in the interview. This first selection was especially focused on recruiting researchers who were physicians involved into academic research (purposeful sampling), in order to selecting information-rich cases who could provide detailed descriptions of the behaviors and motivations of a homogenous group. Due to the scarce number of participants obtained, the sample was increased applying the snowball sampling technique, in which existing study subjects recruited other participants from among their acquaintances who met the eligibility criteria (see below) and could potentially contribute to the study. These new participants belonged to a network of academic and professional contacts working in institutions linked to the University of Barcelona: the University Hospital "Clínic" of Barcelona and the University Hospital of Bellvitge (Barcelona), which are linked to the Faculty of Medicine of the University of Barcelona, and the August Pi i Sunyer Biomedical Research Institute (IDIBAPS) in Barcelona, which is linked to the Faculty of Medicine of the University of Barcelona and to the University Hospital "Clínic" of Barcelona. Participants' profiles included a variety of specialization, work experience and position. Detailed Information about the profiles of the participants is showed in Table III.

Table III Participants of the interview for research area (Med: biomedicine; Bio: biology), specialization, years of experience in research, membership organization, academic position, academic degree, professional group and sex. P: participant.

	Research field	Specialization	Experience (years)	Membership organization	Academic position	Academic degree	Professional group	Sex
P1	Med	Physiology	13	University of Barcelona	Graduate teaching assistant	PhD	Physician	M
P2	Med/Bio	Medical statistics	18	University of Barcelona	Senior lecturer	PhD	Professor	M
P3	Med	Medical statistics	30	University of Barcelona	Senior lecturer	PhD	Professor	M
P4	Med	Biomedical engineering	5	IDIBAPS	no academic position	MD	Engineer	M
P5	Med	Biomedical engineering	1,5	IDIBAPS	no academic position	MD	Engineer	W
P6	Med	Nuclear medicine	13	Clinic Hospital	Graduate teaching assistant	PhD	Physician	M
P7	Med	Nuclear medicine	20	Clinic Hospital	Graduate teaching assistant	PhD	Physician	W
P8	Med	Nuclear medicine	15	Clinic Hospital	Graduate teaching assistant	PhD	Physician	M
P9	Med	Oral Pathology	30	University of Barcelona	Senior lecturer	PhD	Professor	M
P10	Med	Sport medicine	24	University of Barcelona	Lecturer	PhD	Professor	M
P11	Med	Sport medicine	21	Ramon Llull University	Senior lecturer	PhD	Professor	W
P12	Bio	Molecular biology	5	University of Barcelona	Research assistant	PhD	Professor	W
P13	Bio	Genetics	20	University of Barcelona	Senior lecturer	PhD	Professor	W
P14	Bio	Molecular biology	28	University of Barcelona	Senior lecturer	PhD	Professor	W
P15	Med	Nuclear medicine	34	Clinic Hospital	Lecturer	PhD	Professor	M

	Research field	Specialization	Experience (years)	Membership organization	Academic position	Academic degree	Professional group	Sex
P16	Med	Nuclear medicine	15	Clinic Hospital	Graduate teaching assistant	PhD	Physician	M
P17	Med	Nuclear medicine	15	Clinic Hospital	Graduate teaching assistant	PhD	Physician	W
P18	Med	Neuroscience	5	University of Barcelona	Lecturer	PhD	Professor	W
P19	Med	Medical microbiology	22	Bellvitge Hospital	Graduate teaching assistant	PhD	Physician	W
P20	Med	Gastroenterology	34	Bellvitge Hospital	Graduate teaching assistant	PhD	Physician	M
P21	Bio	Physiology	32	University of Barcelona	Professor	PhD	Professor	M
P22	Bio	Bioenergetics	22	University of Barcelona	Senior lecturer	PhD	Professor	W

2.3.2.3 Collection of the interviews' data

In order to answer the questions of the study, data was collected through one-on-one semi-structured interviews. A set of questions was created as a general script of the interview (appendix C, in Spanish). The issues studied in each question were categorized into predefined general themes according to the five general questions of the study, to facilitate the data analysis process. Table IV shows the categorization of the issues. During the interview, the order of the questions was not strictly observed and varied according to the answers of the participants. The approach of the interviewer was focused on motivating participants by insisting on those aspects on which they seemed to be more interested or on which they appeared to be more self-confident and talkative. Furthermore, the participants were free to discuss further topics parallel or tangential to those provided during the interview.

Table IV Grid of the categories of information collected in the interview.

Topic of the question	Image needs	Image use	Related text needs	Related text use	Reading
Background	x				
Work performed	x				
Typical image needs and use	x	x			
Image needs and use in publication	x	x			
Image selection	x				
Image editing		x			
Image tools	x	x			
Textual description needs			x	x	
Textual description use				x	
Textual description creation			x	x	
Textual description guidelines			x	x	
Image reading					x
Textual descriptions reading					x
Reading process					x
Barriers to reading					x
Image search	x	x			
Textual description search			x	x	

The interviews were conducted face-to-face with 22 participants in the interviewees' offices or laboratory located at the University Hospital "Clínic" of Barcelona (six participants); the University Hospital of Bellvitge (two participants); the August Pi i Sunyer Biomedical Research Institute (IDIBAPS) in Barcelona (two participants); the faculties of Medicine, Biology and Odontology of the University of Barcelona (10 participants). One interview took place at a participant's house and another interview was conducted by phone. A copy of consent form and the agreement for being recorded were discussed with each individual and signed prior to the beginning of the interview (appendix D, in Spanish). The oral consent was required to the participant who was interviewed over the phone before any data were collected for the study.

The duration of the interviews was between 20 – 45 minutes and the answers were audio-recorded with some note taking. The notes reported contextual observations about the attitudes of the participants and the artifacts used during their daily work. These observations were useful to detect that the participants used to perform the processes of image creation, editing and search in their personal computer, mainly under Windows (19 participants). It was also observed that they used software as Microsoft PowerPoint to perform common image editing tasks (annotation, framing, etc.). Some researchers showed the image archiving and communication systems (PACS) they used especially for working with medical images. During the interview, participants were also asked to show and comment some examples of common images they used in academic publications.

Most interviews (20 participants) where performed at the workplace of the participants, during the working time. This condition influenced the interviews and due to time constrains, the questionnaire was frequently partially covered or responses were not extensive. Occasionally the interviews were interrupted by phone calls or interruptions by students or by other researchers. As a consequence of this, the interview was mainly focused on the questions addressed to participants as authors. The open-ended questionnaire provided high flexibility to the interview. Probably due to the broad spectrum of the questions covering different fields of research and the lack of specific knowledge in medicine and biology of the interviewer, some participants showed a tendency to describe their work in general terms, probably omitting relevant information. In other occasions, when the information was very technical, time constrains prevented deeper explanations of specific issues.

2.3.2.4 Analysis of the interviews' data

The Framework Analysis method (often called "thematic analysis") was chosen for analyzing the qualitative data collected in the interview. It is considered the most appropriate to study specific information needs and provide outcomes or recommendations (Lacey & Luff, 2001). It allows for the inclusion of a priori concepts as well as emergent concepts, for example in coding (also called "indexing"). Another benefit is that it provides a systematic approach to the analysis process, making clear the stages by which the results have been obtained from the data (Lacey & Luff, 2001).The first step of the process was to transcribe the recordings of the interview, read the interview transcriptions several times and assign a code to each interview (familiarization). The following step was to identify a thematic framework based on both predefined issues (as defined in Table IV) and themes emerging from the familiarization stage. In this step, an extra category related to general practices, difficulties and suggestions in the current image submission process was added to the other five categories of analysis. This and the following steps of the analysis's process were influenced by the answers of the interview's participants, which oriented the author on the cause of the issues and how to solve them.

Once the thematic framework was defined, the process of "coding" was performed by using the qualitative data analysis software NVivo (v.10): specific units of data (e.g. sentences or paragraphs) of the transcriptions were identified and assigned to different themes according to the thematic framework. An example of this kind of data are the answers related to the theme of the image needs in academic articles, as in the following passage extracted from the interview with a participant:

> "*La función [de la imagen] es demostrar y corroborar lo que estoy diciendo es cierto y real*" ("The function of the image is to demonstrate and confirm that what I assess is true and authentic").

During the coding process, specific subcategories inside the general thematic framework emerged, as showed in Table V.

Table V Thematic framework Codes assigned to units of information in the coding process, their related research question and specific issue.

Code name	General topic	Specific Issue
access_img	Reading	barriers to reading images
accessibility_img	Reading	barriers to reading images
element_img	Image use	features of the image
legend_img	Related text needs and use	features of the image
statistical_signif	Related text needs and use	features of the image
graphical_abstract	Image needs	graphical abstract
color_use	Image use	graphical issues
contrast	Image use	graphical issues
resolution_img	Image use	graphical issues
bad_example	Reading	graphical issues
caracter_img	Image needs	image definition
definition_img	Image needs	image definition
designer	Image use	image editing
editing_img	Image use	image editing
tool	Image needs and use	image editing
format	Image needs and use	image format
how_select_img	Image use	image function
why_use_img	Image needs	image function
position_in_article_img	Reading	image reading
reading_img	Reading	image reading

Code name	General topic	Specific Issue
reading_paper	Reading	image reading
reviewer_practices	Reading	image reading
repository_img	Image needs and use	image storage
caption	Related text needs and use	image text description
diagnose	Related text use	image text description
title_img	Related text use	image text description
description_img	Related text use	image text description
metadata_img	Related text needs and use	image text description
mention	Related text needs	image text description
tables	Related text use	image text description
diagram_method_img	Image needs	image type
graphics_img	Image needs	image type
illustration_img	Image needs	image type
microscope_img	Image needs	image type
photo_img	Image needs	image type
typology_img	Image needs	image type
western_blot_and_gel	Image needs	image type
clinical_prac_img	Image use	image use for specific profile
technician_pract	Image use	image use for specific profile
book_edition	Image use	image use in specific case
presentation	Image use	image use in specific case
supplementary_material	Image use	image use in specific case
video	Image use	image use in specific case
copyright	Other topics	publishing practices
editorial_guidelines	Image use - General submission process	publishing practices
publishing_practice	General submission process	publishing practices
submission_problem	General submission process	publishing practices
suggestions	General submission process	publishing practices

Once defined the codes, the headings from the thematic framework were used to identify all the units of analysis for each theme, which were gathered together in a Word document. An example of these documents is provided in Appendix E, where all the coded information related to captions of images was collected. The last step was to search for patterns and associations within the data organized into categories and to create concepts and explanations. For example, from the coded information gathered in the category concerning the use of images in academic articles, a set of image selection criteria was identified and ordered.

Due to limitations in resources, data collection by qualitative techniques was restricted to a limited number of researchers in medicine and biology in universities and research institutes in Barcelona (Spain). However, the total number of the participants of the survey (198) and the number of participants in the interview (22) was considered enough according to the scope of the study: an approach to the image authoring chain from the point of view of the accessibility that, to the knowledge of the author of the thesis, has not attempted before in Spain.

2.3.2.5 Limitations of the interviews

The limitations concerning this study are the same as those of the survey study, as described in section 2.1.3.4. They differ for the fact that the participants of the interview were not exclusively from the University of Barcelona, but also from other research institutes. However, they belonged to a localized territory with specific conditions. For this reason, the scope of this study is limited to this geographical area and its results can be considered reliable only for it. Further research is required in order to extrapolate the results of the analysis to other countries or inclusive to other academic institutions in Spain.

2.4 Summary of the research methods

In this chapter, the quantitative and qualitative research methods applied in this thesis have been presented. They are summed up as follows:

- A literature review of the current solutions for making images accessible to readers with disabilities from a multidisciplinary perspective. This review included a previous definition of the target users who benefit of the improvement of image accessibility in STEM academic articles.
- An audit of policies and practices in the processes of submission and delivery of images in biomedical articles and in other STEM articles (computer science and mathematics). The analysis was applied to 32 selected journals.
- An analysis of the technical characteristics of figures in 30 selected STEM articles was manually performed. A specific lexical analysis (qualitative and quantitative) of captions and mentions was also performed with 9 selected articles.
- An online survey with 198 academic researchers in medicine and biology of the University of Barcelona. The survey aimed at assessing the authors' practices, needs and attitudes concerning the use and functions of images in academic articles.
- Face-to-face semi-structured interviews with 22 academic researchers in medicine and biology from different institutions and research centers in Barcelona, in order to deepen

into the most interesting aspects arisen from the survey and gain a qualitative perspective of motivations and practices in the creation, editing and submission of images in academic articles.

Chapter 3

Definitions and background

3 Definitions and background

This chapter is an introduction to key concepts related to the images commonly used in biomedical academic publications and in other STEM fields, such as computer science and mathematics.

Firstly, general terminology used to define different types of static visual information according to their content is presented, in order to clarify the characteristics of the different digital images used in STEM scientific publishing in general and in academic biomedical publications in particular. After that, the basic technical characteristics of digital images used in academic publications are reviewed, emphasizing the aspects related to the accessibility.

Secondly, the different types and formats of images and documents used in academic publications are described, paying special emphasis on their support and possibilities for accessibility.

Lastly, the textual information related to images in STEM academic publications (caption, mention(s), text inside the image and metadata) are presented, as they are an important element of the images and they have a specific role in process of authoring and publishing image.

3.1 Premise: definition of terms

The terms that will be used along this study are defined in this section, especially those related to the types and characteristics of the image in general and STEM images in particular.

In this study the term "image" is used in the broad sense of any two-dimensional figure, such as a picture, a diagram, a map, a graph, a chart or a painting. We use the term "visual representation" as a synonymous of image. An image may be captured by optical devices (such a camera, microscope, etc.) or by natural objects and phenomena (such as the human eye or water surface). It may be created manually (such as drawing and painting) or rendered automatically by printing or computer graphics technology. The focus of this study is on images created automatically, in particular those generated by a computer. The concept of "image" is wide and includes different subgroups. The terminology associated with the concepts of these subgroups is defined below.

3.1.1 Graphics

In this thesis we assume the definition of graphics as a synonymous of "graphic representation", according to the Engelhardt definition (Engelhardt, 2002): "A graphic representation is a visible artifact on a more or less flat surface that was created in order to express information and with the intention to describe, explain, inform or instruct". This definition of "graphics" excludes as primary goals to amuse, delight, persuade, invigorate, provoke or otherwise stimulate. This study is constrained to the objectives posed by Engelhart definition in order to avoid the additional complexity of dealing with art and advertising.

Examples of graphics are photographs, drawings, charts, diagrams and numbers (e.g. math equations). Graphics often combine text, illustration and color.

3.1.1.1 Diagram

The term “diagram” indicates a two-dimensional visual representation that represents data or information in a simplified and structured way according to some visualization technique. In other words, a diagram is an abstract, simplified and structured visual representation of real world facts, data, concepts, constructions, relations and ideas used to visualize and clarify the topic. For instance, a diagram could be a chart, a flowchart, a schematic illustration of a natural phenomenon, a map or a formal diagram, but not an historical photograph.

A chart (also known as “graph”) is a particular type of diagram displaying numerical data. It represents tabular, numerical data in a simple way, often using lines and curves to show relationship between large quantities of data. Statistical graphics as histograms, bar charts, pie charts, line charts, box plot and scatter plots are among the most common types of charts.

A flowchart is a symbolic representation of sequence of activities.

A schematic illustration is a schematic diagram combining diagrammatic information conveyed by abstract and graphic symbols with schematic drawings reproducing real pictures (for example, human organs).

A map is a visual representation of some space in which relationships between elements such as objects, regions, and themes are highlighted. The geographic map is the typical example of a map considered as a visual representation of an area. A treemap is a particular case of a non-spatial map, which displays logical relationships between items, but no spatial relationships.

*A “*formal diagram” is a diagram with a well-defined data-model behind their visual structure, such as tree structures, Venn diagrams, cause-and-effect diagrams, and Unified Modeling Language (UML), a standard language for visualizing, specifying, constructing and documenting the artifacts of a software system) diagrams.

3.1.1.2 Photograph

A photograph (or “photo”) is an image created by light falling on a light-sensitive surface. Currently, a common type of photo is an optical image obtained from a camera.

3.1.2 STEM graphics

A “STEM graphic” presents Science, Technology, Engineering and Math (STEM) information visually in form of charts, diagrams (for example, illustrating a biological process as the carbon cycle or a chemical, molecular or cellular structure), math equations and tables (NCAM, 2009b). In this study the term includes also biomedical images, described below.

3.1.2.1 Biomedical images

Biomedical images include several types of images captured by different devices according to different techniques. They can be classified according to the acquisition modality (Müller et al., 2012).

Radiology imaging

- X-Ray, 2D Radiography (Figure 12, panel A). It is the most widely used technique to measure the density of X-ray absorption and produces flat x-ray images. It is mainly used in the detection of pathology of the skeletal system. In some cases it is useful for the detection of disease processes in soft tissue.
- X-Ray Computerized Tomography (often referred as CT) (Figure 12, panel B). It measures the density of X-ray absorption and produce tomographic images (virtual 'slices') of specific areas of the scanned object taken in different directions. These cross-sectional images can be combined to generate a three-dimensional image of the inside of the body and used for diagnostic and therapeutic purposes in various medical disciplines.
- Magnetic Resonance Imaging (MRI) (Figure 12, panel C). It uses magnetic fields to polarise and excite hydrogen nuclei (single proton) in water molecules in human tissue, which produce a “resonance” that is detected as a spatially encoded signal, resulting in images of the body. MRI traditionally creates a two dimensional image of a thin "slice" of the body (such as CT does), however modern MRI instruments are capable of producing images in the form of 3D blocks. It is used for the description of organs and soft tissue.
- Positron Emission Tomography (PET) (Figure 12, panel D). This modality is included in the category of nuclear medicine techniques, which measure the density of radioactive molecules injected into the body to diagnose or treat various pathologies. PET technique detects the concentration of gamma rays emitted by the radioactive molecules introduced into the body and constructs a three-dimensional image of functional processes in the body.
- Ultrasonography (Figure 12, panel E). It measures variations in the speed of high-frequency sound waves propagation reflected by tissue to produce (up to 3D) images. It is often used to obtain information about moving organs and includes obstetric sonography and echocardiography.

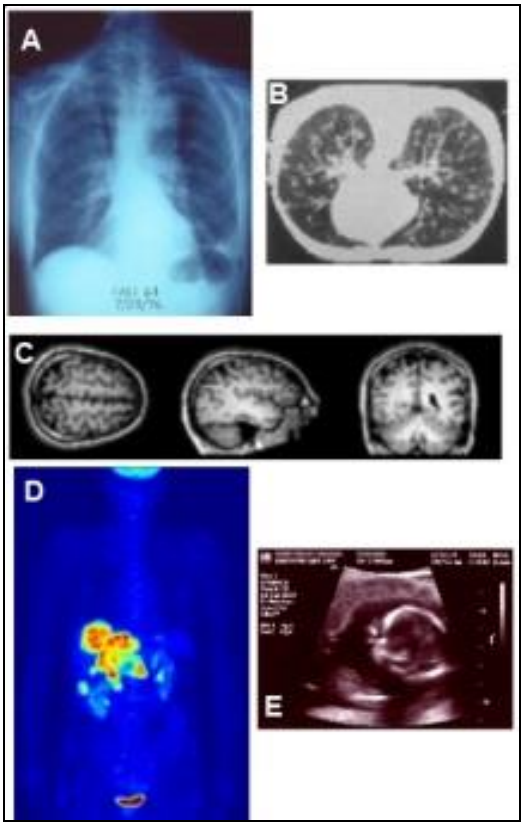


Figure 12 Examples of different image modalities in radiology imaging: X-ray, 2D Radiography (panel A) (Source: (Partin, n. d.); CT (panel B) (source: (Public Health Image Library, n. d.); MRI (panel C) (source: (Wikipedia, 2013); PET (panel D) (source: (Maus, 2012)); Obstetric ultrasonography (panel E) (source:(Kemp, 2007)).

Other types of biomedical images

- Printed signals: a standard graphical representation of signals and waves emitted by the body. Examples of printed signals are electrocardiograms (ECG) (Figure 13, panel A).
- Visible light photography: optical images obtained from camera (Figure 13, panel B). Micrographs are not included in this definition (see below).
- Micrograph: images obtained by means of a microscope, representing cells, cell components, tissues, organs and species (Figure 13, panel C).
- Electrophoretic Gel and blot: a visual map of the distribution of molecules of DNA, RNA or proteins of a tissue on a sheet of nylon or nitrocellulose. They have different modalities, such as Northern blot (for DNA), Southern blot (for RNA) and Western Blot (for protein) (Figure 13, panel D).
- Multipanel image: combination of two or more figures in the same image. (Figure 13, panel E).

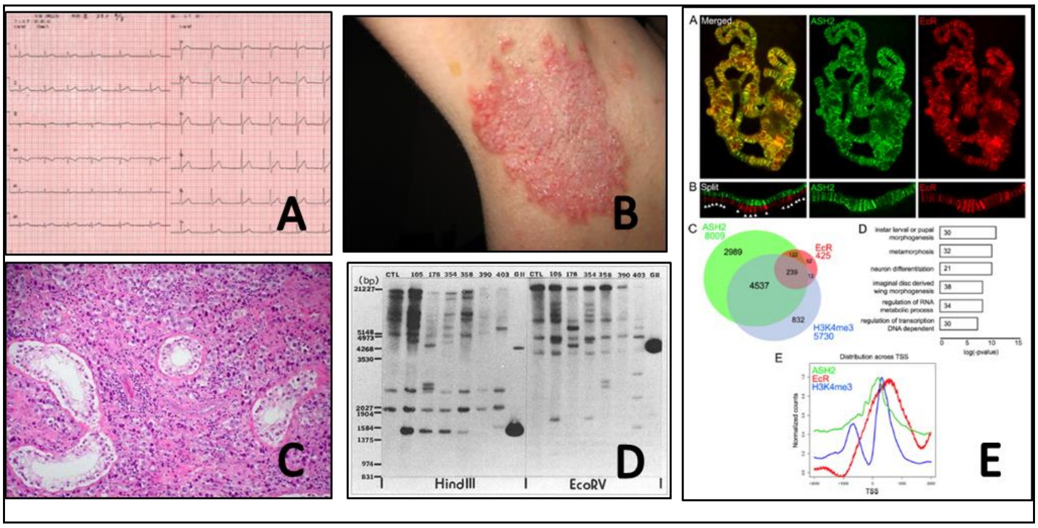


Figure 13 Examples of different image modalities in biomedical imaging different from radiological images: ECG (panel A) (source: (Wikipedia, 2015)); optical image used in medicine (panel B) (source: (Julesmcn, 2006)); micrograph (panel C) (source: (Uthman, 2006)); Western blot (panel D) (source: (IRRI photos, 2007)); multipanel image (panel E) (source: (Carbonell et al., 2013)).

3.2 Characteristics of digital images in academic publications

Authors submitting online manuscripts are required to understand the basic features of the images, in order to be able to adjust their features to meet publication requirements (LaBerge & Andriole, 2003a). These technical characteristics determine the quality of the image and they influence the perception of the visual information, especially to people with visual impairments. Therefore, this section introduces basic concepts related to the technical characteristics of the images commonly used in academic publications and their consequences on accessibility.

3.2.1 Raster vs vector graphics

A digital image is a computer-generated "snapshot" taken of a scene or scanned from documents. The computer represents images in a numerical system, according to two common techniques that generates two types of images, raster graphics (also known as bitmap graphics) and vector graphics.

3.2.1.1 Raster graphics

Raster graphics are images composed of a grid of a fixed number of pixels (dots or picture elements). Each pixel is assigned a tonal value (black, white, shades of gray or color) represented in binary digits (zeros and ones) that are stored in a sequence by a computer and often reduced to a mathematical representation (compressed). This stream of data is then interpreted and reassembled by the computer and transmitted to an output device such as a monitor, video projector or printer. Raster graphics are commonly used for photographs, scanned images and continuous-tone data plots. They cannot be enlarged without losing resolution or image quality.

The key characteristics of a raster image are:

- Digital image dimensions
  It refers to the horizontal and vertical measurements of an image expressed in pixels. It can be expressed as the total number of pixels in the image, equal to the number of columns of pixels (width) times the number of rows of pixels (height). It determines the quality of the digital image, which is the total amount of information in a digital image. Dimensions in pixels can be calculated by dividing the physical dimension of the image (see below) into the corresponding pixel dimension (or resolution, see below) against which it is aligned.

- Physical dimensions
  The physical dimensions of the displayed image, measured in inches or centimeters, equal to the dimensions in pixels divided by the resolution in pixels (see below). For example, an image file with dimensions of 2400 pixels by 3000 pixels and a spatial resolution of 300 pixels per inch has the physical dimension of 8 inches by 10 inches when printed on paper.

- Spatial resolution
  A measure of the image details expressed as pixels per unit length. It expresses the relation between the dimensions of the image in pixels and the actual physical dimensions in inches. Images with higher spatial resolution contain more discernible fine detail. The units of spatial resolution can be expressed in two ways: pixels per inch (ppi), used to express the resolution of a digital file and dots per inch (dpi), used to describe the resolution of the output display device such as a monitor or a printer.

- Bit depth (color depth)
  The number of bits of information available to define the color or gray shade of a pixel. It is expressed in bits per pixel and the greater the bit depth, the greater the number of tones (grayscale or color) that can be represented. For example, a bitonal image (in Black/White) is composed of pixels consisting of 1 bit each assigning respectively 0 for black and 1 for white. A grayscale image is composed of pixels represented by multiple bits of information, typically ranging from 2 to 8 bit or more. In a 2-bit image, the number of tones that can be assigned to each pixel is 4 (2^2). A color image is typically represented by a bit depth ranging from 8 to 24 or higher. A 24-bit image offers 16.7 million (2^{24}) color values.
  Indexed color is a technique for storing color information not directly in the image pixel data, but in a separate color map (called *palette*), an array of color elements in which every color is indexed by its position within the array. In this way, the image pixels do not contain the full specification of its color, but only its index in the *palette.* It is used for reducing computer memory and file size of images with 8-bits color depth or lower.

- Color mode

Color modes are the different methods used for numerically representing the information encoded in the color in terms of three or more coordinates. Among several color modes, the options relevant to online manuscript submission are grayscale, RGB, and CMYK.

The RGB color space uses three channels to encode information about Red, Green and Blue, the three primary colors. This mode is called additive because the primary colors are added together to achieve the final color. It is the most appropriate color mode for images created with transmitted light, such as light emitting computer screens or video projectors.

CMYK mode encodes color on four channel using cyan, magenta and yellow (secondary colors) plus black (the “K” in the acronym). It uses a subtractive approach and it is appropriate for reflective light situations, such as in printed images.

Besides the channel conveying color information, some image formats such as PNG or GIF (see below), support an extra channel called Alpha channel, which defined the transparency of each pixel in the image.

- Digital file size

  The amount of computer memory needed to store the image. It is measured in kilobytes (KB) or megabytes (MB) and is dependent on the number of pixels, the bit depth, the color mode, the file format and the amount of data compression (see below).

- Data compression

  Data compression applied to digital images is a process of reducing size of the image file by abbreviating the string of binary code in the image based on complex algorithms. Data compression can be lossless - which abbreviates the binary code without discarding any information - or lossy – which utilizes reduction strategies that take advantages of the limitations of the visual perception for discarding the least significant information of the image. The compression ratio indicates the size of the compressed file compared to that of the uncompressed file and depends on the image format. Compressed images have the advantages of requiring less storage space and less time to be displayed.

- File format

  A plethora of digital image file formats are currently available. They differ in terms of resolution, bit-depth, color capabilities and support for compression and metadata. Next section is dedicated to the detailed description of the most common digital image file formats in academic publishing.

3.2.1.2 Vector graphics

Vector graphics are images composed of many individuals’ geometrical primitives, such as lines, curves and shapes, defined by mathematical expressions. Colors in vector graphics are defined by values assigned to attributes of the basic shapes. They are commonly used for diagrams, charts and schematic drawings. Dedicated drawing and illustration software, such Adobe Illustrator and Inkscape, can create and edit vector graphics.

Since each geometrical object in a vector graphic is defined not by pixels (as in raster graphics) but by mathematical instructions, vector graphics are fully scalable without any loss of resolution. This is particularly beneficial to users with visual disabilities, since they can access images at any desired display size without any loss of quality.

Furthermore, some vector-based formats such as SVG (see below) maintain textual characters separated from the rest of the image, enabling to search in the textual content of the image. This characteristic makes text inside this type of images available to assistive technologies such as screen readers, which can read the text inside the image to people with visual disabilities. The text is also available to web search engines such as Google, which can access it for image indexing. On the other side, text in raster formats is made by pixels and, unless it is converted from pixels into recognizable characters by an Object Character Recognition (OCR) software, it is not available to assistive technologies.

3.2.2 Encapsulating formats

Encapsulating formats can contain or encapsulate raster graphics, vector graphics and text within the same file. This category includes metafiles and Page Description Languages (PDLs).

Metafiles contain lists of commands for drawing and displaying an image (commonly vector drawings) when they are executed. They allow displaying and using their contents in different computer programs and operating systems. Common metafile formats are the Windows Metafile (WMF) and Enhanced Metafile (EMF) and PICT, a metafile commonly used for encapsulating vector graphics introduced by QuickDraw on the original Apple Macintosh computer, now dropped in favor of PDF format.

Page Description Languages (PDLs) are languages used to describe information about layout, fonts and graphics to a printer or display device. PostScript (PS) is the classic type of this format, designed to provide detailed instructions to computer printers. The Encapsulate PostScript format (EPS) is based on PS, but compared to its antecessor it was developed with the exchange of graphics in mind, making the distinction between PDLs and metafiles ambiguous. In this thesis EPS is referred to as a metafile. As EPS, the Portable Document Format (PDF) is another very common metafile based on PS. Being a main format currently supporting the dissemination of academic articles, it will be reviewed in detail in section 3.4.

3.3 Image file formats commonly used in academic publishing

A file format is a standardized means of encoding the digital information in a computer file. It defines the structure and type of data stored in a file. An image file format encodes the digital information that makes up the image file. The selection of an image file format depends on its intended use, broadly for Internet use, computer display and print press application. According to the literature (Varma, 2012)(Levine, 2010)(Lee, 2014) popular graphic file formats used in academic publications in general and in particular in biomedical publications are:

- Raster graphics: Graphic Interchange Format (GIF), Joint Photographic Experts Group (JPEG), Portable Network Graphic (PNG), and Tagged Image File Format (TIFF), Photoshop Document (PSD).
- Vector graphics: Scalable Vector Graphics (SVG), Adobe Illustrator (AI)
- Metafile: Encapsulated PostScript (EPS).

SVG is a format very relevant from an accessibility standpoint and it will be described thoroughly in a specific section. Apart from the formats cited in the literature, a specific format used commonly used in the medical imaging workflow (DICOM) will be also described.

3.3.1 Raster image formats used in academic publishing

3.3.1.1 Joint Photographic Expert Group (JPEG - JFIF) – JPEG 2000

The Joint Photographic Expert Group (JPEG) is not actually a file type, but the name of a lossy compression algorithm proposed by the Joint Photographic Experts Group. It is used within the JPEG File Interchange Format (JFIF), which is colloquially called “JPEG”. It was created specifically for the storage and transmission of photographic images and it supports many color spaces, including grayscale, RGB and CMYK. It uses a very flexible lossy compression algorithm, which allows adjusting the compression ratio. Compared to other formats, JPEG allows very high levels of compression ratio (typically 10:1) with little perceptible loss in image quality. For this reason, JPEG provides the best quality and lowest file size for continuous tone images and it is a suggested format for web delivery (JISC Digital Media, 2014a). However, the lossy compression technique provokes a loss of data, which may translate into image degradation.

JPEG 2000 is a compression standard and a coding system developed by the Joint Photographic Experts Group to replace JPEG. Although JPEG 2000 can use both lossless and lossy methods of compression, compression ratios are not significantly improved compared to normal JPEG. It is not currently widely supported in web browsers and it looks unlikely that it will replace JPEG in the near future (JISC Digital Media, 2014a).

3.3.1.2 Graphic Interchange Format (GIF)

GIF is the oldest Web-based graphic file format designed for working with simple continuous-tone images such as simple diagrams, logos and banners. GIF supports lossless data compression techniques. It has the limitations of supporting only to 256 colors (or shades of gray) and to be a proprietary format. It is recommended only for displaying a web banner or logo with 8-bit or less color, greyscale or black and white bitonal images or animated images on the Web (JISC Digital Media, 2014a).

3.3.1.3 Portable Network Graphics (PNG)

PNG file is a W3C open source standard created as a patent-free replacement for GIF. It supports indexed-color, grayscale and 24-bit RGB images, lossless compression and transparency. It is very versatile format, since it can be used in either an 8-bit indexed version (as a GIF replacement) or as a 24-bit full color version with lossless compression.

PNG is not recommended for professional-quality print graphics, since it does not support non-RGB color spaces such as CMYK. It has worse performance compared to JPEG in producing high quality and small, full color images for viewing on the web (JISC Digital Media, 2014a). PNG shows best performance when the image has large, uniformly colored areas (e.g., icons) and is especially recommended on the web for displaying high quality continuous-tone full color images, high quality continuous-tone grayscale images, images or logos with transparent layers (JISC Digital Media, 2014a).

3.3.1.4 Tagged Image File Format (TIFF)

TIFF was designed for being compatible with different image processing devices, such as printers, monitors and scanners and currently is widely supported by image-manipulation applications and by publishing and page layout applications. The main strengths of TIFF are:

- to support a wide range of image sizes, resolutions and color depths;
- to support different lossless compression techniques, which allow TIFF files being edited and re-saved without losing image quality;
- to contain more than one image in each file, including compressed JPEG;
- when uncompressed, it maintains the information created by the capture device.

The main shortcoming of TIFF is the large size of the resulting file, even after the use of lossless compression techniques. Another limitation is its lack of support in some of the most common Web browsers. It is a suggested format for capture, archiving, manipulation and commercial and desktop printing, but not for web delivery (JISC Digital Media, 2014a).

3.3.1.5 Photoshop document (PSD)

PSD is the Adobe Photoshop’s native file format for storing layered image files (layering is a technique that enables authors to insert labels and other elements such as arrows on a transparent layer above the background image). PSD is recommended for image optimization and manipulation (JISC Digital Media, 2014b). Its main disadvantage is to be a proprietary format. Although PSD is primarily a raster image format, it can be considered as acting as a metafile, since it can hold vector information, such as type, in its layers.

3.3.2 Metafile formats used in academic publishing

3.3.2.1 Encapsulated PostScript (EPS)

EPS is a metafile format specifically intended to encapsulate graphics and based on PostScript. Its widespread adoption was mainly due to its ability to provide security management for preventing the encapsulated graphics to be manipulated when shared among different platform or sent to commercial printers. However, nowadays EPS has now been largely superseded by PDF (JISC Digital Media, 2014b).

3.3.3 Vector formats used in academic publishing

3.3.3.1 Adobe Illustrator (AI)

AI is the Adobe Illustrator's native file format for representing single-page vector graphics in both the EPS and PDF formats. AI format is commonly used during the editing step of the different layers of vector graphics, while it is usually exported to PDF for generating the production file. Its main disadvantage is to be a proprietary format.

3.3.4 Image formats relevant from an accessibility standpoint

Scalable vector graphics

Scalable Vector Graphics (SVG) is an XML markup language (see section 3.4.5) for describing 2D vectors and mixed vector/raster graphics and it is a W3C open standard. This format uses vector shapes made of a combination of straight lines and curves to represent scenes and can also include text and raster graphics. An example of SVG including a vector shape (a black rectangle), a raster image (a PNG file) and text ("University of Barcelona") is showed in the code below:

```
<svg xmlns=http://www.w3.org/2000/svg
xmlns:xlink="http://www.w3.org/1999/xlink">
   <title> Logo UB </title>
   <rect x="10" y="10" height="130" width="300" style="fill:
   #000000"/>
   <text x="150" y="30" font-size="14.5" font-weight="bold" text-
   anchor="middle" fill="white"> University of Barcelona </text>
   <image x="10" y="40" width="300" height="80"
   xlink:href="http://www.ub.edu/web/ub/galeries/imatges/logo_home_no
   u.png" />
</svg>
```

Being an XML document type, SVG benefits from the rendering of style separated from the content by using CSS or XSL style sheets, which allow users (including user with special needs) to control the display of SVG images according to their needs. Furthermore the text included inside SVG files can be accessed by people with disabilities using assistive technologies.

SVG content can be included within a Web page. Some of the options on how to do this are detailed in section 3.4.1. A new version of SVG (W3C, 2014j) is currently under development and it is supposed to add new ease-of-use features to SVG, as well as more closely integration with HTML and CSS.

3.3.5 Image format specifically used in biomedical workflow

3.3.5.1 Digital Imaging and Communication in Medicine (DICOM)

DICOM is a proprietary standard that was created for the distribution and viewing of medical imaging, such as CT, MRI and ultrasonography. DICOM format file is commonly referred to as the image file compliant with the Part 10 of the DICOM standard (NEMA, 2014) and it is the *de facto* standard for receiving scans from a hospital, especially in radiology (Varma, 2012). The main disadvantage of this format is its large size. For example, a set of pre and post-contrast computed tomography (CT) images of the brain can have a size of 35 megabytes (Graham, Perriss, & Scarsbrook, 2005). The conversion of DICOM file into other formats is most often performed at a diagnostic workstation or at a Web client of an archiving and communication system (PACS) with the "export" function to JPEG or TIFF file (Varma, 2012).

3.4 Document formats including images in STEM academic publishing

STEM images are commonly included in scientific articles inside documents that have specific formats. HTML and PDF are the most common formats to deliver scientific research articles in academic journals (Attwood et al., 2009). They are reviewed in this section, paying particular attention to the mechanisms they support for including figures. Mechanisms for including text alternatives in images are also introduced in this section and thoroughly described in section 4.3.

Besides these formats, other formats not widely used in scientific academic publishing are discussed for their relevance in the accessibility publishing: DAISY Digital Talking Book (DAISY), a standard format for accessible digital books, and EPUB, a new format considered the most promising format for accessible digital publishing (International Digital Publishing Forum, 2014b) and the successor of the DAISY standard.

3.4.1 HiperText Markup Language (HTML)

HTML is the standard markup language used to publish information on the Web. It was originally designed by Tim Berners-Lee at the European Organization for Nuclear Research (CERN) as a language for semantically describing scientific documents, although its general design and adaptations over the years have enabled it to be used to describe a number of other types of documents. The Internet Engineering Task Force (IETF) was responsible to develop and maintain HTML specifications until 1996, when the efforts of the W3C's HTML Working Group to codify common practices resulted in HTML 3.2. The HTML 4.01 specification was published in late 1999 and HTML 5 is a "Proposed Recommendation" at the time of writing (October 2014).

Currently most STEM journals offer an online version of their articles in HTML. Graphics are inserted into an HTML page using the <img> tag, which has two required attributes: *src* and *alt*. The *src* attribute specifies the URL of the linked image.

The *alt* attribute specifies a short textual replacement of the image, in order to make the visual content accessible to users with special needs. When the image is an image of text - an image is used to represent text, such as logotypes that combine textual information and graphic display to convey the brand name – the *alt* attribute may be used to specify the literal transcription of the text represented in the image of text. Another attribute of the <img> element relevant from the accessibility point of view is *longdesc*, which provides a link to a long text alterative of the image describing the content of the image in detail. The *alt* and *longdesc* attributes will be described in depth in the section dedicated to the mechanisms for associating text alternatives to images (section 4.3).

Besides the <img> element, there are other mechanisms for including images in HTML documents.

The <object> tag is used to render an embedded object within an HTML document, including multimedia (like video and Java applets) and images. The body of the <object> tag allows the inclusion of fallback content, content that is to be used when the resource loaded by the element is not rendered by the user agent. For example, when an image in SVG is not supported by the browser, the same image in PNG format (the fallback content) is displayed to the user. The inclusion of text alternatives as a fallback mechanism of images in the <object> element will be described in detail in section 4.3.

CSS can be also used to include images in HTML documents. In particular, this mechanism is recommended for adding purely decorative images (see section 4.2), by using CSS properties (*background, background-image, content, list-style-image*) and without requiring additional markup within the content.

In the development of HTML 5, the <figure> element is another mechanism available for encapsulating images (illustrations, diagrams, photos, code snippets) in a HTML document. Another new element, the <figcaption> element is used to explicitly associate a caption to the image included in the body of the <figure> element. <figcaption> will be described more in detail in the section dedicated to the mechanisms for including text alternatives to images.

Several mechanisms allow SVG files being embedded in HTML documents. In addition to the mechanisms previously described, HTML5 have proposed the <svg> tag, a container for SVG graphics. Other tags that may be used to embed SVG files are <iframe> (an inline frame used to embed another document in HTML 4.01 and HTML5 documents) and <embed> (a container for an external application or a plug-in in HTML5 documents).

3.4.2 Portable Document Format (PDF)

PDF is the native format of the Adobe Acrobat family of products oriented towards the presentation and exchange of electronic documents independently of the application software, hardware and operating system. It is the descendent of the PostScript page description language and relies on the same imaging model to describe text and graphics in a device-independent and resolution-independent manner. PDF includes objects such as annotations, hypertext links, Javascript and multimedia, which are not part of the page itself but are useful for interactive viewing and document interchange (Adobe Systems Incorporated, 2006). The first version of the format was created at the beginning of 90s by Adobe, a leader in the market of publishing software. The fidelity of the format in rendering and printing a file exactly as an author designed it and the strategy of Adobe of offering the software reader under a free license have been decisive for gaining world-wide acceptance.

Since 2008 PDF is an open standard (ISO, 2013). It is also the "umbrella" standard of special-purpose PDF standards, including PDF/A for archiving, PDF/E for engineering, PDF/X or PDF/VT for printing.

Particularly relevant for the accessibility discipline, the PDF variant PDF/UA-1 standard ISO 14289-1 (ISO, 2012) provides a set of guidelines for creating PDF files that are universally accessible. It is addressed to people with disabilities, IT managers in government or commercial enterprises and compliance managers.

The PDF document consists in a sequence of pages composed by any combination of text (referred as text objects), graphics (path objects) and images (external objects) (Whitington, 2011):

- a text object consists in one or more characters and layout information (position, fonts).
- a path object contains vector instructions within the content stream of the PDF document. This integration makes difficult the extraction of vector images as independent objects;
- an image external object (XObject) defines a rectangular array of sample values, each representing a color. Raster graphics are included in a PDF as objects independent of the content stream and can be extracted from it with their original composition. The image model in PDF is very similar to JPEG compression format, and this is the only standard image format that can be placed into a PDF without any modification. For this reason, JPEG images are the only ones that could be extracted from a PDF in its original form.

PDF offers its own mechanisms for associating descriptive texts to images. Text alternatives may be conveyed by the *alternate* and *actual* attributes of the <figure> tag, the tag used for identifying images that provides relevant content within a PDF document. The *alternate* attribute provides a generic text alternative to the image, without any reference to the length of the text included in it. The *actual* attribute conveys the transcription of an image of text. The caption of an image may be tagged by the <caption> element and semantically associated to the figure it refers to.

Due to the relevance of the issue in this thesis, the relations between the accessibility features of PDF, the inclusion of images in this format and their associated text will be described more in detail in the section dedicated to the mechanisms for associating text alternatives to images (section 4.3).

3.4.3 DAISY Digital Talking Book

The DAISY (Digital Accessible Information System) is an open technical standard for structured and accessible digital books developed by the DAISY Consortium. It has been considered for a long time the *de facto* accessible content standard for XML-based e-books, although EPUB3 format (see section 3.4.4) has inherited the accessibility features of the DAISY system and it is currently considered its successor (Garrish, 2014).

DAISY specification is made up of rules and requirements necessary to create a variety of Digital Talking Books (DTBs). DTBs are a multimedia representation of a print publication that enables the access to the text through digitally recorded human voice or synthetic text-to-speech technology. The main aim of DTBs is to provide access to digital books to blind and visually impaired users, but DTBs can also help to improve reading and comprehension skills in readers with learning disabilities (NCAM, 2009a). They overcome most of the barriers usually experienced by users with print impairments during the reading: the content of the books is presented with enabled navigation within a sequential and hierarchical structure consisting of (marked-up) text synchronized with audio. Through this system, users can access text and images synchronized with audio and can navigate documents sentence by sentence, search text, place bookmarks and control the reading speed without distortion.

Currently, “DAISY” refers to two related NISO standards maintained by the DAISY Consortium:

- ANSI/NISO Z39.86-2005 (R2012) (NISO, 2005), Specifications for the Digital Talking Book [DAISY 3], which defines the format and content of the electronic file set that comprises a DTB and establishes a limited set of requirements for DTB playback devices. It defines an XML grammar (originally called DTBook, more recently also DAISY XML) that is intended to represent the textual content of a DAISY DTB. DTbook was designed as a format for creating content (*authoring format*).
- ANSI/NISO Z39.98-2012 (NISO, 2012), Authoring and Interchange Framework for Adaptive XML Publishing Specification (known as *ZedAI*), which defines an XML-based framework with which content producers can represent different kinds of information resources (books, periodicals, etc.), with the intent of producing documents suitable for transformation into multiple universally accessible formats (e-text, Braille, large print, and EPUB, among other formats).

Apart from these standards, the National Instructional Materials Accessibility Standard (NIMAS) 1.1 Technical Specification is a simplified version of the DAISY specification adapted for the creation of print instructional materials in accessible formats to blind or other persons with print disabilities in elementary and secondary schools in the United States (National Center on Accessible Instructional Materials, 2006).

In addition to text and audio files, a DTB can include images that can be presented on players with visual displays. Images included in DTBs must be presented in JPEG (JFIF V 1.02), PNG or SVG (the preferred format in the NIMAS technical specification).

When a DTB contains figures, these are included in the set of source files that form a DTB file. They are referenced from the <img> tag in the file containing the text of the book with appropriate markup, which can be an XHTML file or a “DTBook” XML file (XML and XHTML are described below in this section). The <imggroup> element provides a container for one or more <img>s and associated textual descriptions to images in specific elements. The *alt* attribute is required for all <img>s (DAISY Consortium, 2008).

Currently DAISY consortium has driven the incorporation of DAISY accessibility requirements into the EPUB 3 standard and has advocated for the use of EPUB as the most promising format for accessible publishing, replacing DAISY DTB. For this reason, the DAISY DTB format is currently not used in the digital academic and consequently it is beyond of the scope of the thesis. It has been presented in this section for its relevance from an accessibility point of view.

3.4.4 EPUB

The EPUB specification is a distribution and interchange format standard for digital publications and documents developed by the International Digital Publishing Forum (IDPF) (whose members are companies as Google, Kobo and Adobe and publishers as Hachette, Pearson and Santillana, among others). It is a free and open standard that defines a mean of representing, packaging and encoding structured and semantically enhanced Web content — including HTML5, CSS, SVG, images, and other resources — for distribution in a single-file format (International Digital Publishing Forum, 2014b). It has been endorsed by the International Publishers Association (IPA) as a global standard for digital publishing (International Publishers Association, 2013).

As noted previously, EPUB3 specification is considered the successor of the DAISY standard. It has embraced all accessibility features offered by the DAISY standard and enables features useful for improving the reading by persons with print disabilities (International Digital Publishing Forum, 2014b). The ability of this format to provide full accessible documents to print disabled readers is one of the reasons of its endorsement among several institutions and organizations. The Advisory Commission on Accessible Instructional Materials in Postsecondary Education for Students with Disabilities has recognized EPUB3 as a promising format particularly suitable for post-secondary educational publications (AIM Commission, 2013). A project involving all the stakeholders in publishing industry and the "EPUB ecosystem" (publishers, reading system developers, retailers, service providers and the accessibility community) and other organizations working on the implementation of EPUB3 (such as IDPF, the Readium Foundation, the Book Industry Study Group (BISG), Benetech, the American Printing House for the Blind, the National Federation of the Blind and the DAISY Consortium) is ongoing to define the priorities for the implementation of EPUB3 format in the publishing workflow, including accessibility requirements (Association of American Publishers, 2013).

An EPUB3 document consists of a set of interrelated resources and includes:

- The content documents (HTML5 files, images, CSS files, etc.).
- the root file (.opf) called package document containing information about the publication, a list of all resources used in the EPUB file and their reading order.
- Other files that provide information to the reading systems about the type of format and the location of the package document.

Figures included in EPUB3 documents are marked in the HTML5 file code using the HTML5 elements, which are the same as those reviewed in the section dedicated to HTML documents (<img>, <object>, <figure>,<figcaption>, etc.). A detailed explanation of how to include image descriptions in EPUB3 will be provided in section 4.3.5.

3.4.5 XML

Extensible Markup Language (XML) is a markup language that enables the creation of documents composed of structured data. XML is a W3C standard which supplies a grammar and syntax for representing structured data in text-based form (W3C, 2010). An XML document contains structured information organized into “elements”. Each element is composed of a tag, a label used to mark up the beginning and end of an element and the content of the tag. Compared to another markup language as HTML, while the tags used in HTML are primarily focused on *presentation* (e.g.<H1>Header1</H1>), XML tags are used to indicate *meaning* (e.g. <organisationName>University of Barcelona</organisationName>). XML attributes are used to describe an element or to provide additional information about an element and they are contained within the starting tag of the element.

The rules about the structure and syntax of an XML document are defined in a Document Type Definition (DTD) file. Several DTDs have been developed for defining different types of information: SVG for vector graphics, XHTML for web documents, MathML for mathematics, etc. MathML will be described below for its relevance in rendering accessible mathematical formula and notations, another type of information usually presented a graphic.

3.4.6 MathML

MathML is a markup language used to render mathematical formulas and notations, otherwise depicted as graphics in many digital documents. Mathematical symbols on the Web are not accessible when presented simply as graphics, sometimes even when the mathematical expressions are represented using text. Mathematical Mark-up Language (MathML) is an XML format that encodes the structure and the content of mathematical notations in a way they can be displayed, manipulated and shared in browsers and accessed by screen readers. It is a recommendation of the W3C math working group and part of the HTML5 the specification. The second edition of the current version of MathML (3.0) has been recently released (April 2014).

Since MathML preserves the meaning of the formula components separated from the presentation, the content of an equation written in MathML in a Web page can be rendered normally by a Web browser or be read aloud by screen readers. MathML is referred to as the most effective method of improving the accessibility of mathematical and scientific content in core instructional materials to students who are blind or who have print disabilities (United States Department of Education, 2012). However, in some cases of complex formulae, MathML can be a successful alternative for the screen reader only with the support of mathematical reading rules.

3.5 Image-related texts in academic publishing

Images in academic publications can be associated to different kinds of related text, such as figure captions and mentions (text within the article that refers to the figures), the text labels inside images and metadata. In this section, the role, characteristics and use of these image-related texts in academic publication is described.

3.5.1 Caption

A figure caption is “a title, brief explanation, or comment accompanying an illustration” (Collins, 2015). Sometimes the caption can contain the figure title and serves as both as an explanation of the figure and as a figure title (American Psychological Association, 2010). A figure caption is referred by many authors as “legend”. In this study, the use of the term “caption” is preferred, due to the possible confusion with the legend element that appears within the illustration itself and identifies the symbols used in a map or a chart (University of Chicago Press, 2006)(Figure 14).

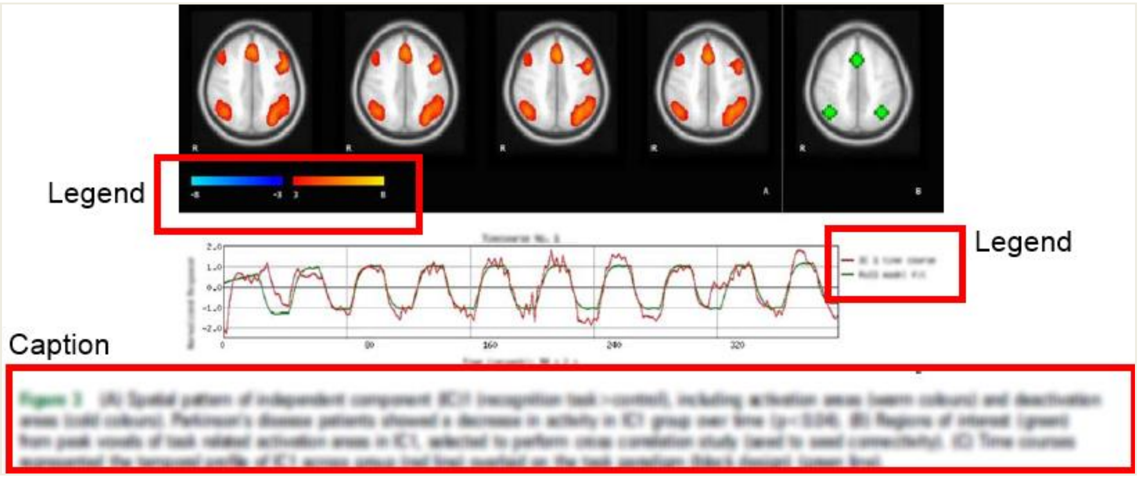


Figure 14 Sample figure showing textual elements related to the figure. Original graphic from: (Segura et al., 2013) labelled by the author.

The role of the caption as an important component of figures has been emphasized by several authors. Captions are necessary for understanding the content of a figure (Agarwal & Yu, 2009) (Yu et al., 2009), often concisely summarize a paper’s most important results as perceived by the author (Cohen, Wang, & Murphy, 2003), provide lexical cues that efficiently represent image content (Rafkind et al., 2006) and contribute to understanding the graphic’s message (Elzer et al., 2008). They are especially useful for locating information about experimental results (Divoli, Wooldridge, & Hearst, 2010).

According to the most widespread used manuals of styles for scientific academic literature (University of Chicago Press, 2006),(American Psychological Association, 2010), (American Chemical Society, 2006), a caption of an image has the following characteristics:

- Function: Its common function is to complement the figures by an informative, descriptive and explanatory textual description.
- Length: It can be a brief full sentence or can consist of two or more sentences or of a title followed by one or more full sentences.
- Position: It appears immediately below an illustration, or sometimes above it or to the side.

Other common manuals such as the style guide of the National Library of Medicine of United States and the IEEE style manual, focused on Electric Engineering and Computer Sciences literature, do not offer any reference to the definition of the main characteristics and use of the caption.

Descriptions of the content in caption were found in the previously cited manuals and in other sources selected for their relevance in the academic publishing in general and in the academic biomedical publishing in particular, according to the scope of the thesis:

- General guidelines on article submission for authors offered by the three most important international publishers, which publish almost half of the articles produced by the academic journal publishing industry (Campbell, Pentz, & Borthwick, 2012): Elsevier (Elsevier Publishing, 2012); Wiley-Blackwell (Wiley-Blackwell Publishing, 2015) and Springer (Springer Publishing, n. d.).
- Guidelines provided by other academic publishers which journals have the best impact factor according to the ranking of the Journal Citation Reports (JCR 2013): Nature (Nature Publishing Group, 2014), New England Journal of Medicine (Massachusetts Medical Society, 2015), Science (American Association for the Advancement of Science, 2015).
- Guidelines provided by the ICMJE (International Committee of Medical Journal Editors, 2014), which are well established guidelines for the preparation of the component of an article that standardize the preparation and formatting of manuscripts (Aydıngöz, 2005).

The features identified that are recommended to be included in the caption are the following:

- Labels: identification and explanation of labels, including symbols, letters, arrows and numbers used to identify in the figure.
- Abbreviations: explanations of abbreviations.
- Modifications: descriptions of modifications applied to the image, such as selective digital adjustments and enhancements (cropping, brightness, color, sharpness, etc.).
- Source/permission: specification of the source of the image, appropriate credits and permissions.
- Unit of measurement: specification of the units of measurement of variables (for axis or scale bars).
- Statistical analysis details: explanation of the details of results of statistical analysis showed in the figure (standard deviation, p value, etc.).
- Area of interest: explanations of key features marked in the image.
- Magnification rate: specification of the magnification rate.
- Captioning modality: specification of the acquisition modality (e.g. x-ray, MR, CT, etc.).
- Clinical information: succinct presentation of relevant clinical information (patient's history, relevant physical and laboratory findings, clinical course, response to treatment, etc.).
- Cut type (plane): specification of the cut type of cross- sectional images (e.g., a CT axial, coronal or sagittal plane).
- Reconstruction type: specification of the reconstruction type (e.g., a rendering 3D or a mip of a CT) and other relevant technical information (coordinate space in which an image has been normalized).

- Method of staining: specification of the staining method used to enhance contrast in micrographics, the preparation of the sample and other relevant information about the microscope.
- Lanes in electrophoretic gels and blots: specification of any relevant information about lanes in electrophoretic gels and blots (modifications applied as cropping, quantitative comparisons between samples, etc.).

These features are synthetized and associated to specific types of images in Table VI.

Table VI Features in captions recommended by the guidelines reviewed: Chicago Manual of Style (ACM), American Psychological Association manual of style (APS); American Chemical Society (ACS); Wiley Blackwell (Wiley); New England Journal of Medicine (NEJM).

Figure element described in the caption	Apply to	CMS	APS	ACS	Elsevier	Wiley	Springer	ICMJE	NEJM	Science	Nature
Identify/Explain labels	all type of images	x	x	x			x	x	x	x	
Abbreviations		x	x	x						x	
Modifications							x		x		
Source / permissions		x	x			x	x				
Unit of measurement of the axis	statistical graphics	x	x							x	
Statistical analysis details			x					x			
Areas of interest	other figures in Medicine and Biology (photos, etc.)		x				x	x			
Magnification rate (scale)		x						x		x	
Captioning Modality			x								
Clinical information									x		
Cut type (plane)	other figures in		x								

Figure element described in the caption	Apply to	CMS	APS	ACS	Elsevier	Wiley	Springer	ICMJE	NEJM	Science	Nature
Reconstruction type	Medicine (photos, etc.)		x								
Method of staining	other figures in Biology (photos, etc.)						x	x			
Lanes in electrophoretic gels and blots							x			x	

3.5.2 Mentions

Mentions are text reference(s) to figures in the main text of an article (University of Chicago Press, 2006). They normally bear numbers in order to locate the figures they refer to (e.g., “Figure 5 shows…,” or “see fig. 3.2”).

After reviewing the sources cited in the previous section, no specific suggestion for the content of captions was found, except for a short recommendation in the Springer guidelines to provide complementary explanations in the main text about key elements of the images.

Mentions should be distinguished from callouts, instructions that indicate where a table or an illustration is to appear during the typesetting of production of an article. Callouts do not appear in the published articles and will not be studied here.

3.5.3 Text labels and text inside the image

In academic publications, text labels are commonly adopted means that allow an author to label, point to or indicate some relevant features of the image (Goede et al., 2004). They consist of textual labels placed over the image to depict an annotation, such as text and letters. Figure 15 shows an example of textual labels and other figure text appearing inside the image, along with visual annotations. Visual annotations can be pointers (line, edge, arrow), circles and polygons used to highlight regions of interest (Antani et al., 2008).

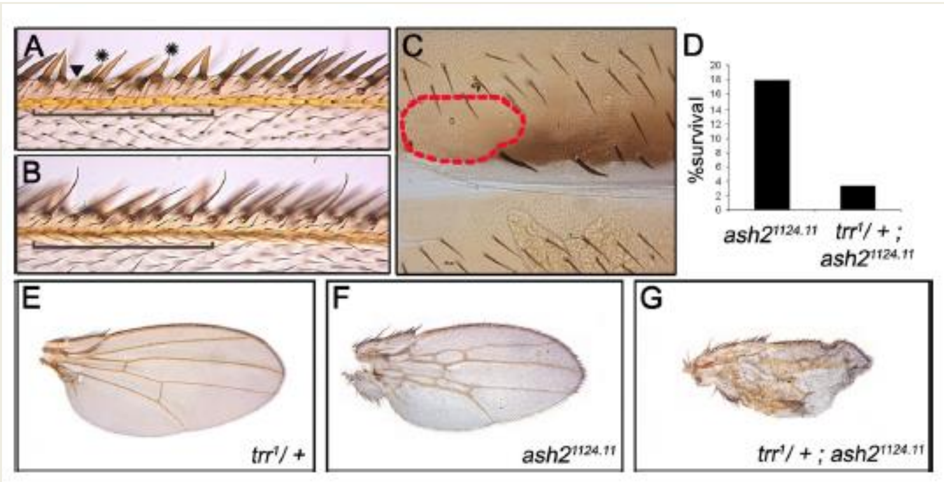


Figure 15 Example of figure with textual labels (letters A-G) and other text (“%survival”, “trr”, “ash2”, etc.) and a visual annotation (a red ellipsoidal-shaped annotation).Extracted from: (Carbonell et al., 2013).

3.5.4 Metadata

The term *metadata* (literally “data about data”) in a digital context usually refers to structured textual information that describes the creation, content or context of digital resources (JISC Digital Media, 2014c). Metadata may be used to organize, describe, find, access, index, collocate, structure, navigate, preserve and manage the resources they describe.

They can be categorized according to the information they describe and their purpose:

- Descriptive metadata describe the data content and are typically used for indexing and retrieve digital resources. In image files, they describe the content by different properties, such as caption, title, author, keywords and capture's location.
- Administrative metadata are used for managing and administration digital objects. They can be distinguished into technical metadata and preservation metadata. Technical metadata are used to record the technical attributes of digital objects. Properties describing technical information of images are commonly related to the characteristics of raster and vector images described in section 3.2.1, such as file format and size, spatial resolution, dimensions, bit depth and color mode. Other technical information concerns the acquisition parameters (such as the date of acquisition and the brand, name, and model number of the source), since technical metadata can be generated by most modern image-capture devices, such as digital cameras and digital microscopes. Image processing software (or other authoring tools) can also store metadata information into the image file. Preservation metadata provide information related to the preservation management of information resources, such as licensing or rights usage term.
- Structural metadata provide information about the internal organization of the digital resource and the relationships with other resources. They can contain, for example, information about different versions of an image file prepared for different screen sizes.

Metadata automatically generated by the capture device or the processing software are called "intrinsic" or "implicit", while metadata manually added and edited by users are called "extrinsic" or "explicit".

Concerning visual resources in general and still images in particular, metadata offer a selective or simplified description of the image usually structured according to some generic categories (also called *fields*) (e.g. Creator, Title, Subject). The full set of categories is the standardized structure of metadata called metadata "schema". An example of a common metadata schema is Dublin Core (DC), a schema maintained by the Dublin Core Metadata Initiative (DCMI)(Dublin Core Metadata Initiative, 2014) intended to describe any type of resource, including Web-based resources. It includes a set of fifteen elements (title, creator, subject, etc.) with optional qualifiers, used to structure descriptive records and facilitate information sharing.

Metadata can be kept within a digital file in different ways: within the digital file (internal or *embedded* metadata), within a database or in a separate XML-encoded file (external metadata) or all of the above at once.

A variety of *image metadata formats (*standard protocols of encoding) are used to store metadata into digital image. Most widely used metadata formats embedded in image files are (Metadata Working Group, 2010):

- The Exchangeable Image File Format (EXIF) is a standard widely used to store image data captured by digital cameras and other systems (such as scanners) and to enable data compatibility and exchangeability. It defines a set of tags that describes photographic images, including the camera model, its serial number, the date and time of image capture, the shutter speed, the aperture, the lens used and the ISO speed setting. Other information, such as authorship and copyright is generally not provided in the camera's output and it can be filled in during later stages of the image processing as extrinsic metadata. EXIF can be found primarily in TIFF and JPEG files. It is not supported in PNG and GIF files.

- The Information Interchange Model (IIM), developed by the International Press Telecommunications Council (IPTC), is a standard designed to capture information on the activities of news gathering, reporting and publishing. With the widespread adoption of the Adobe Photoshop software, the use of embedded IPTC tags in image files became extensive (NCAM, 2011). A major update of the standard, commonly known as "IPTC Core", enables IPTC data to be incorporated in a range of image formats, such as JPEG, TIFF via Adobe's XMP schema (see below).

- The Extensible Metadata Platform (XMP) is a standard metadata format (ISO 16684-1) developed by Adobe System Inc. for the creation, processing and interchange of metadata in a variety of applications, including desktop applications and back-end publishing systems. XMP metadata can be located in different file formats, including TIFF, JPEG, PNG, GIF, EPS and PDF. Additionally, XMP data can be stored in external XMP sidecar files.

  As the image metadata formats previously described, XMP can be useful for recording the history of an image as it passes through multiple processing steps, from the moment they are photographed or scanned to the editing step and the publication into a final document such an academic article. A figure supporting XMP can receive new metadata properties from each software program or device along the authoring and publishing workflow and carry them along, under the condition that all involved editors support XMP or do not delete it from the resource.

  In a PDF file, XMP metadata may be used to describe the whole document or individual parts of the document, including images. XMP metadata properties are serialized into XML/RDF (see section 4.6) and embedded in a metadata stream contained in a PDF object. The following code extracted from Renz, Brandtzaeg, & Hornef (2012) shows how metadata associated to an image in a PDF of an academic article are included into the code. Among other information, the code shows the title of the image ("nri3112-f3") expressed in the property "title" of the Dublin Core (dc) schema:

```
55 0 obj
<< /Subtype /XML /Type /Metadata /Length 44437 >>
stream
<?xpacket begin="ï»¿" id="W5M0MpCehiHzreSzNTczkc9d"?>
<x:xmpmeta xmlns:x="adobe:ns:meta/" x:xmptk="Adobe XMP Core
5.0-c060 61.134777, 2010/02/12-17:32:00        ">
   <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-
syntax-ns#">
      <rdf:Description rdf:about=""
            xmlns:dc="http://purl.org/dc/elements/1.1/">
         <dc:format>application/pdf</dc:format>
         <dc:title>
            <rdf:Alt>
               <rdf:li xml:lang="x-default">nri3112-
f3</rdf:li>
            </rdf:Alt>
         </dc:title>
      </rdf:Description>
      <!—The rest of serialized RDF goes here. It has been
removed for brevity. -->
  </rdf:RDF>
</x:xmpmeta>
<?xpacket end="r"?>endstream
Endobj
```

The information about embedded metadata is available to users in the advanced options for metadata visualization of the Adobe Reader, a PDF viewer application (Figure 16).

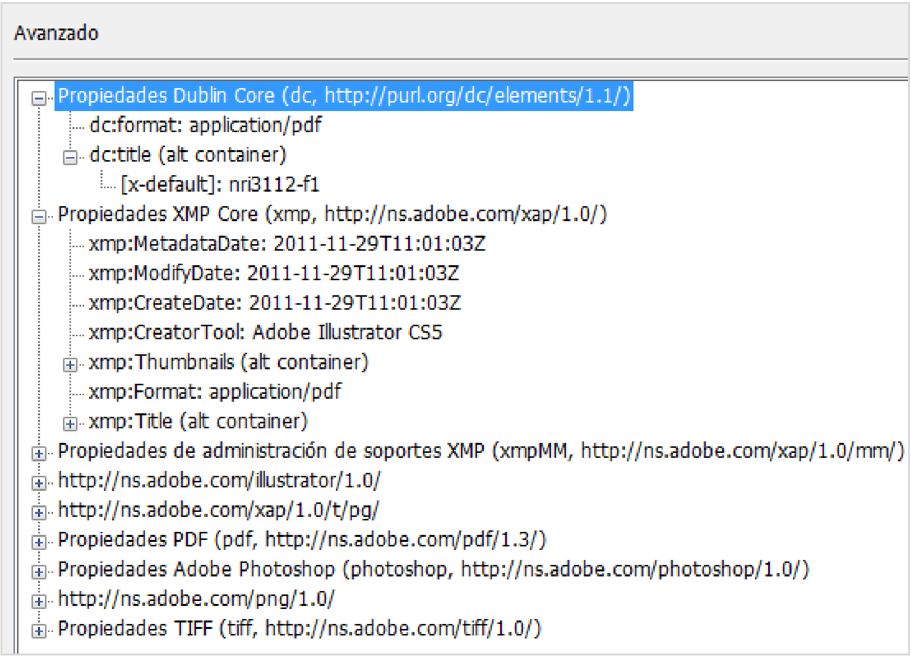


Figure 16 Screenshot of the advanced options of metadata in Adore Reader software.

In DICOM files, metadata are stored in the header section. The header stores a large variety of metadata information about the patient (e.g., name, ID, Sex, Date of Birth), the study (e.g., ID, Referring physician, date, description), the series (e.g., number, date, modality, description), the acquisition equipment (e.g., Institution, manufacturer) and other technical metadata (e.g., transfer syntax, photometric interpretation, image width, image height, bits per pixels, frames), etc. However, DICOM says little about the content or meaning of the pixel data

(Channin et al, 2010). Furthermore, header's metadata are loss during the conversion to other format formats (Varma, 2012).

3.5.5 Associated text of images in biomedical academic papers

Traditionally, the function and use of image-related textual descriptions in academic publications (especially in the area of biomedicine) have been analyzed from the perspective of improving the efficiency and satisfaction of the image retrieval process (Divoli, Wooldridge, & Hearst, 2010); as an element of the figure in the hybrid (text and image) biomedical document retrieval process (Christiansen, Lee, & Chang, 2007) (You et al., 2011) (Apostolova et al., 2013); and for summarizing the image content (Yu & Lee, 2006) (Agarwal & Yu, 2009) (Bhatia, Lahiri, & Mitra, 2009). For example, caption extraction from PDF documents in the domains of chemistry, computer science, physics and astronomy has been studied with the objective of improving image indexing and classification (Choudhury et al., 2013). Captions have been combined with visual features analysis in order to improve accuracy in establishing the modality classification of medical images (Kim, 2010) (Kahn et al., 2012). Studies analyzing the characteristics of image-related texts in academic documents from the perspective of accessibility were not found. The current use of metadata in PDF documents in academic repositories and the consequences for accessibility has been studied by Hewson & Tonkin (2010). However, their study has reviewed metadata at PDF document level, but not at image level.

This chapter has been dedicated to define basic concepts related to digital image publishing. The characteristics of digital images, the most used image and documents formats and the image-related texts commonly used in academic publications have been defined. In the next chapter the fundaments of the digital accessibility discipline will be presented and the solutions offered by this discipline and other disciplines on how to make images accessible and their application to academic publishing are thoroughly described.

Chapter 4

Literature review

4 Literature review

This chapter presents a review of the contributions offered by the accessibility discipline concerning the motivations, methods and techniques on making universally accessible images. Alternative representations and in particular text alternatives of images are described, as they are the most relevant technique for providing access to the visual information to all readers. Parallel contributions (theories, models and methods) from the Semantic Web and Library and Information Science disciplines about the issue are also explored.

In section 4.1, definitions and scope of the digital accessibility discipline are detailed. In particular, the issues related to image accessibility are presented and the target users of the thesis are detailed, according to their characteristics and the barriers they face when they access the content of visual information in STEM academic articles. Then, the accessible design principles and recommendations concerning images are identified and described, according to the most adopted guidelines for the creation of accessible content (WCAG 2.0).

In section 4.2, the characteristics and use of alternative representations of images are presented.

In section 4.3, the accessibility mechanisms currently available for the application of text alternatives are thoroughly detailed.

In section 4.4, image-related texts in STEM articles are explored as potential mechanisms for conveying text alternatives.

Sections 4.5, 4.6 and 4.7 present a review of the research on the use of metadata, semantic annotations and natural-language summarization techniques for the textual description of images.

In section 4.8, the guidelines and the best practices provided by the accessibility discipline for describing the content of images are reviewed.

Finally, section 4.9 presents the models offered by the Library and Information Science research fields for the analysis of image.

4.1 Introduction to digital accessibility

Digital accessibility is a relatively young discipline born in the early 90's as a response to the access problems to digital content caused by the expansion of the Web. For this reason, some authors refer to it as *Web Accessibility*. A widely accepted general definition is provided by the Web Accessibility Initiative (WAI), funded by W3C (Web Accessibility Iniciative (WAI), 2005):

"Web accessibility means that people with disabilities can perceive, understand, navigate, and interact with the Web, and that they can contribute to the Web. Web accessibility also benefits others, including older people with changing abilities due to aging".

The tight connection between accessibility and usability is evident in the definition offered by the ISO 9241-171:2008 (ISO, 2008), in which accessibility is defined as "the usability of a product, service, environment or facility by people with the widest range of capabilities". The difference between the two concepts resides in the target audience of the product: while usability concerns the design of a user interface that is effective, efficient and satisfying for a narrowly defined target audience, "accessibility makes sure the user interface is designed…for more people—especially people with disabilities, in more situations—including with assistive technologies"(Henry, 2007).

Digital accessibility depends on several components working together in the development and interaction of the web and other documents:

- Content: the information in a web page, web application and other documents. This information can include text, images and sounds.
- Web browsers, media players and other user agents.
- Assistive technologies: software used by people with impairments for accessing digital content.
- Users with their specific profile, knowledge and experience.
- Developers: designers, developers and authors who create contents.
- Authoring tools: software that create web sites and other documents.
- Evaluation tools: digital accessibility evaluation tools, HTML validators, CSS validators, etc.

The framework of the digital accessibility is showed in Figure 17.

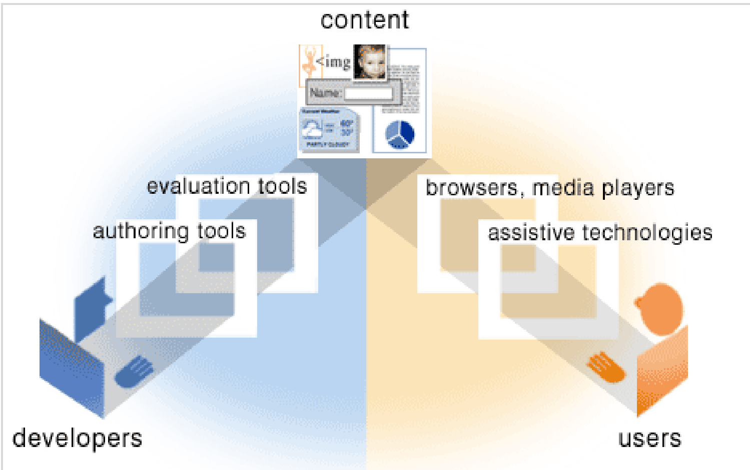


Figure 17 Essential components of digital accessibility. Source (Shawn Lawton & Duffy, 2005).

The two main models currently followed in designing accessible digital content are the Universal Design (UD) model and the User-Centered Design (UCD) model (Ribera, 2010a).

Universal Design (UD) (or Design For All – DFA) is defined as "the process of creating products (devices, environments, systems and processes) which are usable by people with the widest possible range of abilities, operating within the widest possible range of situations (environments, conditions, and circumstances)" (Vanderheiden, 2000).

In other terms, the UD is a design paradigm whose aim is to develop products and environments accessible by every user, regardless of any physical, sensory or intellectual specific characteristic or context. Products designed according to the Universal Design principles and methods do not require a posterior adaptations or specialized design

(Stephanidis, 2009). This paradigm is applied in the Web Content Accessibility Guidelines (WCAG), the official guidelines created by the W3C that regulate the creation of accessible content on the Web.

On another hand, the User-Centered Design (UCD) approach considers a process of designing digital products that meet the needs of all users to be not feasible (Ribera, 2010a). It focuses on the profile, equipment and context of use of specific users in the design of a usable interface**.** According to this paradigm, a web is accessible “if it can be used by specific users with specific disabilities to achieve specific goals with effectiveness, efficiency and satisfaction in a specific context of use” (Petrie & Kheir, 2007). It is based on an early focus on users and tasks, empirical measurements and iterative design (Rubin & Chisnell, 2011).

According to (Ribera, 2010a), the two design paradigms currently present incoherencies that may be bridged by focusing on the integration of two basic needs: on one hand, the need for sharing common basic principles of accessible design; on the other hand, the need for paying attention to any difference dictated by the specific conditions under which the user accesses the web content. This approach includes the adoption of common general rules to ensure the distinction between content, structure and presentation of the digital material, which will enable the technical capability to adapt digital contents according to the needs of different user profiles, equipment and contexts of use.

4.1.1 Disability

The concept of disability has evolved in the last few years. Traditionally, a disability was defined as “a measurable impairment or limitation that interferes with a person’s ability and may refer to a physical, sensory, or mental condition” (Cavender, Trewin, & Hanson, 2014). It was therefore considered as a specific characteristic of a person (Palacios & Romañach, 2007). Currently disability is defined as the condition that results “from the interaction between persons with impairments and attitudinal and environmental barriers that hinders their full and effective participation in society on an equal basis with others” (United Nations, 2006). This vision of the disability considered as dependent from social and contextual aspects has been adopted by The International Classification of Functioning, Disability and Health (ICF), the WHO framework for measuring health and disability at both individual and population levels. It has been also embraced by the regulation on digital accessibility of Spain (Real Decreto Legislativo, 2013).

This semantic shift of the concept of disability is reflected in digital accessibility, whose principles, methodologies and technique are addressed not only to the “traditional” categories of users with the disabilities (as previously mentioned), but also to users under situational limitations that arise from circumstances, environment and conditions while accessing digital content, such as (W3C, 1999):

- The use a text-only screen, a small screen, or a slow Internet connection;
- a situation where eyes, ears, or hands are busy or interfered with (e.g., driving to work, working in a loud environment, etc.);

- the use of an early version of a browser, a different browser entirely, a voice browser, or a different operating system.

4.1.2 Assistive technologies

Digital assistive technologies are all the digital technologies – including software and hardware tools - supporting people with disabilities on the interaction with digital products in a computer device. In the legislation of United States it is defined as "Any item, piece of equipment, or product system, whether acquired commercially off the shelf, modified, or customized, that is used to increase, maintain, or improve the functional capabilities of individuals with disabilities" (Public Law 100-407, 1988). An extension of this definition is provided by the ISO 9999:2011: "Any product (including devices, equipment, instruments and software), especially produced or generally available, used by or for persons with disability: for participation; to protect, support, train, measure or substitute for body functions/structures and activities; or to prevent impairments, activity limitations or participation restrictions" (ISO, 2011).

An example of assistive technology is a screen reader, a software that convert the information displayed on a graphical interface into synthesized speech or into Braille characters on a refreshable Braille display. A screen reader allows blind people accessing, navigating and interacting with the content on the screen, without the need to rely on other individuals. Screen readers interact with the browser's Document Object Model (DOM) and the Application Programming Interface (API) and generate a virtual representation of the page elements (text, links, button, form fields, etc.). The content of this representation is rendered to the user as a linear stream of information. Blind users usually navigate through the content of a page by skimming through the active elements of the page (link or form controls), headings, paragraphs, etc. Screen readers can read not only web content but also the operating system, word processing programs and other applications. The audio presentation conveyed by screen readers may also be beneficial to people with cognitive disabilities during the reading process.

4.1.3 User groups

Some people with disabilities tend to face barriers when they try to access digital content. According to Brajnik (Brajnik, 2009), a barrier is "any condition that hinders the user's progress towards achievement of a goal, when the user is a disabled person". Web accessibility is about removing those barriers, in order to enable people with disabilities using and contributing to the Web. As commented in the previous section, the primary focus of digital accessibility is on people with visual, auditory, physical, speech, cognitive and neurological impairments. People accessing digital content under "situational limitations" are another user group that benefits of the solutions offered by the digital accessibility.

Static images in STEM academic articles convey information that is expressed mainly in a visual language, often combined with text (as described in section 3.2). Visual information in static digital figures presents specific barriers that make it difficult or impossible to access the information they convey to some groups of people with disabilities. These barriers concern the perception and comprehension of the content of the figures and may especially affect users with any type and degree of visual impairments, users with cognitive impairments or users accessing digital content under some “situational limitations”. In general, people with physical disabilities do not experience specific barriers in perceiving and understanding static images, unless their physical disability is combined with one or more disabilities cited before. Users with photosensitive epilepsy are another group with disabilities which generally does not face specific barriers in perceiving and understanding static images, since static images do not present any flickering or flashing effect.

In this section, the groups of users particularly sensible to the barriers presented by information in static figures are reviewed according to their characteristics, the barriers they face in accessing figures, the assistive technology they use and the solutions proposed in accessing the figures by the accessibility discipline.

4.1.3.1 Blind users

Blindness means “anatomic and functional disturbances of the sense of vision of sufficient magnitude to cause total loss of light perception” (Stephanidis, 2013). It implies substantial, uncorrectable loss of vision in both eyes. Chronic blindness can be caused by several factors, including cataract, glaucoma, age-related macular degeneration, corneal opacities, diabetic retinopathy, vitamin A deficiency in children or infections. A person with total blindness is one who has a visual acuity of less than 3/60 (0, 05) to light perception, or a visual field of less than 10 degrees, in the better eye with the best possible correction (WHO, n. d.).

Blind people cannot access information presented only in a visual format. The most common solution proposed by the accessibility discipline to overcome this barrier is the use of a text alternative associated to the visual content. This text makes the image content available to screen readers, which can render it in speech output and/or Braille. Since the creation and application of text alternatives (and other alternatives) is one of the most discussed problems in the accessibility discipline, it will be described further in the section “alternative representations to graphics” (4.2).

4.1.3.2 Low-vision users

Low-vision users have relatively moderate visual impairments that interfere with daily activities and cannot be fully corrected by glasses. The definition include users with poor acuity (vision that is not sharp), tunnel vision (seeing only the middle of the visual field), central field loss (seeing only the edges of the visual field) and clouded vision (Altmanninger & Wöß, 2008). A person with low vision is one who has a visual acuity of less than 6/18 (0,33) to light perception, or a visual field of less than 10 degrees, but who uses, or is potentially able to use, vision for planning and/or execution of a task (World Health Organization, s. f.).

This group includes users with color-blindness, who have difficulties to distinguish certain colors under normal lighting conditions. Most people with color-blindness experience a certain degree of limited perception of particular shades of reds and greens. Color blind people includes people less sensitive (protanomaly) or not sensitive to red light (protanopia) and people less sensitive (deuteranomaly) or not sensitive (deuteranopia) to green light. Another type of color-blindness with a very low incidence is the blue-yellow color-blindness (tritanopia), which limits the ability to distinguish blue from green and yellow from violet.

Low–vision users experience limitations in perceiving and understanding visual content. Their ability to recognize and interpret images and the text inside images may be reduced under the following conditions:

- If contrast between a text in an image and its background is too low. Contrast could affect all users if an image is to be printed in a black and white printout.
- If color is the only way to distinguish an item. For example, when a green arrow indicates a positive amount and a red one indicates a negative amount in a balance, color-blindness users could be unable to distinguish the information conveyed by the arrows, such as any kind users. The ability of any kind of users in perceiving different elements inside a color image printed in a black and white could be also hindered.
- If text is too small or the font family has low legibility.
- If the image has low resolution, it loses sharpness under the high levels of magnification applied by low vision users during the viewing.

Assistive technologies commonly used by people with some degree of visual impairment are screen magnifiers, hardware and software that increase the size of the display. Screen magnifier software use standard display monitors and enlarges information of the screen by pre-determined incremental factor (for example, 1x magnification, 2x magnification). Most of them can magnify the full screen, parts of them or provide a zoom of the area around the cursor or pointer. They usually include extra features such as the ability to change screen colors, enhance mouse pointers or cursors.

The access to visual information by users with visual impairments can be facilitated by different techniques, such as avoiding the use of color as the sole mean to distinguish two or more different information items, providing sufficient visual contrast between all visual elements and their background and in general making all the elements inside an image (including text) distinguishable and legible.

4.1.3.3 Users with cognitive disabilities

Cognitive disabilities are impairments in cognitive processes (Stephanidis, 2009) and can stem from language and learning disabilities (e.g. dyslexia) to brain injury, Alzheimer's disease (i.e., memory retention problems) and dementia (Stephanidis, 2013). Thus, they vary greatly in their impact in the use of information technology.

Users with cognitive disabilities have limited ability to perceive, understand and memorize information, apart from other limitations in cognitive functions, such as attention and mathematical thinking. They may experience difficulties in understanding the information presented in visual format, especially in graphics such as diagrams that are complex and poorly constructed (Murphy, 2005). An example of information conveyed by images particularly difficult to grasp by people with cognitive impairment may be “a correlation between variables in a complex scatter plot” (Brajnik, 2009).

Users with dyslexia - which is a neurological problem affecting language and commonly dealing with reading and writing - can benefit from accessing the content of web pages by different modalities, for example combining listening simultaneously with reading visually (RNIB, 2013). For this reason, people with dyslexia may use specialized screen readers that read websites aloud to them, such as blind users commonly do. Consequently, they can access the textual equivalent of an image (see next section) (Cunningham, 2012).

Comprehension of visual information by users with cognitive impairments can be facilitated by paying particular attention to the design of visual features such as color contrast, position and size of visual elements, type, size and spacing of text inside the image. Some types of images such as images of text do not allow the adaptation of text presentation and users with a cognitive disability could find them particularly hard to perceive and interpret.

Abbreviations - including those used in text inside figures - present barriers in comprehension to readers with cognitive impairments and in general to all readers and should be explained.

4.1.3.4 Users with a situational limitation

Some users may experience barriers in accessing digital content that are determined by the context of access. These barriers can depend on some “disabling technology” or environmental conditions. Example of “disabling technologies” may be mobile devices with reduced screen, screen with limited color palette or grayscale screen (such as some types of e-reader devices) and text-based web browsers (which do not display images) (Brajnik, 2009). Examples of users subjected to environmental limitations are people accessing digital text in poor or excessive lighting conditions or under stress conditions (e.g. in a hurry, in a noisy and distracting environment, while carrying out some other important task).

4.1.3.5 Search engines

Even though they are not “human”, in many cases spiders of search engines work similarly to web users who are blind (Pemberton, 2003) and face the same kinds of barriers that humans face (Brajnik, 2009).

4.1.3.6 Synthesis of target users

The review of the users helps us to identify the profile of the audience who benefits of the improvements in image accessibility in the workflow of biomedical academic articles (point F in the figure publishing chain, see Figure 4) and to answer to the first question (Q1) of the thesis: “Which are the profiles of the target readers with disabilities of STEM academic articles?”. After reviewing the bibliography, five categories of users can be considered the target groups who benefit of improvements in the accessibility of images:

1. blind users;
2. users with visual impairments;
3. users with cognitive disabilities;

4. users with situational disabilities;
5. search engines.

4.1.4 Definition of accessible design principles addressed to images and image-related texts

Once defined the main target users groups, next step is to analyze the most adopted accessibility guidelines for content creation (WCAG 2.0), in order to identify which guidelines deal with the content of images and which barriers these guidelines help the target users groups to overcome. The aim of this analysis is to identify which aspects should be taken into account by authors of STEM academic article for creating accessible images.

The current version of the WCAG (2.0) (W3C, 2008) defines the accessibility of a web site in terms of its adherence to four basic principles:

- Perceivable: "Information and user interface components must be presentable to users in ways they can perceive" (W3C, 2008). In other terms, digital content should be perceived by users with different preferences and abilities. For achieving this, alternatives in different modalities should be provided to users with some sensory and perceptual impairment. For example, text alternatives should be provided to all non-text content (including images) for being accessible to blind people.
- Operable: "User interface components and navigation must be operable" (W3C, 2008). Interaction and navigation should be designed for being used by every user. For example, digital contents should be designed for being accessible by keyword to blind people.
- Understandable: "Information and the operation of user interface must be understandable" (W3C, 2008). Content and interface should be designed to be easily understandable to all kind of users. For example, text content should be designed according to readability principles in order to be accessible to people with cognitive impairments.
- "Robust: "Content must be robust enough that it can be interpreted reliably by a wide variety of user agents, including assistive technologies" (W3C, 2008). The compatibility of web sites with different user agents (web browsers, media players, plug-in, etc.) and assistive technologies (see below) should be maximized.

A set of 12 guidelines defines the application of these four theoretical principles and a group of 61 success criteria related to different guidelines helps validating the accessibility of web content, according to three levels of conformance: A (lowest), AA and AAA (highest). In synthesis, according to the WCAG 2.0 approach, a web site is accessible "if it is perceivable, understandable, operable by users despite their impairments and if it is robust with respect to user agents and assistive technologies" (Brajnik, 2011).

Being the principles, guidelines and success criteria of WCAG 2.0 technology-neutral, they are relevant to the development of web content, web applications and digital publications in general (Brewer, 2013). These guidelines have reached a world-wide acceptation and legislations of many countries reference them or are based on them, including the Section 508

Amendment to the Rehabilitation Act in the United States and the "Ley general de derechos de las personas con discapacidad y de su inclusión social" in Spain (see section 1.2).

Table VII shows the WCAG 2.0 principles, guidelines and success criteria (SC) concerning images (taking into account their visual and textual elements and the image-related elements) and the barrier types and user groups specific for each guideline according to Brajnik (Brajnik, 2009).

Table VII WCAG 2.0 principles, guidelines, success criteria and conformance Level related to content in images

WCAG 2.0 Principle	WCAG 2.0 Guideline	WCAG 2.0 SC	Conf. Level	Barrier	User group
1. Perceivable	1.1. Text Alternatives: Provide text alternatives for any non-text content so that it can be changed into other forms people need, such as large print, braille, speech, symbols or simpler language	1.1.1 Text alternatives	A	Rich images lacking equivalent text	blind, cognitive, contextual, search engine
1. Perceivable	1.4. Distinguishable. Make it easier for users to see and hear content	1.4.1 Use of color	A	Color perception is necessary to understand information	blind, low vision, contextual
1. Perceivable		1.4.3, 1.4.6 Contrast	AA, AAA	Insufficient visual contrast	low vision, contextual
1. Perceivable		1.4.4 Resize text	AA	Text cannot be resized	low vision, contextual
3. Understandable	3.1. Readable: Make text content readable and understandable	3.1.3 Unusual words	AAA	Complex text	cognitive
3. Understandable		3.1.4 Abbreviations	AAA	Acronyms and abbreviations without expansions	cognitive

The six barriers defined in Table VII are caused by the lack of application of accessible principles in the design of the images. These design principles concern specific elements of the image, such as its color, contrast, text, etc. WCAG 2.0 provides guidelines on how to overcome these barriers. Information concerning causes, effects on users with disabilities and WCAG 2.0 recommendations are detailed below, according to the model of analysis of accessibility barriers proposed by Brajnick (2009):

Barrier n.1 Rich images lacking equivalent text.

- Cause: Images provide information only in a graphical format. No equivalent textual description is provided.
- Addressed to: images conveying relevant information or having a relevant function (see section 4.2.2).
- Failure mode: Blind users have no way to get the information presented (only) via graphics. Users with cognitive disabilities might not be able to grasp the meaning of some of the information displayed in the image. Users that have images turned off to save downloading costs in a mobile dispositive cannot access information presented (only) via graphics. Search engines will have less text to analyze and rank appropriately the page in which the image is located.
- Fix: Add an equivalent text description to the image.

Barrier n.2 Color is necessary.

- Cause: Color is used as the sole mean to distinguish two or more different information items inside an image.
- Addressed to: Color images.
- Failure mode: People with low-vision can be unable to perceive colors as those with normal vision capabilities. Such people would be unable to distinguish between two information items, such as the red and yellow lines in Figure 18. People with protanopia could be unable to distinguish between two or more information items shown in green and red. Readers using devices with grayscale screen or reading black and white printed version of the image might not be able to distinguish between two or more information items.
- Fix: Any information presented in color should be also presented in a way that does not depend on color perception, for example using patterns (Figure 19).

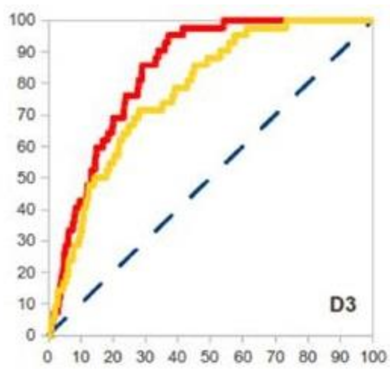


Figure 18 Example of a line chart where colors are used as unique mean to differentiate information among lines. Extracted from: (Anighoro & Rastelli, 2013).

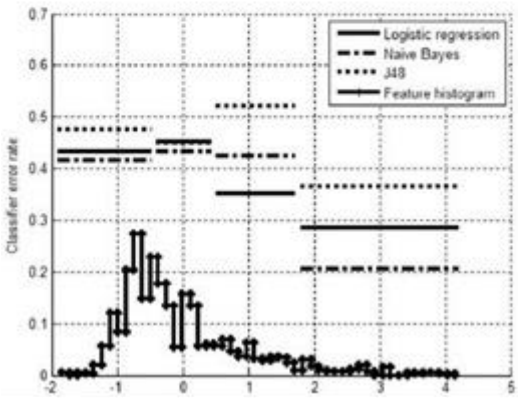


Figure 19 Example of line chart where patterns are used to differentiate information among lines. Extracted from: (Audhkhasi & Narayanan, 2013).

Barrier n.3: Insufficient visual contrast.

- Cause: Contrast with background is too low.
- Addressed to: Visual features and text inside images.
- Failure mode: Users with low-vision will have to make an effort to distinguish between foreground and background visual and textual items inside the image and in some cases will be totally incapable of recognizing, and perhaps even detecting, the foreground items. Users reading images in poor lighting condition will have to make an effort to distinguish between foreground and background items. The general lack of quality of image (due for example to low resolution) can get worse the ability to distinguish items inside the image to all readers.
- Fix: Provide sufficient visual contrast between all visual information (including text inside images) and its background.

Barrier n.4 Text cannot be resized.

- Cause: Text inside images is too small or illegible and cannot be resized since it is painted in pixel within images.
- Addressed to: Text inside images.
- Failure mode: Low vision users cannot read text that is painted in pixel within images and has low resolution. Depending on the browser or other reading application being used, users might not be able to increase the font size of the text.
- Fix: Text should be presented in easy-to-read fonts.

Barrier n.5 Complex text

- Cause: Images contain text that is complex to read because of the complexity of the sentences, phrases or words.
- Addressed to: Text inside images and other image-related texts (captions and mentions).
- Failure mode: Users will struggle in understanding the content and in navigating through it.
- Fix: Simplify as much as possible the structure of sentences and the lexicon used. Avoid using jargon. Make sure there are no spelling errors.

Barrier n.6 Acronyms and abbreviations without expansions

- Cause: The image contains acronyms and/or abbreviations that are not expanded.
- Addressed to: Text inside images and other image-related texts (captions and mentions).
- Failure mode: Users might not understand what the abbreviation means. Even if it was explained in previously seen pages, the user might not remember it.
- Fix: Always add the expansion to every abbreviation and acronym.

As commented before, the accessible principles in the design of the images concern specific elements of the image (including image-related texts): text alternatives, use of color, color contrast, resolution, dimensions, font size and font family of text in images, content of text in images and in caption/mentions and content of text in images and in caption/mention. Table VIII details these elements and relates them to their corresponding accessible principles according to the view of the author of the thesis. These are the elements that should be analyzed for assessing a universal accessibility of images in STEM articles.

Table VIII Image-related accessibility issues according to WCAG 2.0.

WCAG 2.0 Principle and guideline(s)	WCAG 2.0 Success Criterion	Level	Barrier	User group	Related to
1. Perceivable 1.1. Text Alternatives	1.1.1 Text alternatives	A	Rich images lacking equivalent text	blind, cognitive, contextual, search engine	text alternatives
1. Perceivable 1.4. Distinguishable	1.4.1 Use of color	A	Color perception is necessary to understand information	blind, low vision, contextual	use of color
1. Perceivable 1.4. Distinguishable	1.4.3, 1.4.6 Contrast	AA, AAA	Insufficient visual contrast	low vision, contextual	color contrast, resolution, dimensions
1. Perceivable 1.4. Distinguishable	1.4.4 Resize text	AA	Text cannot be resized	low vision, contextual	font size and font family of text in images
3. Understandable 3.1. Readable	3.1.3 Unusual words	AAA	Complex text	cognitive	content of text in images and in caption/mentions

WCAG 2.0 Principle and guideline(s)	WCAG 2.0 Success Criterion	Level	Barrier	User group	Related to
3. Understandable 3.1. Readable	3.1.4 Abbreviations	AAA	Acronyms and abbreviations without expansions	cognitive	content of text in images and in caption/mentions

In summary, in this section the design principles for making accessible image according to the guidelines of the WCAG 2.0 have been defined. These principles are general and apply to all types of images, including STEM images in academic articles. They define which elements of the images (including visual elements, text inside the image and image-related texts) should be analyzed in order to assess the current accessibility of images in STEM academic articles. These accessible design principles should be followed by the actors of the process of creation, editing and submission of images in STEM articles in order to create and publish accessible images.

Next section will discuss more in detail the current solutions for making accessible images and in particular the techniques for applying alternatives representations to images.

4.2 Alternative representations to graphics

In this section, the solutions offered by the accessibility discipline for providing access to the visual information to people with impairments (especially with severe visual impairment) are presented. The definition and use of different alternative representations to images are assessed and in particular the characteristics, function and application of text alternatives. The use of text alternatives is explicitly recommended by WCAG 2.0 and several mechanisms for their association to images in common publishing formats of academic articles (HTML and PDF) are currently available.

The aim of this and the following sections (4.2 –4.8) is to answer to the second question (Q2) of the thesis: “Which are the current solutions for making visual content in STEM academic articles accessible to people with disabilities”?

4.2.1 Definition of alternative representations of graphics

Images have implicit knowledge that is expressed in a visual language. This language is available only to those capable of perceiving and understanding it. People with disabilities in general and in particular people with visual disabilities can undergo limitations or be excluded from accessing the knowledge. Digital accessibility offers a variety of solutions for making information conveyed by images available to users with visual disabilities, which mainly consist on creating alternative representations of the images. These alternatives rely on converting information normally acquired by one sense to a modality suitable for another sense.

The main solution proposed by accessibility design for the presentation of the information conveyed by figures is to provide an appropriate text alternative to figures. As commented in the previous section, Guideline 1.1 of WCAG 2.0 (W3C, 2008) recommend to "provide text alternatives for any non-text content so that it can be changed into other forms people need, such as large print, Braille, speech, symbols or simpler language". A text alternative is defined as "a text that is used in place of non-text content for those who cannot view the non-text content" (W3C, 2012b). The lack of a text alternative is an accessibility barrier to users with difficulties to perceive non-text content (such as people with visual disabilities), people using browsers that do not support the display of images and search engine bots.

The application of appropriate text alternatives (and other alternative representations) should be tailored to different kinds of images, their function, the content they convey and the context in which they are presented (W3C, 2014g). In the next section, different categories of images are described according to their specific content and function, emphasizing the characteristics of the group of images to which static images in STEM articles belong.

4.2.2 When to apply alternative representations

WCAG divide images into two main categories, depending on their function: decorative images, which have only an aesthetic function, and informative images, which provide informative content, enhance greater understanding of the content or play a relevant role into the context where the image is presented.

Decorative images do not present important content and are used for layout or non-informative purposes (WebAIM, 2012). This type of image is defined as "purely decorative", since it serves only an aesthetic purpose, provides no information and has no functionality (W3C, 2014l). This category includes:

- Images used only for visual formatting, such as lay-out spacers or bullets.
- Images used as part of a text link, such as an icon of a house next to the link "Home".
- Images used to add ambience or visual interest to the page ("eye-candy"), such as the image of an exotic plant that merely reinforce visually the advertisement of a greenhouse.

WCAG do not require text alternative for this type of images. In contrast, informative images have the function of illustrating concepts and information graphically and require an alternative representation (and in particular, a text alternative). This category includes different types of functional and informative images, defined according to their specific content and function.

Functional images are images that provide important functions, such as those used as part of the navigation bar (e.g., an image containing the text "Products"), as navigational icon (e.g., an arrow with the function of passing to the next page of a series), as self-explanatory icon (e.g., an icon showing the type of document to download) or as a linked logo, and are presented alone without any accompanying text. The text alternative of these images should explain their function.

Other specific types of functional images are:

- Image maps, which contains multiple clickable areas.
- Image of text, which consist in text that has been rendered in a non-text form in order to achieve a particular visual effect (e.g., logotypes which combine textual information and graphic display to convey the brand name), as commented in section 3.4.1.
- Artistic image: a visual representation that cannot be clearly included in one of the two main categories (decorative/informative). It conveys a specific sensory experience that is not purely decorative and does not primarily convey important information or perform a function. An example of this type of image could be an aquarelle painting inspired by a poem on a page reciting that poem.

Informative images can be distinguished between simple and complex according to their level of complexity:

- "Simple" informative images**:** images conveying simple concepts or information that can be expressed in a short phrase or sentence in a text alternative. Examples of these images include:
  - Image used to supplement other information. For example, an image showing a dog wearing a bell supplements the adjacent text that explains the purpose of this bell.
  - Image conveying succinct information, e.g., a simple diagram that illustrates an anti-clockwise direction for unscrewing a bottle top or cap.
  - Image conveying an impression or emotion, such as a photograph showing a smiling family group intended to give the impression that the website or the company it represents is family friendly.

- Complex informative images**:** images that contain substantial, rich and/or high-structured information. They require an alternative text that should be a full text equivalent of the data or information provided in the image and cannot be expressed in a short phrase or sentence. This category can include:
  - Graphs showing statistical data, including line graphs, pie charts, scatter plots, etc.
  - Charts showing a sequence of activities and a process, such as flow charts.
  - Maps showing locations or other information such as weather reports.
  - Screenshots;
  - Diagrams and illustrations in general, where the page text relies on the user being able to understand the image.
  - Groups of images used to convey a single piece of information.
  - Mathematical notations and formulae.
  - Visual representations in scientific domains. For example:
    - Optical images obtained from a camera, such as a photograph of a skin disease or a micrograph of cells obtained by a microscopy.
    - Medical images obtained by different modalities, such as the magnetic resonance imaging (MRI) of a brain or the ultrasonography of a foetus.
    - Drawings illustrating of complex phenomena, such as a natural process (e.g., the carbon cycle or a chemical, molecular, or cellular structure).

- Visualization of phenomena in 3D, such as the rendering of the Raleigh-Taylor instability.

Images in STEM academic publications, such as those presented in section 3.1.2, are commonly semantically dense and convey professional/technical information (Burford, Briggs, & Eakins, 2003). Since they are usually informative and complex, they require to be presented by text alternatives or by other alternative representations in order to make them accessible to all types of readers.

4.2.3 Typologies of alternative representations to graphics

This section presents the following alternative representations to figures currently used to make images accessible to people with disabilities:

- large print;
- alternatives in audition modality, which includes speech and sonification;
- alternatives in tactile modality, which includes tactile graphics and haptics;
- audition-tactile modalities combined.

4.2.3.1 Large print

Documents in large print are documents which text has a minimum recommend size of 16 point, ideally minimum of 18 point (UKAAF, 2012b). Images in documents in large print format are images designed in an amplified version intended to be usable by partially sighted people and people with learning difficulties. Although large print documents are widely used, they present several disadvantages (Swenson, 2013):

- Their production is expensive and they are often printed in black and white.
- Because of their size, they are heavy and bulky to carry.
- Sometimes they are produced enlarging photocopies and the resulting letters and pictures are blurred or hard to read.
- Captions of images are often not correctly enlarged.
- Compared to the option of reading text (and images) with optical aids (magnifiers and reading glasses) or in digital devices that allow the user to choose the magnification level, large print limits the ability of accessing regular print by readers and narrow their range of options for reading.

4.2.3.2 Audition modality

Speech

This modality depends on the creation of a textual description to the visual content. Audio presentation of text can either be speech recordings (such as those found in most audio books and some digital talking books), or synthesized speech produced by text to speech (TTS) technology. The TTS technology integrated into screen reader software, such as VoiceOver or Jaws, can automatically read text displayed by a computer (Power & Jürgensen, 2010).

__Sonification of complex data__

Sonification refers to the use of non-speech sound to convey information and render data, either to enhance the visual display or as a purely audio display. An example of sonification is the one routinely used in hospitals to keep track of the rhythms in electroencephalogram (EEG) signals. Non-speech sound can be used to improve accessibility in different ways, for example by supporting the navigation of structured documents in the Web, the review of data in spreadsheets or in auditory graphs (Peres et al., 2008). Sonification is more effective in the presentation of diagrams and schematic representations (auditory graphs) in which it is much easier to separate the meaning from the presentation of the information. Therefore, it has been mainly proposed as an alternative presentation to statistical diagrams. An interesting review on the use of diagram sonification (auditory graphs) has been presented by Abu Doush (Abu Doush et al., 2010).

The sonification of complex data in diagrams has the advantage of dealing with a wide range of objects without being constrained by a specific lexicon. However, it has serious limitations:

- It can be effective in summarizing trends, but it is less effective in creating fine-grained presentations (Abu Doush et al., 2010), for example in presenting the comparison of two related lines on a line graph (University of Delaware, 2013).
- There is no agreement among listeners on how perceptual properties such as pitch should be deciphered. For example, “listeners might agree that increasing pitch suggests increasing temperature, yet the same group of listeners may feel that decreasing pitch offers a more intuitive representation of increasing size” (Stephanidis, 2009).

Some authors have recognized the limitations of this alternative representation and recommend it mostly as a supplemental tool to convey visual information, not as the central mode of communication (Keane & Laverentz, 2014).

4.2.3.3 Tactile modality

__Tactile graphics__

Tactile graphics are physical counterparts of their printed equivalents displayed in a 2D space, with various materials providing depth or texture to the graphic by reproducing contours or raised areas. One of the most common techniques for creating tactile graphics is embossing, where raised dots are printed within a small distance to create 2D structures. Embossing printers generates Braille text intermixed with graphical elements and the tactile graphic is usually presented with a text explanation written in Braille characters.

Tactile graphics can reproduce different types of images, including diagrams, illustrations, drawings and maps. They are not mere transcriptions of print illustrations or raised versions of printed graphics, they are representations of images adapted for the sense of touch. Due to the fact that the tactile sense is considerably less sensitive than the visual sense and works in a more serial manner than vision (Stephanidis, 2011), the graphic needs to be re-designed in a simplified way to make sense in a tactile form for blind and partial sighted readers.

Besides the traditional 2D tactile graphics, it's worth mentioning the recent 3D printer technology to print tactile graphics. A 3D printer creates accurate graphics that have the ability to emphasize a specific line or part of the graphic by reproducing substantial variations in height (Stangl, Kim, & Yeh, 2014). In general, it can produce three-dimensional object from digital models. Among other applications, 3D objects reproducing STEM material such as molecular structures are used as teaching material for blind student (Siu, 2014). The use of touch for exploring 3D objects is also employed in haptics technology (described in the next section).

Tactile graphics have some disadvantages that obstruct/hinder its widespread adoption (Power & Jürgensen, 2010):

- they are substantially larger than standard print documents and they are hard to transport and store.
- They often provoke a loss of details and information, due to the space and resolution limitations in the tactile medium.
- They have high costs of production (including the price of the embossing printer, the tactile developing system and the support material) that limit their availability.
- They are hard to update, since every change into a tactile document requires regeneration either of a portion or the whole of the document.
- It has been estimated that no more than 10-25% of braille readers can read a tactile graphic (Gardner, 2014).

Haptics

Haptic interfaces (or simply haptics) are tactual interfaces that allow the user to feel via touch kinesthetic forces, vibration, temperature, shape and other tactile surface properties of the virtual object by the use of a force feedback interface. The representation of various forms of picture characteristics (boundary, shape, colour, texture) or the size and orientation of 3D geometrical objects can be rendered by haptic techniques. Haptics help blind people to explore virtual textures and 3D objects through the sense of touch and some studies have shown that the use of haptic modalities enables blind people to spatially interpret the information in an image (Roth et al., 2001) (Lévesque, 2005). However, most haptic technology is a solution still not affordable for the home users and its adoption into mainstream acceptance and use may still be several years away (Darrah, 2013).

Tactile and audition modalities combined

Although tactile graphics remain a primary method of providing blind learners with access to visual material, tactile alone is considered insufficient to convey the content of the image and supplementary information is required alongside the figure. To potentially overcome the limitations of a pure tactile (or a pure audio) presentation, the combination of audio and tactile feedback has been proposed by some authors (Gardner & Bulatov, 2006). Tactile/audio presentation has been tested with bar charts and it has been recognized particularly beneficial for users who are not proficient in reading Braille (Goncu, Marriott, & Hurst, 2010). A particularly successful implementation of audio/haptic presentation is the IVEO touch sensory device, where the tactile graphic is placed on a press-sensitive screen and when the user touches different areas of the graphic a relevant audio feedback is provided as a complementary description. The image digital format utilized by the IVEO system for generating tactile graphics is SVG. IVEO is an actual implementation of alternatives to graphics in STEM fields, including health and biology.

4.2.4 Conclusions on alternative representations of graphics

The review of the alternative representations of images has emphasized some limitations that currently hinder their effective adoption in the academic publishing workflow:

- They are expensive to produce and to update (large print, tactile graphics).
- They have limited availability, adoption and acceptance (large print, tactile graphics and haptics).
- Their effectiveness in rendering visual information is limited to specific types of images and they use perceptual properties whose meaning is not clear to all users (sonification).
- While the audio-tactile solution has been recognized as beneficial, it presents the same limitations of the tactile graphics.

These limitations may be overcome by adopting text alternatives, which are considered the a primary way for making information accessible (W3C, 2014l). They can be rendered by a variety of user agents through any sensory modality (for example, visual, auditory or tactile) and therefore they convey the content and function of images according to the needs of different user profiles.

4.3 Text alternatives to images

This section presents the functions and characteristics of text alternatives and the mechanisms used for including them in the most common documents of STEM articles according to accessibility principles and guidelines.

4.3.1 What is a text alternative?

A text alternative is a mean for providing a textual replacement for non-text content. In particular, its function is to serve the same purpose and present the same information as the original non-text content. When a text alternative is correctly provided, it can replace the non-text content and no functionality or information would be lost (W3C, 2014k). Non-textual content is “any content that is not a sequence of characters that can be programmatically determined or where the sequence is not expressing something in human language” (W3C, 2012b). The definition of non-textual content includes all the types of images described in section 3.1.1, audio and video content (excluding controls, time-dependent multimedia, captcha, etc.), ASCII Art (pictures pieced together from the 95 printable characters defined by the ASCII Standard), emoticons and leetspeak (an alternative symbolic alphabet for the English language used primarily on the Internet).

A non-text content, such as a static 2D images in a biomedical article in HTML, can be "programmatically determinable". This means that when a figure in a HTML document is properly marked up according to specific techniques (that will be described below), the "semantics" of the images, such as its content and position, can be determined in a standard, machine or software readable form. Therefore, user agents (including assistive technologies) can programmatically extract this information and present it to users in different ways.

There are two main types of text alternatives to images:

- A short text alternative, which conveys the purpose and information of the image in a short phrase or sentence.
- A long text alternative, which provides a rich, expressive description necessary to explain the details of the graphic.

When shall we use a short and when a long text alternative? The application of short and long text alternatives to images depends on the type of image, its function and the context where the image is presented. As a general rule, if it is possible to create a text alternative in a short sentence this option is always preferable. If it is not possible to provide a text equivalent in a short phrase or sentence, then the short text alternative should provide a brief overview of the information and a long text alternative would be used in addition to convey the full information (W3C, 2014k). For example, if a picture shows how to change a tire, a short text alternative describes what the picture is about, while a long text alternative describes the exact procedure.

4.3.2 Requirements of the mechanisms for applying text alternatives

The mechanisms used to provide these alternatives should follow basic requirements for conveying text alternatives efficiently according to different use cases (W3C, 2014h) (Groves, 2013) (Watson, 2014):

- Availability: they must provide a description for images.
- Discoverability: they must allow discovering and accessing the text alternative. In other words, they must allow people to be able to find out that a text alternative is available and provide a way to access it. Ideally, the "call to action" should be visible (for sighted people) and programmatic (for non-sighted people).
- Structured markup: They must allow the inclusion of rich markup and semantic structures in the text description. For example, a graph or chart is likely to need a structured detailed description that includes headings, lists, data tables and so forth. Markup is also important as you may want to provide alternative representations of the data, alternate media to describe the data, etc.

The mechanisms for applying text alternatives vary according to the document format where the images are included. In the next section the mechanisms for applying text alternatives to images in HTML and PDF documents (the most common formats currently used for disseminating academic articles, as commented in section 3.4) are described in detail. The mechanisms for figures in other formats relevant to accessible publishing, such as EPUB and SVG, are also presented.

4.3.3 Mechanisms for conveying text alternatives in HTML

In this section, the following main techniques to provide text alternatives to images in HTML documents are presented:

1. *Alt* attribute in the <img> element for a short description
2. *Longdesc* attribute in the <img> element.
3. Long description in a link immediately adjacent to the image
4. Long description in the text near the image, with a reference to the location of the long description in the short description

The following mechanisms alternative to the main techniques are also presented: <object>, <title>, <figure> and <figcaption> tags and *wai-aria* attributes.

4.3.3.1 Alt attribute in the <img> element

The primary and most used technique for conveying a text alternative is by including a text in the *alt* attribute of the <img> element in HTML documents. It is meant to replace the image when it cannot be rendered or accessed. The value of this attribute is commonly referred to as *alt text* and WCAG 2.0 does not specify any particular length for it. Accessibility experts do not agree on its maximum length, with recommendations varying from 50 to 1024 characters depending on the author (Clark, 2003)(Korpela, 2012) (Social Security Administration, 2010) (W3C, 2000b). However, some experts recommend keeping alt text shorter than 125 characters, since popular screen readers such as the Freedom Scientific JAWS split the *alt* attribute into distinct chunks of 125 characters each and read each chunk separately as if they were separate graphics (DO-IT knowledge database, 2013). A test conducted by the author of the thesis has detected a similar performance: JAWS 16 splits the *alt* attribute into distinct chunks of 254 characters (including spaces) each and read each chunk separately. The test has been performed in Mozilla Firefox 35 and Google Chrome 40.

The function of the *alt* attribute varies depending on whether the information provided in the image can be summed up in few words or not.

When the information provided by the image can be summed up in few words, the *alt* attribute serves as a replacement for the image and provides an equivalent description of the content. In this case, it could be possible to remove the image and replace it with the text alternative and no functionality or information would be lost. The example image (Figure 20) and the code below show the use of the *alt* attribute just described:

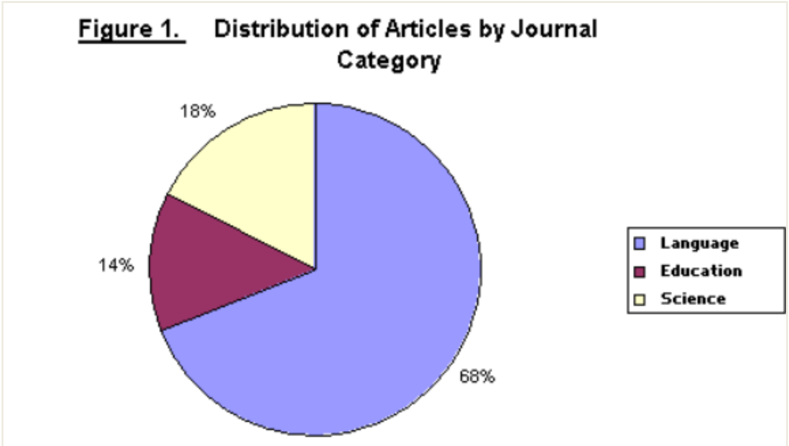


Figure 20 Example of a pie chart Extracted from: (W3C, 2014g).

```
<p>
<img src="figure1.gif" alt="Distribution of Articles by Journal
Category. Pie chart: Language=68%, Education=14% and
Science=18%." />
</p>
```

When a short description cannot serve the same purpose and present the same information as the non-text content, the *alt* attribute provides information that serves as an identification of the content, the *alt* attribute serves as a short description of the content and is accompanied by a programmatically associated longer text alternative. An example of this application is provided in the section focused on the use of *longdesc* attribute (see next section).

The *alt* attribute can be also used to provide the literal transcription of the text represented in images of text (images that are not graphical or illustrative in nature but can be all, mostly, or heavily made up of text). The example image (Figure 21) and the code below show this use of the *alt* attribute for an image of text used as an unlinked logo.

Figure 21 Example of an image of text. Extracted from (W3C-WAI Act Project, 2014)

```
<img src="wai.png" alt="Web Accessibility initiative">
```

The *alt* attribute mechanism has the following advantages (W3C, 2014g, p. 5):

- It is a “programmatically determinable” text. For example, when a screen reader recognizes an image, it looks for its *alt* attribute and reads aloud its textual content to the user.
- Its application does not impact upon the default visual display of content.
- It is supported in most assistive technologies and in non-graphical web browsers.

However, it has the following drawbacks:

- Its content is generally not visible to some users who may find it useful. For example, people with cognitive disabilities, who can take advantage of accessing textual descriptions for enhancing the comprehension of images, are not aware *of* the existence of the *alt* text unless they are users of screen readers. However, some user agents provide ways to make the *alt* attribute discoverable by sighted users (for example, the “Popup Image Alt Attribute” extension in Chrome).

- It does not support the inclusion of semantic structures, such as headings, paragraphs and lists, nor the indication of language.
- If its content is too long, it may cause display issues in some browsers or reading issues in some assistive technologies. For example, as commented before, screen readers such as JAWS split the *alt* attribute into distinct chunks of limited characters.

In synthesis, considering the basic requirements of text alternatives defined in section 4.3.2 (availability, discoverability and structured markup), the *alt* attribute meets the availability criterion and partially the discoverability criterion (with the programmatic “call to action”, not with the visible “call to action”). It does not meet the structured markup criterion.

4.3.3.2 *Longdesc* attribute in the <img> element.

Unlike the *alt* attribute (which provides essential information), the *longdesc* attribute provides a rich, expressive description necessary to explain the details of the graphic (Clark, 2003). This is implemented with a special kind of link pointing to a longer description of the image that may be located in a separate page or within the same document as the image. Assistive technologies supporting the *longdesc* attribute announce the user that a description is available, for example announcing “has long description” or “Press ‘enter’ to hear long description” and the user can choose to hear the description or not.

The Figure 22 is an example of a chart requiring an *alt* and a *longdesc* attribute. This is the case where the short description provided in the *alt* attribute is inadequate to serve the same purpose and present the same information of the image (as explained in the previous section) and a longer description is required.

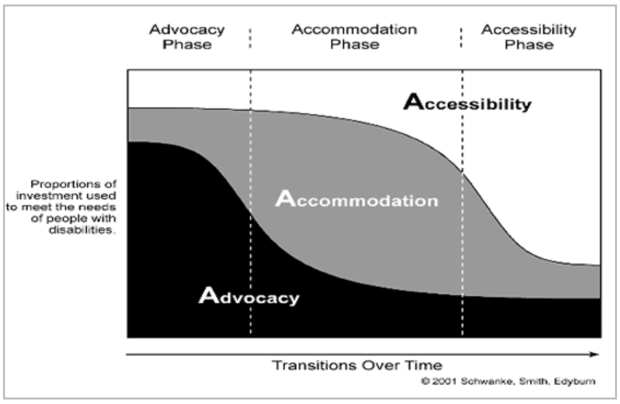


Figure 22 Example of an area chart. Extracted from: (ACCESS-ed, 2011).

The following code presents an example of the syntax of a *longdesc* attribute for Figure 22 pointing to an external page, along with an *alt* attribute that provides information used as an identification or as a very short description of the image.

```
<!-- The description is within an external page -->
<img src="a3_diagram.png" alt="Diagram representing the A3
Theoretical Model" longdesc=“a3_diagram_explained.html” />
```

The following code presents the syntax of a *longdesc* attribute of Figure 22 linking a long description in the same document, along with the *alt* attribute (which use is the same as the previous example).

```
<!-- The description is somewhere on the same page as the image
-->
<img src="a3_diagram.png" alt="Diagram representing the A3
Theoretical Model" longdesc="#a3_diagram_explained" />
<div id="a3_diagram_explained">
 <h3>Long Description: the A3 Theoretical Model</h3>
 This graph or diagram serves is a representation of the A3
 Model, a theoretical model that…<!-- Full Description of Graph
 -->
 <p>Long description ends.</p>
</div>
```

Compared to the option of providing the description in an external page, the presentation of the description in the same document may have the disadvantage that the user is not able to identify the final of the description. For this reason, the end of the description should be clearly indicated in the text, as showed in the example.

The content of the long description of Figure 22 could be the following, adapted from (ACCESS-ed, 2011):

“This chart is a representation of the A3 Model, a theoretical model that demonstrates how organizations provide services for the inclusion of people with disabilities. The x-axis is labeled “Transition Over Time”. The triptych-like diagram is labeled at the top, from left to right, with "Advocacy", then "Accommodation", then "Accessibility". The y-axis is labeled “Proportions of investment used to meet the needs of people with disabilities”. The phases are presented with the 3 panels divided by dotted lines. The intervention approach of advocacy is represented as the bottom layer across the diagram in black. The intervention approach of accommodation is represented as the middle layer across the diagram in gray. The intervention approach of accessibility is represented as the top layer across the diagram in white. The diagram looks like a rolling wave. The intervention approaches of advocacy (in black), accommodation (in gray), and accessibility (white) sum to 100% in each of the three parts or phases, depicting the proportional balance of the organization’s approach at any point in time. The left panel shows primarily black then gray in lesser quantity and then white in even lesser quantity. The middle panel shows the largest portion as gray (or accommodation-based). The right panel shows primarily white space and less of gray and even lesser of black”.

The *longdesc* mechanism has the advantage to be explicitly associated with the image for users of assistive technology and to allow the inclusion of structured markup. Recent research has also found that most common screen readers have good support of the *longdesc* attribute (W3C, 2014h). Even though it is not part of the core specification of HTML5, it has been recently advanced to a Recommendation by the WC3, as the attribute supporting the link descriptions to images in HTML5 content (W3C, 2015) .

However, some browsers do not currently make visually obvious that the *longdesc* attribute is available on images, thus it is only partially discoverable by sighted people (University of Minnesota Duluth, 2014).

In synthesis, the *longdesc* attribute is a mechanism that could meet the criteria of availability, structured markup and programmatic discoverability, while it has limitations in making visible the “call to action” (“visible” discoverability).

4.3.3.3 Techniques alternatives to the *longdesc* attribute

Long description in a link immediately adjacent to the image

Another technique for providing a long description is by including a link immediately adjacent (before or after) to the image pointing to the long description in another location than the image (W3C, 2014c).

The advantage of this technique is that it can be used in any technology and meet the availability, discoverability and structured markup conditions. However, an important drawback of this technique is that the small image with the link pointing to the long description can be considered as a “visual encumbrance” in those situations where the design requirement restricts the use of visible links on a page (Groves, 2014).

Long description in the text near the image, with a reference to the location of the long description in the short description

The long description can also be provided as part of the standard presentation, near the image both visually and in the linear reading order but it has not to be the very next item of the image. For example, there may be a caption under a chart with the long description provided in the following paragraph. In this case the *alt* attribute identifies the image and specifies where the long description is located. The following example is adapted from that provided in (W3C, 2014d):

There is a bar chart on a Web page showing the sales for the top three salespeople. The *alt* attribute says: "October sales chart for top three salespeople. Details in text following the chart". The paragraph immediately below the chart says: "Sales for…<!-- Full Description of Graph -->”.

This technique meets the requirements of availability, discoverability and structured markup conditions. It makes the long description visible to all users, without requiring the user to jump off to another location for the description.

4.3.3.4 <object> element

The body of the <object> element can be used to embed objects such as images, audio, videos, PDF and Flash within an HTML document. A text alternative can be provided in the <object> element as a fallback mechanism for those browsers or user agents unable to render the image. The code below is an example of the use of the body of the <object> element to render a text alternative of Figure 20 in SVG included in a HTML document:

```
<object type:"image/svg.xml" data="figure1.svg">Distribution of
Articles by Journal Category. Pie chart: Language=68%,
Education=14% and Science=18%.</object>
```

However, the <object> element is not designed as a specific mechanism to convey text alternatives and other methods specific for conveying text alternatives (as those presented previously) are preferable to it (W3C, 2014f).

4.3.3.5 title attribute

The title attribute is contained in the <img> element and "offers advisory information about the element for which it is set" (w3c 1999). Its purpose is conveying supplementary (not essential) information and it can be helpful for providing optional details about the image or adding some verve to the description of the image (Clark 2003). Browsers may render the values of the <title> attribute by displaying the title as a tool tip, a small hover box with information about the item being hovered over. However, accessibility experts consider the <title> attribute not to be a useful technique for conveying important information to all users (Faulkner, 2012).

4.3.3.6 <figure> and <figcaption> elements

As commented in section 3.4, the <figure> element is a self-contained element that can be used to encapsulate images (illustrations, diagrams, photos, code listings, etc.) introduced in HTML5. The <figcaption> element is used to explicitly associate a caption to the <figure> element. It may be nested in the <figure> element in such a way that it is programmatically associated with the image it describes. Furthermore, if the <figcaption> content is sufficient to describe the image to non-visual users, it may be used as a longer description together with the text alternative provided in the *alt* attribute (DIAGRAM Center, 2014b), as shown in the example code below created by the author :

```
<figure role="group">
<img src="/figure1.jpg" alt="Panoramic view of Barcelona" />
<figcaption> A panoramic view of the city of Barcelona on
December. Original by <a
href="http://www.flickr.com/photos/barcelona/">Joan
Martorell</a>
</figcaption>
</figure>
```

The <figure> element permits to programmatically determine the caption of a figure, providing a description of the image. It supports the inclusion of structured markup, including links to external resources. This is an advantage compared to *alt* attribute in the <img> element.

However, the <figure> element has currently some limitations. Browser support for these new elements is currently limited to interpreting them as generic <div> elements and no user agent is available that take advantage of the specialized tagging (e.g., to provide the option to ignore) (International Digital Publishing Forum, 2014a).

In synthesis, this technique could meet the three requirements of availability, discoverability and structured markup. However, currently there are some limitations in their support and intended use.

4.3.3.7 WAI-ARIA techniques

WAI-ARIA 1.0 (Web Accessibility Initiative – Accessible Rich Internet Applications) is a W3C specification created for making more accessible dynamic content and controls developed with Ajax, HTML, Javascript and other related technologies. Compared to the techniques previously described, which are focused on the description of static images, the mechanisms offered by this specification have been developed for improving the accessibility of dynamic content in general, and thus they can be also applied to images. WCAG 2.0 currently recommends two ARIA techniques for associating text alternatives to image: the *aria-labelledby* attribute for short text descriptions and the *aria-describedby* attribute for long text descriptions.

The *aria-labelledby* attribute may be used to provide a short text description for complex graphics. The following example (W3C, 2014a) shows the syntax of a read-only complex graphic of a star rating pattern, composed of several image elements. The text alternative for the graphic is the label, visible on the page beneath the star pattern:

```
<div role="img" aria-labelledby="star_id">
<img src="fullstar.png" alt="" />
<img src="fullstar.png" alt="" />
<img src="fullstar.png" alt="" />
<img src="fullstar.png" alt="" />
<img src="emptystar.png" alt="" />
</div>

<div id="star_id">4 of 5</div>
```

The aria-describedby attribute may be used to link to a more verbose description of the image in the same location of the figure, when the content or function of the image cannot be provided in a short text or requires markup (including lists, data tables and MathML). Its function is similar to the *longdesc* attribute, but this attribute cannot reference a description outside of the page containing the image and the technique of inclusion is different. It associates an element with text that is visible elsewhere on the page by using an ID reference value, as shown in the next example (DIAGRAM Center, 2014b):

```
<img src="/watermolecule.jpg" aria-describedby="watermolecule-
description" alt="A water molecule." />
<p id="watermolecule-description">A water molecule is made of
two hydrogen atoms and one oxygen atom. The two hydrogen atoms
are positioned on the oxygen atom and are separated by
approximately 105 degrees. </p>
```

The advantages of these techniques is they are available, discoverable (both programmatic and visible) and can include HTML structured markup. However, they have still limited support by some assistive technologies and browsers.

Another attribute called *aria-describedat* is currently a proposal and it will allow including the URL of a page containing the longer description of the image.

4.3.4 Mechanisms for conveying text alternatives in PDF

Before reviewing the mechanisms currently available for including text alternatives to vector and raster images in PDF documents, it is necessary to provide information about the general accessibility of PDF documents.

Since version 1.4, PDF documents have improved their accessibility by including structural tags to define headers, tables, lists, etc. By using the appropriate tagging, an assistive technology such as a screen reader can present a summary of the document, facilitate navigation, provide structural information of the content, etc. (Ribera, 2008).

Figures providing relevant content must be tagged with the <figure> tag. Figures which do not contribute to the vital document contents (such as background images and decorative images in general) must be marked with the <artifact> tag (which is also used for other contents not relevant to the document's meaning, such as headers and footers). Figure 23 shows the tags of an academic article in PDF as displayed in Adobe Acrobat 9 Pro. The tag assigned to a figure (<figure>) in the article and the corresponding figure in the document are highlighted.

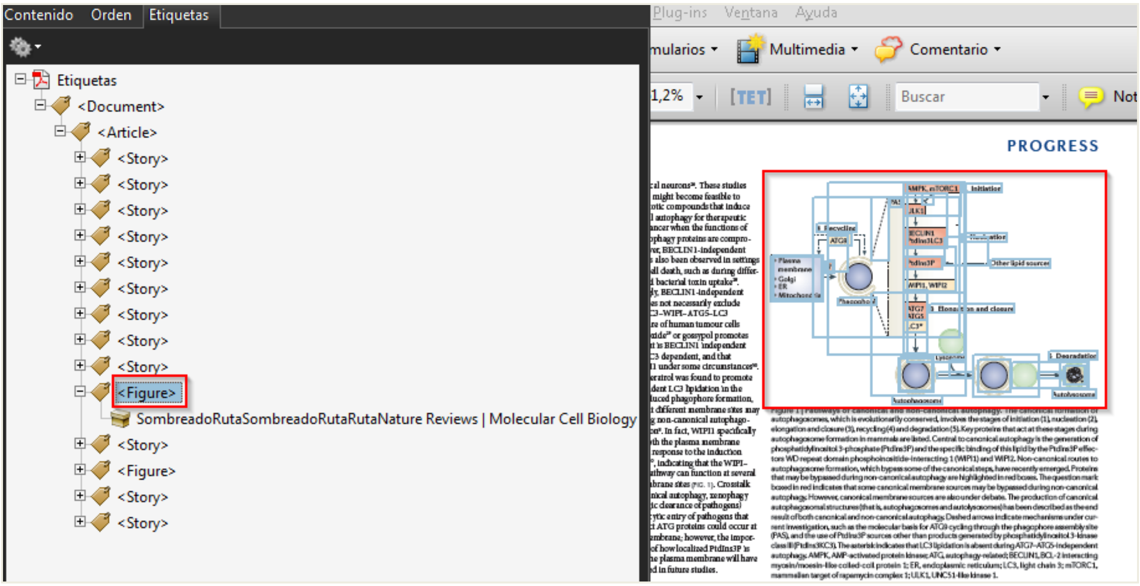


Figure 23 Screenshot of a dialogue box showing tags in Adobe Acrobat Pro9. The figure reproduces a page of the article of (Codogno, Mehrpour, & Proikas-Cezanne, 2012).

When raster graphics are correctly tagged by the <figure> tag, they can support the inclusion of the *title*, *alternate text* and *actual text* attributes (Figure 24):

1. The *alternate text* attribute allows the inclusion of text alternatives to images. PDF specification and WCAG do not recommend any specific length of text alternatives included in the *alternate text* attribute. In a simple test performed with Adobe Acrobat 9 Pro, the author of the thesis has detected that the value of the *alternate text* attribute allows the inclusion of 31900 characters (included spaces) maximum.
2. The *actual text* attribute supports the literal transcription of an image of text.
3. The *title* attribute assigns a title to the image in the same way as the title attribute in HTML.

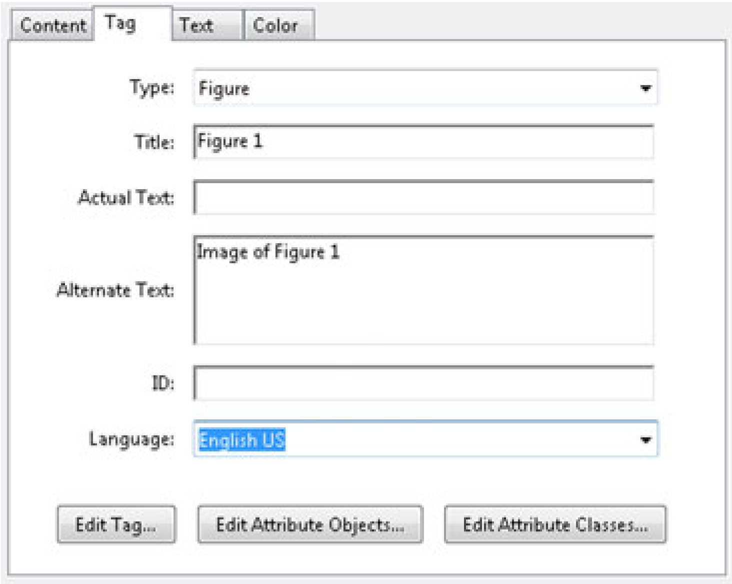


Figure 24 Screenshot of the dialog box for editing title, actual text and alternate text attributes in Adobe Acrobat software.

PDF also supports the <caption> tag, which semantically marks the caption of images in PDF.

The use of the different techniques for providing text alternatives to images in PDF is recommended in the PDF/UA-1 specification according to the following recommendations:

- Graphics objects other than text objects and artifacts must be tagged with a Figure tag.
- A caption must be tagged with a Caption tag.
- An “informative” figure must have a text alternative.
- Actual text should not be used for a figure when an alternate text is more appropriate.

The main limitation of the mechanisms offered in PDF is that while raster graphics can be tagged as a whole, each line on a vector image is identified as a path. Consequently, the association of *title*, *alternate text* or *actual text* attributes to a vector graphic is a highly complex manual process which requires the regrouping of geometrical primitives.

PDF programs or assistive technologies provide access to tagged PDF documents. However, being the PDF/UA technical specification recent, they do not still offer a full PDF/UA compliance (PDF Association, 2013). Promising collaboration between Adobe and assistive technologies manufacturers are currently ongoing, such as the partnership of the NVDA screen reader producer with the PDF Association (PDF Association, 2012) for the implementation of full support to PDF/UA.

Both alternatives (*alternate text* attribute and *actual* attribute) meet the requirements of availability and programmatic discoverability, but they do not meet the visible discoverability and structured markup requirements. The *caption* attribute meet also the requirement of visible discoverability.

4.3.5 Mechanisms for conveying text alternatives in EPUB

As described in section 3.4.4, the EPUB3 content specification is based on a subset of HTML5 and the techniques for including text alternatives to images described previously for HTML documents can be applied to images in EPUB3 documents.

The recommended methods for including accessible informative images in EPUB3 are proposed by the DIAGRAM Center (2014b):

1. If the image is sufficiently described in the surrounding text, use the *alt* attribute to label the image as the one being referenced and described in the surrounding text.
2. If the image is sufficiently described in the caption and the caption is programmatically associated to the <figure> element using the <figcaption> element, use the *alt* attribute to label the image as the one being referenced in the caption.
3. If the image is not sufficiently described in the surrounding text:
    a. If the image description is brief and can be presented without any markup, present the description using the *alt* attribute.
    b. If the image description is lengthy or requires markup (including lists, data tables or MathML), present the description using *aria-describedby*.

Advantages and disadvantages of these methods are the same as those described previously when applied to images in HTML documents.

4.3.6 Mechanisms for conveying text alternatives in SVG

SVG format supports the inclusion of specific mechanisms for embedding textual descriptions.

The <title> element provides a human-readable title for the element that contains. It may be rendered by a graphical user agent as a tooltip or as speech by a speech synthesizer. The <title> can be used also to include the literal text alternative of an image of text. The <desc> element supports a longer text description of an element that contains it. Authors are recommended to provide this kind of description for complex images or other content that has functional meaning, as in the following example (W3C, 2000a):

```
<?xml version="1.0"?><!DOCTYPE svg PUBLIC "-//W3C//DTD SVG
20000802//EN"
  "http://www.w3.org/TR/2000/CR-SVG-20000802/DTD/svg-
20000802.dtd">
<svg width="6in" height="4.5in" viewBox="0 0 600 450">
  <title>Network</title>
  <desc>An example of a computer network based on a hub</desc>
</svg>
```

The <title> and <desc> tags can be used to different graphical object levels, since they can be associated to a shape or a specific graphical element, a group of elements or a whole document. As an SVG image may consist of several components combined hierarchically, the hierarchy and text equivalent descriptions can be combined for helping who cannot see to create a rough mental model of an image (W3C, 2000a). An example of this application is provided by Schepers (Schepers, 2013) and reproduced in the following code:

```
<svg xmlns="http://www.w3.org/2000/svg"
```

```
         xmlns:xlink="http://www.w3.org/1999/xlink"
         width="100%" height="100%" viewBox="0 0 160 180">
    <title>Star</title>
    <desc>2 overlapping triangles, forming a blue and gold 6-
pointed star, with a blue circle in the background</desc>
    <circle cx="80" cy="90" r="75" fill="none"
stroke="cornflowerblue" stroke-width="5"/>
    <g style="stroke-width:5; stroke-linejoin:round;">
      <title>Triangles</title>
      <polygon id="triangle-up" points="80,10 150,130 10,130"
fill="gold" stroke="cornflowerblue">
        <title>Upwards triangle</title>
        <desc>A gold equilateral triangle, pointing upwards, with a
blue outline.</desc>
      </polygon>
      <polygon id="triangle-down" points="150,50 80,170 10,50"
fill="cornflowerblue" stroke="gold">
        <title>Downwards triangle</title>
        <desc>A blue equilateral triangle, pointing downwards, with
a gold outline.</desc>
      </polygon>
    </g>
  </svg>
```

SVG is a very good option for graphics and it will be presented as a choice in the conclusions, as its support is increasing. Despite the advantages just mentioned, SVG solutions for conveying text alternatives have currently some limitations in reading systems, browser and screen reader support. The evaluations performed by the Book Industry Study Group (Book Industry Study Group, 2013) shows different levels of support of SVG features among e-book apps, device and reading system. A test commissioned by Adobe (Watson, 2013) reports the following results:

- Jaws 15 and Internet Explorer 10 is the best combination at present, since the screen reader can navigate to the SVG image, identify it as an image, and announce the accessible name (in the <title> element) and accessible description (in the <desc> element). However, nested SVG images are not well supported.
- NVDA and Internet Explorer combination present the <role>, <title< and <desc>, but its support is erratic (multiple images are announced and title and desc are repeated).
- In all other browser/screen reader combinations the SVG is ignored.

Besides the limitations regarding the current support of SVG, another limitation is that SVG document contains the semantics of the image only implicitly. It cannot overcome the absence of information behind the graphical representation. Although SVG allows the author to include a text description for each logical component of the image, without any contextual information provided in the image it is difficult or even impossible to infer what the author of the figure intended it to mean.

In synthesis, currently the described mechanisms for including text alternatives in a SVG file meets partially the requirements of availability, discoverability and structured markup, due to the current lack of support. However, it is expected that this limitation will decrease as the adoption increases.

4.3.7 Conclusions about recommended mechanisms for applying text alternatives

In this section, the most common mechanisms for providing text alternatives as they are recommended in accessibility guidelines have been described. Their advantages and limitations have been reviewed:

- The *alt* attribute supports a short description used as a textual substitute of the image; it is basically useful only for screen readers and its content is not available to some sighted people who may find it useful; if its content is too long, it may cause display issues in some browsers or reading issues in some assistive technologies. Finally, it does not support the inclusion of semantic structures.
- The *longdesc* attribute supports a long description for the image. However, it is not part of the core specification of HTML5 and it cannot be generally discoverable and actionable by sighted users.
- The *alternate* text and the *actual text* attributes in PDF are not discoverable by sighted people and do not support structured markup.
- The techniques recommended in EPUB3 have the same limitations as those applied to images in HTML documents.
- Due to the current lack of support the mechanisms for including alternative description in SVG meets partially the requirements of availability, discoverability and structured markup.

Table IX shows a summary of the most common mechanisms for each type of alternative description (short, long, actual text) and document format and their conformance to the three main criteria defined for alternative descriptions.

Table IX Common mechanisms for providing text alternatives and their conformance to the availability, discoverability and structured markup requirements. "Partial" conformance means that the technique conforms to the requirement but currently it is partially supported by browsers and screen readers.

text alternative	Document format	availability (mechanism)	discoverability (programmatic)	discoverability (visible)	structured markup
short description	HTML	@alt in <img>	yes	partially	no
	PDF	@alternate text In <figure>	yes	no	no
	EPUB3	@alt in <img>	yes	no	no
	SVG	<title>	partially	partially	partially

text alternative	Document format	availability (mechanism)	discoverability (programmatic)	discoverability (visible)	structured markup
long description	HTML	@longdesc in <img>	yes	partially	yes
	PDF	@alternate text In <figure>	yes	no	no
	EPUB3	@longdesc in <img>	yes	partially	yes
	SVG	<desc>	partially	partially	partially
actual text	HTML	@alt in <img>	yes	partially	no
	PDF	@actualtext in <figure>	yes	no	no
	EPUB3	@alt in <img>	yes	no	no
	SVG	<title>	partially	partially	partially

In synthesis, current mechanisms recommended for conveying text alternatives to figures in HTML and PDF documents have some limitations in meeting the criteria of availability, discoverability and structured markup. Therefore, they may fall short in conveying text alternatives to the complex scientific figures commonly included in STEM articles.

In order to explore other possible mechanisms for conveying text alternative, in the next section image-related texts are explored as potential text alternatives to images in STEM academic articles.

4.4 Image-related texts in STEM articles as potential text alternatives

In the previous section, it has been emphasized that the main mechanisms offered for providing text alternatives to HTML, PDF and EPUB3 documents have some limitations and their application in digital documents in general, and in scientific literature in particular, is incorrect.

In this section other possible solutions related to the current practices in the workflow of the scientific academic publishing are explored. As described in section 3.5, the figures in STEM academic articles may be associated to related texts, such as captions, mentions, text in the images or metadata. Two are the main reasons that justify the exploration of these texts as possible alternatives:

- They are already included in the practices of authors when submitting images to STEM articles and in the image workflow in general, as discussed in section 3.5.
- Accessibility guidelines acknowledge the use of contextual or surrounding text as a way for conveying a long text alternative for complex image in specific cases.

The aim of this section is to explore the viability of the use of image-related texts in STEM articles as other options for conveying text alternatives. The study of the viability concerns the criteria of availability, discoverability and structured markup.

4.4.1 Caption

Although WCAG do not explicitly recommend the use of caption as a text alternative, captions are considered a way for supplying a text based description meaningful for blind people "when properly used" (Ault et al., 2002). Its use under specific conditions is acknowledged by the guidelines of different accessibility organizations, such as WebAIM (WebAIM, 2012) and DIAGRAM Center (DIAGRAM Center, 2014b), and it has been proposed by the W3C among the techniques useful for text alternatives in HTML5 (W3C, 2014e).

For example, as commented in the previous section, the <figure> and <figcaption> elements can be used to semantically associate a caption to the image it refers. If the <figcaption> presents an equivalent alternative for the image, the *alt* attribute may be used just to identify the image (W3C, 2014e). The same solution is acknowledged for figures in EPUB3 documents, also under the condition that the <figcaption> content is sufficient to describe the image to non-visual users (DIAGRAM Center, 2014b), as showed in the example below:

```
<figure>
<img src="watermolecule.jpg" alt="Figure 1">
<figcaption>A water molecule is made of two hydrogen atoms and
one oxygen atom. The two hydrogen atoms are positioned on the
oxygen atom and are separated by approximately 105 degrees, as
shown in Figure 1.
</figcaption>
</figure>
```

This mechanism meets the conditions of availability, discoverability and structured content. However, as discussed in the previous section, it is subject of the same limitations in supporting <figure> and <figcaption> elements.

PDF specifications and WCAG do not provide any explicit reference to the use of the <caption> tag for including a text alternative to images in PDF. However, from the point of view of the author of the thesis, the use of caption suggested in HTML documents can be extended to PDF (and EPUB3) under the same conditions.

4.4.2 Mentions

Mentions explicitly refer to images and they are part of the contextual text of the figure.

The use of contextual text is allowed as a text alternative in HTML and, as commented in section 4.3, it may be semantically associated to the image by several techniques (the *longdesc* attribute, a link immediately adjacent to the image, a reference to the location of the long description in the short description in the *alt* text, the *aria-describedby* attribute). In EPUB documents, if the image is sufficiently described in the surrounding text, the *alt* attribute of the <img> element can be used in order to label the image as the one being referenced and described in the surrounding text (DIAGRAM Center, 2014b).

While there is no mechanism to semantically connect contextual text and the image it refers to in PDF documents, it should be noted that usually mentions explicitly refers to the figures by using numbers (e.g., "Figure 5 shows…," or "See fig. 3.2") in academic articles.

In synthesis, as contextual text, mentions could be technically used as long descriptions, meeting the conditions of availability, discoverability (both programmatic and visible) and structured markup.

4.4.3 Text labels and text inside the image

As described in section 3.5.3, textual labels and all the text included in images (excluding captions, which are considered in this study as an image-related text on its own) in academic articles convey textual information used to label, point to or indicate some relevant features of the image.

The main technique used to convey this information in HTML and EPUB3 documents is the *alt* attribute, which is used in this particular case to provide the literal transcription of the text represented in images of text. In PDF documents, this text can be provided in the *actual text* attribute. These techniques meet the availability and programmatic discoverability requirements, but they fail in matching the visible discoverability and the structured markup requirements.

4.4.4 Metadata

As discussed in section 3.5.4., image metadata can provide descriptive, administrative and structural information to images in academic articles. Accessibility guidelines do not explicitly address the use of metadata as text alternatives for figures. However, as noted by some authors (Kawanaka et al., 2009), metadata can serve as a means to provide text alternatives and contribute on the improvement of image accessibility in digital documents. The potential of metadata as a mechanism for conveying text alternatives to images has been explored in several research projects and it will described in detail in section 4.5.

4.4.5 Conclusions about image-related texts as text alternatives

Table X shows a summary of the most common image-related text in academic articles reviewed until now and how their meet basic criteria for text alternatives.

Table X Mechanisms for conveying text alternatives in mage-related texts and their conformance to the availability, discoverability and structured markup requirements. “Partial” conformance means the same as in Table IX.

Image-related text	Document format	Availability (mechanism)	Discoverability (programmatic)	Discoverability (visible)	Structured Markup
caption	HTML	<figure> + <figcaption>	partially	partially	partially
	PDF	<figure> + <caption>	yes	yes	yes
	EPUB3	<figure> + <figcaption>	partially	partially	partially
mention	HTML	contextual text referred from @alt, @longdesc, a link	yes	yes	yes
		@aria-describedby	partially	partially	partially
	PDF	no	no	no	no
	EPUB3	contextual text referred from @alt @longdesc, @aria-describedby	yes	yes	yes
		@aria-describedby	partially	partially	partially
text in figure	HTML	@alt in <img>	yes	partially	no
	PDF	@actualtext in <figure>	yes	no	no
	EPUB3	@alt in <img>	yes	no	no

In synthesis, image-related texts may be technically supported by mechanisms currently available for associating text alternatives to images, including figures in academic articles in PDF and HTML formats. The analysis of the viability of image-related text as alternative texts has been assessed only as a conformance to the criteria of availability, discoverability and structured markup. It has not dealt with the content and function of these texts, which should be also analyzed, after the review of the guidelines concerning the function and content of text alternatives currently available (section 4.8).

Besides the options explored in this section, the review of the mechanisms for providing text alternatives is extended to the possible solutions offered by research fields different from accessibility, such Semantic Web and Library and Information Science. Next sections (4.5- 4.7) present the possibilities offered by the application of metadata (section 4.5), Web Semantic technologies (section 4.6) and natural-language summarization (section 4.7) in providing text descriptions of the content of images and their potential as text alternatives.

4.5 Projects on metadata as text alternatives

Two main approaches have been proposed for providing metadata as text descriptions of images: internal or *embedded* metadata, stored in the same file or structure of the image, and external metadata, located in a separate file or field from the image (for example, in a database or in separate XML-encoded file).

4.5.1 Internal metadata

The potential of embedded metadata as a mechanism for conveying and maintain text alternatives to images along the digital books publishing workflow has been assessed by the DIAGRAM Center (NCAM, 2011), citing the proposal of the Metadata Working Group, a group formed by Adobe Systems, Apple, Canon, Microsoft, Nokia and Sony (Metadata Working Group, 2010). After reviewing the most used metadata formats supported in the image formats used in publishing (EXIF, IPTC, XMP, Dublin Core, SVG), some key description fields in three common metadata formats has been identified as potential image description containers. The content descriptor fields identified for conveying a textual description of the image's content are the Dublin Core *dc:description* property in XMP format, the *ImageDescription* property in EXIF and the *Caption* property in IPTC. Although all of the "description" fields identified in the three metadata specifications are independent of the others, they have enough commonality amongst them for making possible "to develop a coherent workflow to harness certain metadata fields for use as image description containers for accessibility" (NCAM, 2011).

4.5.2 External metadata

External metadata can be created by experts or by social tagging or semi-automatically generated by combining techniques of automatic recognition and human tagging. Some projects based on these approaches and aiming at providing text description to images are described below, with an especial emphasis on the automatic generation on tags.

The Webinsight system (Bigham et al., 2006) has been developed by the Rochester University for the automatic creation and insertion of text alternatives into web pages on-the-fly. When a blind user loads a Web page, WebinSight systems retrieves text alternatives from a database when it is available. Then alternative text is automatically inserted into the value of the *alt* attribute of the associated <img> tag. When the text alternative is not present in the database, the system calculates the alternative text automatically by three techniques:

- Context analysis. The content of the *title* and *longdesc* attributes and text found on the pages pointed to by linked images are analyzed to formulate a text alternative. For example, the text in the title of a page linked by the logo of an organization commonly describes the content of the image.
- Optical character recognition (OCR), used to recognize the text inside images of text.
- When images cannot be labeled automatically, they can be optionally sent to human labeling services with fee if the user desires it.

The system developed by Keysers (Keysers, Renn, & Breuel, 2007) provides a fully automatic text alternative to images in HTML documents. Given an image without a description, firstly it finds a set of images from a collection of labeled images that are visually similar to the untagged image. Then the descriptions of the similar images are used to predict a suitable description for the new image. Category labels are also supplied to the image, based on the terms that often occur in Web pages, like *logo*, *icon* or *photo*. Although the system does not attempt to perform a deep semantic analysis of the input images, it provides some useful information including color, location and the presence of people. The limitation of this approach is the high dependency on a well labeled image database for providing good results in automatic generation of text alternatives.

Other projects have proposed collaborative metadata authoring as a solution for providing text alternatives to visually impaired users under request (Bigham et al., 2010). This method does not require permissions to modify the code and theoretically anyone can improve the accessibility simply by creating metadata (Kawanaka et al., 2009).

4.5.3 A metadata content model for text alternatives

Besides the techniques reviewed previously concerning how to provide metadata for describing images textually to people with visual impairments, a content model has been proposed by the Diagram Center defining how the alternative representations of images (text alternatives, tactile graphics, etc.) are incorporated into digital publications through metadata (DIAGRAM Center, 2014a) in a structured, standard way. The proposal aims to improve the efficiency in which alternative representations are handled by different authoring and reading tools and therefore to reduce the cost of creating the alternatives to students with print disabilities. The content model is based on the ANSI/NISO Z39.98-2012 standard (see section 3.4.3) and it is defined according to the information requirements of the three general profiles involved in the production and consumption of accessible digital documents: producers (e.g. publishers, authors and alternative format production specialist), facilitators (educators, parents, disabilities specialists) and consumers (blind users, low vision users and cognitively disabled users). From these use cases, information requirements are defined. Afterwards, metadata and content are defined according to the requirements. The metadata define the factors that describe, for example, the purpose of the image (why the image is included in the book), the age or grade level of the person to whom the image is addressed, the version of the content model, etc. The content describes the actual information of the image alternative and could include such elements:

- Long description (<longdesc>): a comprehensive verbal representation of the image exposing the constitution, structure, and significance of the image. This is the primary description of the image and thus cannot be omitted. Its function is similar to the *longdesc* attribute's function in HTML documents.
- Summary (<summary>): a brief and concise version of the long description giving its central idea. This description complements the primary description and provides a short description, such as the one included in the *alt* attribute in HTML documents.
- Simplified Language (<simplifiedLanguageDescription>): description of the image in basic, easy to understand language. This alternative description complements the primary description. It is different from the techniques previously reviewed and it is particularly addressed to users with cognitive disabilities.
- Annotation (<annotation>): relevant explanations to the subject matter that cannot be included in the main description. They may include clarification of some words used in the main description.

These instances of image descriptions are included in or referenced by the original graphical elements of digital books.

The DIAGRAM Content Model has been used as a basis for allowing users to add image alternatives to a file in the Tobi software (DAISY Consortium, 2014). This is a free open-source authoring tool developed by the DAISY Consortium for the production of EPUB3 and full-text full-audio DAISY DTB files. It offers a framework wherein the author can supply text and audio describing the image(s) in a full-text, full-audio DTB or EPUB.

4.5.4 Potential contributions and limitation of metadata as text alternatives

Currently the use of metadata in different image workflows, from consumer sharing experiences to professional-level asset management is "critical" (Metadata Working Group, 2010). However, large-scale industrial image metadata assignment is notoriously complex, due to the following factors (NCAM, 2011):

- The large number of metadata formats and specifications for embedding information into digital images are developed independently and provokes a lack of interoperability and integration of metadata in the publishing workflow.
- Since the authors and the electronic publishing industry use different tools for authoring images, the support of metadata handling by these tools is variable and presents some limitations in embedding image metadata and preserving them along the workflow.

The use of metadata (internal or external) for the description of the content of the image has the potentiality of maintain the description embedded in the image file or associated to it and preserve it by being lost along the editing and publishing image workflow. However, the lack of interoperability among formats and tools for authoring images and metadata handling currently prevents this method from being totally effective.

4.6 Image annotation using Sematic Web

This section explores the possibilities offered by the Semantic Web techniques for the semantic description of figures by text. The discipline of Semantic Web and its technologies are introduced and their use for image annotation is detailed. Promising solutions offered by the Semantic Web for providing text alternatives to images are presented and evaluated. Finally, projects for improving image accessibility by generating alternative representations using Semantic Web technologies are reviewed.

4.6.1 Introduction to Semantic Web

Until ten years ago (and *still today, to a lesser extent), the* Web was primarily a diverse document-centric set of distributed hyperlinked HTML pages that formed a virtual repository of information (Lacy, 2005). Most Web content developers targeted their information for direct human consumption and the Web was basically a medium where HTML documents, served up to browsers, were read by humans. Essentially it was like a huge global library of documents available for reading. Although the simplicity of the Web has been a key factor for its astonishing growth, it has become one of its weak points, once the Web has attained a size that makes it unmanageable for users without further assistance. To help users in dealing with this huge amount of information, computers need to have a clearer knowledge about what Web content is about. The greater part of pages currently forming the Web contains semi-structured HTML text, used to indicate the structure and lay-out of documents, but lacks of information explaining the meaning (semantics) of their content. Therefore, the meaning of web content is still largely not machine-accessible and Web content isn't characterized and described appropriately for efficient automated interpretation by machines.

Recently, the new vision of the Web as an environment of machine-human interaction has driven the creation of the Semantic Web, which can be defined as "an *extension* of the current Web from documents to data" (Herman, 2009). It allows data to be defined and linked in such a way that they can be used by machines for the automatic processing and integration of data across different applications. It aims at improving the structured information representations of the Web content by providing semantic mark-up and make them more easily machine-processable.

One of the aims of the Semantic Web is to capture meaning and consequently apply Knowledge Representation (KR) concepts at a web scale. Knowledge Representation in the context of Computer Science is about using a formal language to make knowledge as explicit as possible from a formal and machine-aware point of view. This is necessary because knowledge is usually captured by humans in implicit form (García, 2010), for instance as social conventions that are required to understand the text in a web page.

To develop semantic-based technologies, meanings should be managed autonomously and separated from data, content and application code, in order to enable machines as well as people to understand, share and work with meanings. The implementation of Semantic Web solutions requires the production of structured information that complies with representation standards, over and above the features provided by HTML, databases and XML solutions. This need is solved by a set of technologies - Semantic Web Technologies – organized into layers and based on open W3C standards, the *W3C architecture stack* (W3C, 2007), where each layer exploits and uses capabilities of the layers below (Figure 25).

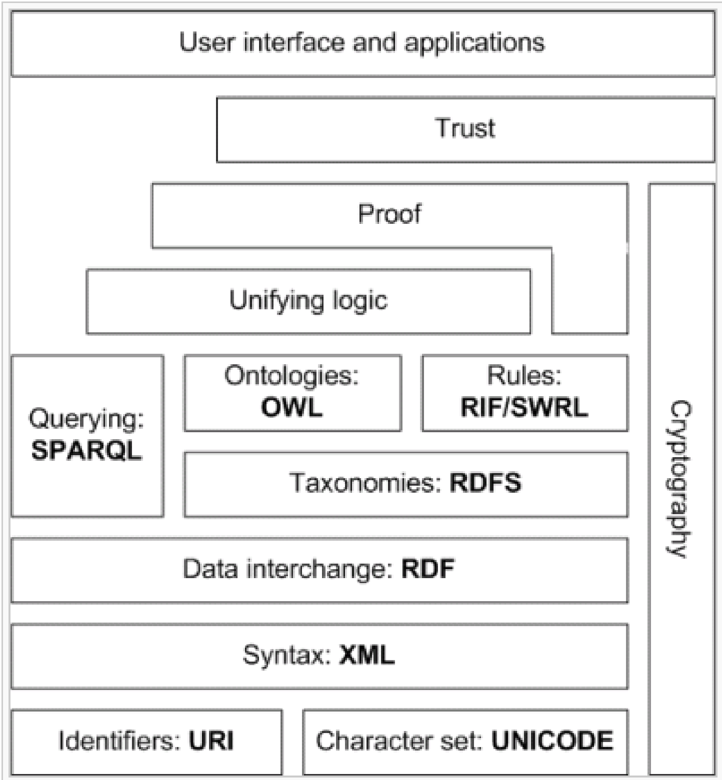


Figure 25 Semantic Web stack layer cake (Extracted from:W3C architecture stack 2010)

The lower layer (the symbol/reference layer) contains technologies that are well known from the hypertext Web and that provide the basis for the Semantic Web. Examples of these technologies are: the Uniform Resource Identifier (URI), used to unambiguously identify a resource in a standard (uniform) way; Unicode, an ISO standard that provides a common representation and technical encoding for text in any language.

The syntax layer (next layer up) is used to serialize the language syntax and includes the following technologies, used to describe not only documents and the links between them, but also every arbitrary thing contained in a digital document:

- As described in section 3.4.5, XML is a standard markup language that provides a grammar and syntax for representing structured data in text-based form. As elements in different XML documents might have the same name (for example the "title" tag belongs to the HTML schema and to the DTBook schema as well), users can define a namespace (xmlns) for an XML schema in order to reuse the tag in another context. In this case, they will use the namespace to disambiguate this tag apart from other tags with the same name.
- Resource Description Framework (RDF) is a W3C standard that defines the base language used in the Semantic Web for describing data, metadata, and even other data languages in Web resources. RDF gives a formal definition for the interchange of an interconnected set

of data distributed at global scale across the Web (Linked Data). RDF uses a graph data structure, in contrast to the relational data structure (such as most databases) and to the hierarchical data structure (such as XML). RDF graph is based on the idea that every data item should have an URI and that every data item can be connected to every other item. RDF makes URI relationships between data items the central attribute of the overall data model. Semantic Web programmers create data with URIs and link them together using relationships that are also named with URIs. In short, RDF model is used to describe a "thing" (the resource you want to describe) by making assertion about its properties. The properties of the resource make up the description of the resource. The RDF conceptual model consists of the following fundamentals:

- An RDF graph is a set of RDF triples.
- An RDF triple, called *statement* (AKA *fact* or *assertion*), has three components:
  - a RDF subject, (the resource the statement describes), which is a RDF URI reference or a blank RDF node;
  - a RDF predicate (the property of the resource described), which is a RDF URI reference;
  - a RDF object (the value of the property), which is a RDF URI reference, a blank RDF node or an RDF literal. An RDF literal can be a RDF plain literal - a character string with an optional associated language tag describing the language of the character string - or a RDF typed literal - a character string with an associated RDF datatype URI. An RDF datatype defines the syntax and semantics of a set of character strings that represent data such as Booleans, integers, dates, etc.

A graph representation of a basic RDF triple is the following:

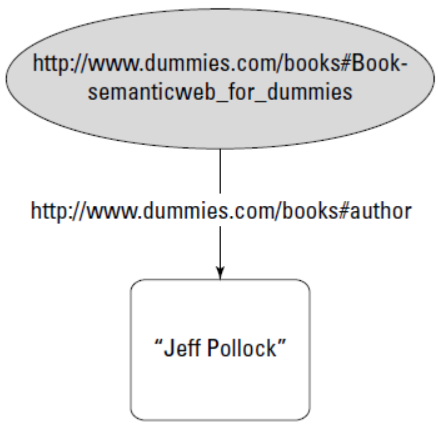


Figure 26. Example of a basic RDF triple. Extracted from (Pollock, 2009).

The statement represented in this graph is: the "Semantic Web for Dummies" book (subject) has an author (predicate), Jeff Pollock (object):

- The resource is an RDF URI reference: http://www.dummies.com/books#Book-semanticweb_ for_dummies.
- The property is a RDF URI reference: http://www.dummies.com/books#author.
- The value of the property is an RDF literal: Jeff Pollock.

These RDF triples are connected into RDF graphs, so that a non-literal RDF object of one RDF triple may be the RDF subject of another triple, as in the example below:

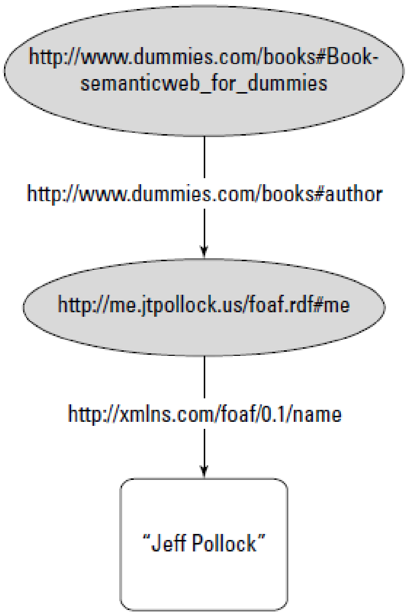


Figure 27. Example of a basic RDF triple connected into RDF graphs. Extracted from (Pollock, 2009).

RDF can be shared by serialising the RDF graph as XML. In this case XML is used as a standard format designed for exchanging a RDF model, though there are other alternatives like N-Triples or Turtle.

The Ontological primitive layer (next layer up) is composed by RDF Schema (RDFS), which provides a standard way to define vocabularies that describes RDF classes and properties (predicates). It allows you to group the RDF data into more complex sets that can be organized and queried via different query languages. RDF and RDFS are the foundation of the Semantic Web.

The Logical layer (next layer up) includes:

- The Web Ontology Language (OWL), which supports formal semantics and reasoning. It gives to the Semantic Web an element of grounding and stability for defining the meaning of data in an unambiguous powerful data model. For example, OWL enables the kind of formal semantics to express in a data model a piece of logic like, "A backpacker's destination is the intersection of all destinations that have budget accommodations and some type of sports or adventure activities"(Lacy, 2005). OWL is a language for encoding ontologies into files. The term *ontology* historically originated in philosophy and metaphysics and referred to studies of the science of being. Ontologies provided an exhaustive classification of objects and their relationships in all spheres of being. Nowadays, computer science community uses the term "ontology" in the context of information sharing to refer to formal descriptions of particular domains. An ontology defines concepts and relationships between concepts used to describe and represent an area of knowledge (Herman, 2009). It is considered a shared and common conceptualization of a domain that can be communicated between people and heterogeneous and distributed application systems"(García, 2010). An ontology is a formal specification of a conceptualization (Gruber, 1993):

- Formal, because it is an abstract model of a portion of the world.
- Explicit specification, that is machine-readable and understandable.
- Of a shared conceptualization, because is based on a consensus and is expressed in terms of concepts, properties, etc.

Beyond this definition, the term "ontology" has another meaning and indicates the vocabulary of declarative formalisms describing a model of a domain. In this study, the term "ontology" integrates both meanings and is used as a "specification of the conceptualization and the corresponding vocabulary used to describe a domain" (van Harmelen & Fensel, 1999). Although the most common form of ontology is taxonomy, which is easier to understand than more complex representation schemas, taxonomies lack the expressiveness needed to describe complex relationships. A taxonomy typically provides a hierarchy of concepts related with a specialization ("is-a") relationship and is normally represented as a tree that has a single root node (which is easily represented by a common database). Ontologies can be used to describe the semantic structure of much more complex objects and therefore they are well-suited for describing heterogeneous, distributed and semistructured information sources such as those commonly found on the Web. The purposes of an ontology are:

- To communicate understanding of a domain. Various communities of interest provide authoritative descriptions of their domain with the aim to reduce misunderstandings.
- To declare explicit semantics. Explicit semantics reduce ambiguity generated by the documentation of a domain's concepts with representational (or modeling) primitives and semantic relationships.
- To support information sharing. By sharing an ontology, applications can build a shared understanding of conforming information.

As a vocabulary, an ontology declares a set of specific terms and at the same time expresses formal definitions of the concepts and the relations between concepts (they are "self-describing") that describe a model of a domain. It typically consists of a comprehensive set of concept classes, relationship between them, and instance information showing how the classes are populated in the application domain:

- Individuals (sometimes called instances). An individual represents the class object instance in the described domain. Individuals may be members of one or more classes. They represent physical or virtual concepts that the ontology is describing. Examples of physical individuals include Mark, instance of person, MyApplePie, instance of pie, etc. Virtual individual include PhoneNumber, Order557, etc.

- Properties (sometimes called attributes or slots or roles). A property is a binary association that relates an object (individual) to a value and describes the relationships between individuals. Examples of properties include price, size and name.
- Classes (sets, collections, concepts, types of objects, or kinds of things). A class represents a group of individuals that share some properties or characteristics. Examples of classes could include Food, Menu Item, Person, Pie and Restaurant.

Semantic relationships are relationships within classes, properties and individuals that make expressive statements about the domain model. Examples of relationships are synonymy (connects concepts with similar meaning); antinomy (opposite concepts); hyponymy (describes a specialization or generalization) and meronomy/holonomy (represent the concept of aggregation/composition). The definition of those relationships allows for a better and automatic interchange of data. RDF gives a formal definition for that interchange.

- The Simple Protocol and RDF Query Language (SPARQL) is a query language for RDF. It can be used to query any RDF-based data.

The upper elements in the Semantic Web stack layer cake include technologies that are still immature.

Besides the technologies described previously, currently there are specialized languages that allow the encoding of semantics within existing content in Web pages by annotating HTML elements with machine readable tags.

4.6.2 Promising solutions for semantic markup of images in HTML

This section presents Microdata and RDFa, two promising solutions for extending HTML syntax to create machine-readable semantic markup about different types of objects, including images. Microdata and RDFa are compared to the main techniques offered by digital accessibility explored in section 4.3 as possible solutions for providing text descriptions to images.

4.6.2.1 Microdata

Microdata standard is a Web Hypertext Application Technology Working Group (WHATWG) specification that uses new attributes in HTML tags to describe the meaning and facilitate the discovery of specific types of information, such as photographs, movies, events, music, organization, places, products. It was proposed by Ian “Hixie” Hickson, the editor of the HTML5 specification, as an alternative to RDFa (see below), since RDFa was for a long time threated not to be incorporated into HTML5. The Microdata standard was primarily designed to offer an easy-to-implement syntax for page authors.

Microdata terminology considers things described in an HTML page as *items*. Each item is made up of one or more key-value pairs: a property and a value. Its syntax is made up of three new HTML attributes, which can be used on any valid HTML element (often <span> or <div>):

- *Itemscope*, which is used to create a new item.

- *Itemtype*, which specifies the type of item.
- *Itemprop*, which adds properties and values to the item.

A simple example of microdata code is the following:

```
<div itemscope itemtype="organization">
  <span itemprop="eponym">University of Barcelona</span>
</div>
```

This snippet is inserted in a HTML page and a user browsing the page would only see the text "University of Barcelona". However, the code provides the following machine-readable assertion: there is an "organization" with the "eponym" "University of Barcelona."

The main strength of Microdata is that it is endorsed by the most widely used search engines (Google, Bing, Yahoo and Yandex). These companies have created a common set of schemas for structured data markup on Web pages, *schema.org* (Google, Inc. et al., 2011), with the aim of making the meaning of web content understandable by search engine spiders and improves the display of search results. *Schema.org* defines a broad, web-scale, shared vocabulary focusing on popular concepts that the promoters of the initiative consider will receive a special treatment through search engines in the near future. Search engines are the main "users" of Schema.org data and have a stated preference for Microdata (Ronallo, 2012). Below, the example of Microdata provided previously is rewritten to use the *Organizationtype* vocabulary of *Schema.org*. In this case the information is efficiently processed by search engines.

```
<div itemscope itemtype="http://schema.org/Organization">
  <span itemprop="name">University of Barcelona</span>
</div>
```

If we want to describe an image in a webpage by using Microdata syntax and schema.org vocabulary, we can use the itemtype called *imageobject,* which is a more specific type of *mediaobject*. *Imageobject* is defined by a set of properties that can include caption and EXIF data associated to the image, a thumbnail for the image and a *representativeOfPage* property (which indicates whether the image is representative of the content of the page).

```
<div itemscope itemtype="http://schema.org/ImageObject">
<img itemprop="contentUrl" src="LIS_faculty.jpg" alt="UB logo"
/>
<span itemprop="caption">Faculty of Library and Information
Science of the University of Barcelona </span>
<meta itemprop="exifData" content="Resolution : 600 x 445" />
<meta itemprop="representativeOfPage" content="true" />
<meta itemprop="thumbnailUrl" content="thumbnail.jpg" />
</div>
```

An advantage of schema.org is its extensibility, since it offers the possibility to specify additional properties or sub-types to existing types. Recently, a set of accessible metadata proposed by the Accessibility Metadata Project (Benetech, 2013) has been included into schema.org. It enables search engines to index information about the accessibility of a resource (e.g., a video, e-book or other digital publication) without ambiguity and thus makes that resource discoverable by its accessibility attributes. A scenario for the application of the schema is when a blind user may want to locate a text-book that is available with particular accessibility features and formats, or a user with hearing impairments may need to locate a captioned version of a video. An example of application of the schema to the description of a biomedical academic article in PDF format with images and text description is detailed in the following code (adapted from (Myers & Capiel, 2014).

```
<div itemscope
itemtype="http://schema.org/MedicalScholarlyArticle">
<meta itemprop="accessMode" content="textual"/>
<meta itemprop="accessMode" content="visual"/>
<meta itemprop="accessMode" content="textOnImage"/>
<meta itemprop="accessibilityFeature"
content="alternativeText"/>
<meta itemprop="accessibilityFeature"
content="longDescription"/>
</div>
```

The *accessibilityFeature* property describes the content features of the article that make it accessible to a broader range of users. The property values indicate that the image has a short alt text (“alternativeText”) and a longer description for the image (“longDescription”). The *accessMode* property values indicate that the document contains text (“textual”), images (“visual”) and scanned images of text (“textOnImage”).

A limitation of Microdata is that currently Microdata items describe a single web page and they are not linked to other items across the Web and they cannot be independent outside the page (Ronallo, 2012).

4.6.2.2 **RDFa (RDF in Attributes)**

Resource Description Framework Attributes (RDFa) is the standard set of extensions to XHTML proposed by the W3C for embedding RDF-based data in XHTML documents. The current version (RDF 1.1) can be used with XML document types and non-XML document types, such as HTML 4 and HTML 5. RDFa Lite 1.1 (W3C, 2012a) is a simplified version of RDFa that includes a reduced set of attributes (*vocab, typeof, property, resource* and *prefix*) and it is intended for authors who want to express fairly simple data in their web pages. In the next example (W3C, 2012a), the vocabulary provided by schema.org has been adopted to define properties of specific parts of information in a web page:

```
<p vocab="http://schema.org/" typeof="Person">
My name is
<span property="name">Manu Sporny</span>
and you can give me a ring via
<span property="telephone">1-800-555-0199</span>
or visit
<a property="url" href="http://manu.sporny.org/">my
homepage</a>.
</p>
```

The *Vocab* attribute specifies the vocabulary used. The *Typeof* attribute has the value “Person”, which corresponds to the type “Person” defined in *schema.org.* The *property* attribute has the value “name”, which is a property of the type “person” in *schema.org* (such as *@telephone* and *@URL*).

A possible application of RDFa is the semantic annotation of an image on the Web, beyond the possibilities that X(HTML) tags usually offers. For example, the following syntax can be used to specify the license and the date of creation of an image:

```
<img src="image1.png" rel="license"
href="http://creativecommons.org/licenses/by-sa/3.0/"
property="dc:created" content="2012-05-05" />
```

The *rel* attribute indicates the reference to the Creative Common Licence and the property *created* (defined in the Dublin Core vocabulary) specifies the date of creation of the image. To identify a specific resource inside an (X)HTML page, RDFa adds the attribute *about*, that can have *rel* and *property* information attached to it. Thus, for example, the following syntax can be used to specify the license and the creator of the “CreativeCommons” logo linked from a web page (Birbeck, 2009):

```
<a about="http://www.slideshare.net/fabien_gandon/rdfa-in-a-
nutshell-v1" rel="license"
  href="http://creativecommons.org/licenses/by/2.5/"
  property="dc:creator" content="Fabien Gandon">
<img src="http://i.creativecommons.org/l/by/2.5/80x15.png" />
</a>
```

To add another property to the anchor, the solution is the creation of an element that contains the context in which all the RDFa operate, as in the example below:

```
<div about="http://www.slideshare.net/fabien_gandon/rdfa-in-a-
nutshell-v1">
 <h1>RDFa in a Nutshell</h1>
 <ul>
   <li>Author:
     <em property="dc:creator">Fabien Gandon</em></li>
   <li>Created:
     <em property="dc:created" content="2007-01-01">
     Jan 1st, 2007</em></li>
   <li>License:
     <a rel="license" href="http://creativecommons.org »
/licenses/by/2.5/"><img src="http://i.creativecommons.org »
/l/by/2.5/80x15.png" /></a></li>
 </ul>
</div>
```

In synthesis, the *about* property sets the context for all contained properties and relationships. If there is no *about* value set, then all properties and relationships will be in reference to the current page. This example shows that RDFa embeds structured data in HTML as part of a graph which links items together across the Web. This is the main difference between RDFa and other methods such as Microdata.

The principal strength of RDFa is the ability to use properties from clearly defined vocabularies, using prefix mappings (indicated in the namespace where the vocabulary is specified), such as Microdata does. However, this mechanism is more complex than in Microdata, thus the search engines behind schema.org have currently given preference to Microdata over RDFa or other formats (Ronallo, 2012).

As emphasized by Bizer (2013), the adoption of specific formats and vocabularies seems to depend mainly on the major consumers the data is targeted at: Google, Facebook, Yahoo! and

Bing. For example, Microdata is dominated by the vocabularies promoted by Google, Yahoo! and Bing via schema.org, while RDFa data is dominated by the Open Graph Protocol vocabulary promoted by Facebook. The Open Graph protocol (Facebook,2014) enables the integration of external content into the social networking platform. To enable any web page to have the same functionality as a Facebook page, <meta> tags in the <head> are added to the web page by using RDFa. The Open Graph protocol uses its own vocabulary with specific properties (*og:title*, *og:type, og: image*, *og:url* are the basic ones).

Microdata and RDFs have currently considerable adoption on the Web. However, the topical domains of the published data mainly concerns persons and organizations, CMS-related and navigational metadata and product data (Bizer et al., 2013). Their use for describing the content of images on the Web, in general, and in academic articles in particular, seems not to be relevant.

4.6.3 Image description and annotation in Semantic Web

Semantic Web technologies can be used in a variety of application areas, such as: *data integration*, whereby data in various locations and various formats can be integrated in one, seamless application; *resource discovery and classification,* to provide better, domain specific search engine capabilities; *cataloguing,* for describing the content relationships available at a particular Web site, page, or digital library (Herman, 2009). All of these applications can be useful when the resource that should be semantically annotated is a static image.

From the point of view of Semantic Web, the issue of how to textually represent the information conveyed by an image is interpreted as the problem of making the implicit knowledge captured in images more easily processable by machines and, consequently, available in other forms. Knowledge (or information) is stored in implicit form and "the encoded knowledge can be completely lost if the receiver agent cannot understand it. For instance, if two agents exchange a written message but the second one does not understand the used language, nothing of the original codified knowledge can be retrieved" (García, 2010). The same statement is applicable to images. Images have implicit knowledge that is expressed using a visual language. People with visual disabilities and machines lacking computer vision capabilities are excluded from accessing this "tacit knowledge". In order to overcome these limitations of the visual language, it is necessary to make the knowledge embedded in the image explicit and provide alternative ways to access it, taking into account the perception and cognition capabilities naturally required. Some of these alternative representations might be based on formal approaches that avoid ambiguity and facilitate computer-based processing. Semantic Web technologies are powerful means of expressing the semantics of an image in a systematic representation and precise definition of the information behind images. Far beyond the metadata description detailed in section 4.5, they enable the conversion from the descriptions of images to formal representations of the semantics of the image. The standardization and unification of formal representations of the image's content offers many advantages: provide an integration of information about the image from different resources in various locations and formats; improve the ability to search, classify and display images according to their syntactic and semantic content; makes more powerful the cataloguing process of visual resources; facilitate the automatic generation of alternative representations of the image. Examples of use cases of semantic annotations in images are, among others: to manage personal digital photo collections, cultural heritage, television archive, large-scale image collections at NASA and biomedical images (W3C Incubator Group Report, 2007); to comment images for educational purposes (DIAGRAM Center, 2014a); to point out or highlight some aspects of the image (Thyssen, 2012).

Semantic information can be added to images by using tags, provided by individual manual indexers, automated keyword generators or collaborative-built folksonomies or by using predefined ontology concepts. It is even possible to combine both ways. They make explicit the syntax and semantics of the metadata related to the images.

The next example (W3C Incubator Group Report, 2007) could be useful for understanding the application of the Semantic Web in image annotation.

Figure 28 Example of image annotated by using the Dublin Core Metadata Element Set. Extracted from: (W3C Incubator Group Report, 2007)

The image in Figure 28 is annotated using RDF and the Dublin Core Metadata Element Set. The RDF/XML example code represents the statement: "There is an image *Ganes.jpg* created by *Jeff Z. Pan* whose title is *An image of the Elephant Ganesh"*.

```
<rdf:RDF xml:base=http://example.org/
xmlns=http://example.org/
xmlns:dc=http://purl.org/dc/elements/1.1/
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">

<rdf:Description rdf:about="Ganesh.jpg">
  <dc:title>An image of the Elephant Ganesh</dc:title>
  <dc:creator>Jeff Z. Pan</dc:creator>
</rdf:Description>
</rdf:RDF>
```

In this example the meaning of all metadata values is expressed in informal strings, which cannot easily be interpreted by automated procedures. The use of the ontology to describe this image permits to create a formal, machine processable description of the image based on a formal agreement of a community of users.

For example, suppose that the community interested in the picture represented in Figure 28 has agreed to use the concepts from the WordNet (WordNet search 2012) ontology to annotate the image. In the following example, the two lines of RDF (added under the *dc:creator* line of the example above) refer to the formal definition of *Indian Elephant* and *Ganesh* in WordNet:

```
<dc:subject
rdf:resource="http://www.w3.org/2006/03/wn/wn20/instances/synset
-Indian_elephant-noun-1"/>
<dc:subject
rdf:resource="http://www.w3.org/2006/03/wn/wn20/instances/synset
-Ganesh-noun-1"/>
```

Thanks to the reference to the ontology, “even applications that have never encountered these concepts before could request more information about the WordNet definitions of the concepts and find out, for example, that *Ganesh* is a hyponym of the concept *Indian Deity*, and that *Indian Elephant* is a member of the *Elephas* genus of the *Elephantidae* family in the animal kingdom”(W3C Incubator Group Report, 2007). This example makes clear how the ontology makes available formal description of the semantics of the image.

4.6.4 Contribution of semantic technologies to image description and accessibility

It is generally acknowledged that developments connected to the Semantic Web area can bring about substantial progress to the creation of accessible content on the Web (Burzagli & Gabbanini 2009). The ability of Semantic Web technologies to add semantic information to Web content (“document enhancement”) represents an opportunity to provide information readable by assistive technology and improve the user experience of people with disabilities in accessing web content (Harper & Yesilada 2012). Machine-understandable information provided by Semantic Web technologies has therefore the potential of enhancing the accessibility of Web content, including images. In particular, semantic annotations have been recognized as a technology that can be used to reveal the knowledge encapsulated in images in educational digital textbooks (Capiel, 2014) and in other different scenarios (Ribera & Splendiani, 2014), especially for improving the access to the visual content of images by people with visual impairments. According to the scope of the thesis, Semantic Web technologies can be used to facilitate image annotation and generation of alternative representations to images, besides image retrieval and reuse. In the next two sub-sections:

- the Web Semantic mechanism described in section 4.6.4 (microdata and RDFa) are analyzed as possible solutions for adding text alternatives to images and compared to the solutions offered by digital accessibility.
- A literature review of the proposals on the use of Semantic Web technique for the automatic generation of alternative representations of images is presented.

4.6.4.1 Comparison of Semantic Web and accessibility techniques for providing text descriptions to images

In section 4.3, the mechanisms currently offered by digital accessibility for providing text alternatives of images in general and in academic articles in particular have been described and compared according to the requirements of availability, discoverability (programmatic and visible) and structured markup.

In order to find the best solution for image annotation/description in academic articles in HTML documents, Microdata and RDFa solutions may be evaluated under the same requirements plus the extra requirements proposed by Dasiopoulou et al. (2011): granularity and use of annotation vocabulary. Furthermore, according to the scope of the thesis, the author of the thesis suggests taking into account their current adoption in academic articles as a requirement for the evaluation. The complete set of requirements is:

- Availability: to provide a textual description of the image.
- Discoverability (programmatic): to allow discovering and accessing the text alternative by programmatic access.
- Discoverability (visible): to allow discovering and accessing the text alternative by a call to action visible for sighted people.
- Structured markup: to support the inclusion of rich markup and semantic structures in the text description.
- Granularity: to support a text alternative describing the whole image or a part of it.
- Use of annotation vocabulary: to provide a text description connected to an annotation vocabulary (e.g. lexicon / thesaurus, taxonomy, ontology).
- Adoption: to be currently adopted in academic articles in HTML documents.

Table XI summarizes the conformance of the main mechanisms offered by digital accessibility (*alt*, *longdesc*, caption) and Semantic Web (Microdata and RDfa) to the requirements.

Table XI Techniques provided by the accessibility (A11y) and the Semantic Web (SW) disciplines and their conformance to the discoverability, structures, granularity (levels), use of vocabulary and adoption requirements. "Partial" conformance means that the technique conforms to the requirements but currently it is partially supported by browsers or screen readers. Note (1): The text description in RDFa may be visible to sight users by transcoding.

Requirement	**Technique**				
	alt	***longdesc***	**caption**	**RDFa**	**microdata**
Availability	yes	yes	yes	yes	yes
Discoverability (Programmatic)	yes	partially	partially	partially	partially
Discoverability (visible)	partially	yes	partially	partially (1)	no
Structure	no	yes	partially	yes	yes

Granularity (Levels)	One, whole image	One, whole image	One, whole image	One or more, whole image or region-specific	One, whole image
Vocabulary	no	no	no	yes	yes
Adoption	Medium	Low	High	Low	Low
Discipline	A11y	A11y	A11y	SW	SW

Compared to the accessibility mechanisms, Semantic Web solutions offer the possibility to semantically describe the content of images using a structured vocabulary. Furthermore, RDFa offers a mechanism for tagging not only the whole image but also specific parts of it, for example using the *ImageAnnotator* JavaScript (Milowski, 2012). Finally, the text description conveyed by RDFa may be visible to sight users by transcoding (Mirri, Salomoni, & Prandi, 2011). On the other side, their current adoption for describing the content of images is low, as commented previously.

4.6.4.2 Use of Semantic Web technique for the generation of alternative representations of images

As specified in section 4.2, an "alternative representation" is a different possible modality of presentation of an image, such as speech, text, alternative image, etc. The main aim of creating multi-modal descriptions for images is to allow the access to the semantics of the images to people with disabilities.

Many researchers have explored different approaches to make graphics accessible to blind or visually impaired people through the automatic generation of alternative representations by using semantic technologies, following two approaches:

- Bottom-up approach: the image is analyzed and interpreted by a predefined system and/or moderator, and then annotations are added to help understand the information.
- Top-down approach: this approach starts with the meaning behind the image instead of its graphical representation, and then generates one of its possible representations without losing access to the initial meaning.

Projects with bottom-up approach

The bottom-up approach starts from a graphical representation, attempts to infer or extract information from it and then generates the alternative presentation. The image is analyzed and interpreted by a predefined system and/or moderator, and then semantic annotations are added to help understanding the information. The task of inferring additional information of images and generate alternative representations from the original image (given by the author) can be accomplished by:

- Human intervention. A moderator, not necessarily the author, who provides a description of the graphical information, or a group of users collaboratively (folksonomies).

- Semi-automatic analysis. This analysis is based on the definition of the graphical elements of a graphic and the relations between them. The automatic interpretation and presentation is based on the process of analysis of the image's components, structured in a layered composition of different abstraction levels.
- A combination of both.

Compared to the techniques explored in the section dedicated to the generation of external metadata for image description, the following projects make use of Sematic Web technologies for the formalization of the semantic annotations extracted from the image.

Graphics Accessible To Everyone (GATE 2012) system aims to develop annotations of pictures and allow blind users to explore images. It offers an annotator connected to a *graphical ontology* (an ontology defining global properties representing important visual aspects of graphical objects) that describes the image by the use of graphics recognition algorithms. For instance, "if graphics recognition process recognizes a graphical object as a swimming swan, it annotates the object as "swan" and fills in the status attribute field with "swimming". The human annotator only checks the correctness"(Kopeček & Ošlejšek, 2008).

This approach is used in the TeDUB system (TeDUB 2004), in which "starting from the lowest level of geometric primitives like lines and curves, a logical inference mechanism is used to stepwise aggregate lower level syntactic constructs to increasingly domain dependent semantic units like "room" in the architectural domain or "pie chart" in a business chart domain. This aggregation process is performed according to a previously defined model for each domain" (Lorenz & Horstmann, 2004).

A limitation of this approach is that it requires writing difficult programs in order to "discover" the structure or semantics of the graphics and this is considered an onerous task for the author (Fredj & Duce, 2007).

Projects with top-down approach

Compared to the bottom-up approach, the top-down approach is based on a different concept of image. An image is not thought of as a set of connected graphical elements on a display (as ink on the paper or pixels on the screen), but as an abstract entity that has intrinsic structure and semantics. Starting from this theoretical assumption, this approach manages to overcome the limitations of the bottom-up approach by making the information on the structure and the semantics of the graphics "part of the graphics". It starts from the meaning behind the diagram instead of its graphical representation, and then generates one of its possible representations without losing access to the initial meaning.

An interesting proposal based on this model is the Graphical Structure Semantic Markup Languages (GraSSML) (Fredj & Duce, 2007). The basic idea behind the GraSSML approach is that if the structure and semantics of diagram are made available at the creation stage, the graphics (defined as *smar*t) can be automatically transformed in different modalities. GraSSML is a group of languages organized into three levels: semantic level, structure level and presentation level (Figure 29). The general model of GraSSML generates presentations in different modalities from a predefined domain specific XML markup language. The process of the image generation with GraSSML starts from the semantic layer, where a domain expert first studies the domain and the class of the diagram with the aim of identifying what concepts, properties and rules the author(s) of the diagram seeks to represent in the domain. Therefore, he determines what fundamental information needs to be provided by the author of the diagram in order to have a full understanding of the information conveyed. Then, the domain expert embodies these concepts, properties and rules in an ontology expressed in OWL and finally creates a domain specific markup language expressed in XML (MyLanguage). At the next level (Structure level), the generic language ZineML (not domain dependent) is used to capture the structure of the diagram. The notational conventions that control the diagrammatic representation of the elements of MyLanguage are captured in rules that govern the generation of ZineML. These rules, provided by the domain expert using an appropriate authoring tool, are expressed in XSLT. At the last level (Presentation Level), a graphical representation of the diagram is generated using SVG or other modality-specific languages such as XHTML. At the same time a textual representation is provided automatically by extracting information from the structure of the diagrams and the semantics behind the diagram. At this level the system allows the creation of a query system, which offers the possibility to interactively explore the diagram and formulate queries concerning specific parts of the diagram. At this stage adaptive presentations of the diagram are offered (textual, graphical and query- based).

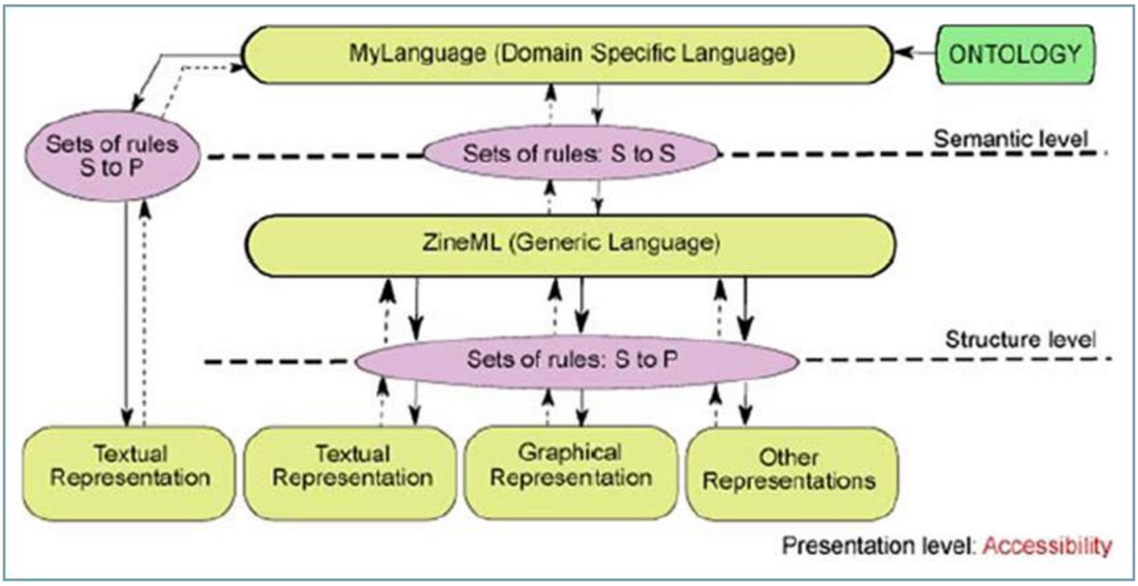


Figure 29 Conceptual architecture of GraSSML (Fredj & Duce, 2007)

Concerning the limitations of the proposal, the model is limited to a specific type of image as defined by the class of "structure diagram" in the taxonomy proposed by Lohse et al. (1994). Furthermore, the model does not propose solutions on how to provide a textual description to graphics already created (Demir et al., 2010), such as the images commonly managed by authors in the academic publishing workflow.

Lastly, the proposal takes only partially into account the knowledge offered by accessibility discipline about what kind information is needed and how to describe graphics. This information will be reviewed in the section dedicated to the proposal from accessibility discipline about the content of text alternatives.

4.6.5 Potential and limitations of the current solutions on image semantic annotation

In this section, the possibilities offered by the Semantic Web technologies for textually describing images have been explored and most promising solutions have been presented. In particular, the potential of solutions for improving image accessibility has been assessed. Semantic Web technologies may offer an opportunity to enrich alternative descriptions to images within the accessibility field, by providing structured and machine-understandable information readable by assistive technology and the automation of the process of assigning text alternatives.

The semantic annotation and enrichment of scholarly documents using Semantic Web technologies is currently a hot research topic, since it has the potential of providing rich data on the internal structures of the academic documents and their components (including images) and semantically connect them with other documents (Ciancarini et al., 2013).

However, Semantic Web technologies currently present some important drawbacks for being effectively adopted in the scientific image publishing workflow in the short-term. The management of the lifecycle of annotations and their related ontologies is a complex process in general and in the open Web environment in particular. As commented previously regarding metadata, also the semantic annotation task may be accomplished during different stages in the authoring and publishing process and using a wide range of file formats and tools, provoking a lack of interoperability. In general, as observed by Lord (Lord, Cockell, & Stevens, 2012), 333 the task of adding semantics into the current scientific publishing process, which is still based on the generation of articles converted from Microsoft Word or LATEX formats to a final PDF “which looks, feels and is something designed to be printed onto paper”, is difficult or impossible to achieve without a revision of the academic publishing environment.

4.7 Natural-language summarization of graphics

Natural language and statistical text processing techniques for generating text summarization of images showing data, such as photographs or other scientific visualizations, have traditionally been investigated to improve figure comprehension and retrieval in specific domains. The aim of summarization is to improve the efficiency and satisfaction of image retrieval: for example, a concise and relevant summary of an image saves end-users’ time when they are examining search results to find something that satisfies their information needs (Bhatia et al., 2009).

The use of natural language summarization to communicate information about numerical data and graphs is another alternative for conveying the content of graphics to people without sight. The use of statistical and natural language processing techniques for the automatic generation of text summarizations of graphics has been focused on the most common figures representing data, especially statistical graphs such as line charts and bar charts.

The Interactive SIGHT (Summarizing Information GrapHics Textually) system proposed by Demir et al. (2010) automatically captures the intended message of bar charts in graphics available on the Web and in popular media and conveys it to visually impaired users. The system is implemented as a browser extension and can be launched by a keystroke combination, which conveys a brief initial summary of a bar chart to the user and, if desired, responds to follow-up requests for additional information about the chart. It works by recognizing bar charts and analyzing their component (e.g., the height and color of each bar, the tick-mark values and textual components, such as any annotations on a bar or axis and the caption of the graphic) using image processing techniques. Then it creates an XML representation of the extracted components of the image, to which a Bayesian inference is applied for assessing the intended message of the graphic, such as a change in trend and the rank of an entity. Finally the message is organized into a coherent text and summarized into natural language, which can be acceded by JAWS in Internet Explorer.

The system could also be used to generate appropriate *alt* attributes, according to the evaluations performed by the authors with visually impaired users. However, it has the limitation of handling only bar and line charts in articles in popular media (such as magazines and newspapers), which have different characteristics form those in scientific graphics.

The iGRAPH-Lite System developed by Ferres (Ferres et al. 2010) was designed to make the information conveyed by statistical graphics in common applications (such as MS Excel) available to users by using natural language through key commands and a Text-To-Speech (TTS) engine. Its aim is to enable users to easily explore the graph by keyboard commands and infer its intended message given the user own needs. The main advantages of the system are:

- It is focused on providing access to scientific graphics;
- It allows an access to the point-level information of the graph;
- It was developed after several studies involving professional statisticians and statistical communication officers interacting about graphs with blind and visually-impaired individuals, decanting the vocabulary most likely to help graph accessibility;

However, it does not provide the user with the high-level content of the graphic (the communicative message of the graphic) and the navigation system requires the user to build a “mental map” of the graphic, a task that can be difficult for congenitally blind users, since they have no personal knowledge regarding the appearance of information graphics (Elzer et al., 2008). In general, it is unclear how to scale results obtained with line charts and bar charts to other types of graphs (e.g., pie charts, scatter charts) (Abu Doush et al., 2010).

The figure summarization system developed by Agarwal & Yu (2009) automatically aggregates information from associated text of figures in biomedical articles, such as captions and mentions, to generate a structured text summary for improving the comprehension of the figures' meaning. The system extracts the most informative and relevant sentences from a full-text article that best describe a figure from each of the four rhetorical categories of the Introduction, Methods, Results and Discussion (IMRaD) model, a format for structuring the content of scientific articles commonly used in publications in the biomedicine discipline and in other research fields. The four sentences are extracted according to the similarity with the caption of the figure and the relationship to the central theme of the article. The result is a summary that textually describes the background of the figure, the methods used to obtain the figure, the outcome and the significance of the figure. Summaries generated by the system are available via "figuresearch" engine, a search engine that searches published medical literature to generate a list view of the results with relevant images, abstracts and summaries. The major problem of the extractive approach applied in this system is that it does not take into account the semantics of the sentences and certain sentences included in the summary are difficult to understand because they are taken out of context (Agarwal & Yu, 2009).

In synthesis, natural-language summarization techniques have been mainly focused on textually describing highly-structured graphics, such as statistical graphics. When they are extended to all kind of figures in biomedical article (such as in the case of the solution proposed by Agarwal), the automatically extracted text cannot avoid the limitation of being difficult to understand, since it is taken out of context.

4.8 Guidelines and best practices for describing images in accessibility

In the previous sections of this chapter (4.3- 4.7) , the issue of how to provide text alternatives to images in STEM academic articles has been addressed by reviewing the solutions offered by accessibility and other research fields. The mechanisms for assigning text alternatives to images and their potential for being integrated into the academic publishing workflow have been stated.

Besides these mechanisms, it is important to assess which information should be included inside a text alternative in order to make it effective. In this chapter the most relevant accessibility guidelines suggesting the content of text alternatives are reviewed, together with the best practices an author should follow to effectively writing them.

4.8.1 The complex task of creating effective text alternatives

As stated previously in this study, academic publications commonly include high semantic images that require long and detailed descriptions. The definition of an equivalent text for images in general, and for complex images in particular, is one of the most difficult issues to deal with in the accessibility field for several reasons:

- Images in STEM academic publications are commonly semantically dense and convey professional/technical information. They usually present a high "conceptual complexity" (RNIB, 2010b), since they represent complex subject matter (ideas, processes and objects) and require prior knowledge to understand them. They require specific technical expertise derived from learning and training to interpret or extract their meaning, and rely on detailed specialist knowledge and vocabulary (Burford, Briggs, & Eakins, 2003).

- The task is challenging and time-consuming, since it consists of including “as much information as possible in the fewest characters without sacrificing intelligibility” (Slatin, 2001). It has been estimated that it can take 50 minutes to describe in detail a moderately complex line graph (Open University, 2013).
- Personal interpretation: currently “determining appropriate, equivalent, alternative text is often a matter of personal interpretation” (WebAIM, 2012).
- It is contextual. Effective text alternatives depend not only on the content of the image (what the image is intended to represent) and its function, but also on the context in which they are set. According to the opinions of blind and visually impaired users, the context is “vitally important” (Tomashek et al., 2013), since text descriptions are preferred only in cases in which the information is not redundant to the text. When the description provided in the surrounding text is sufficient, it is considered preferable by users who are blind or visually impaired to just provide a concise statement in the text alternative, such as "Figure 10 is a graphical depiction of the information just described" (Sacco, Corona, & Johannesen, 2004).
- Consider the audience. Different kinds of user have specific needs and even individuals in the same group can have different preferences concerning how a figure should be described by words. Although users with different types of visual impairments agree on the need for good text descriptions, they differ on the content (Tomashek et al., 2013). The lack of agreement includes:
    a. Amount of information provided in the description. Too much information seems to be not good, but it is preferable to the lack of description.
    b. Level of information. Some individuals prefer more basic description that does not convey the meaning of the image, in order to be able to form their own interpretation. However, the kind of information provided depends on the current and past level of vision. For example, when congenitally blind users access the textual description of an image, they build a personal “mental map” of the graph. This task can be very difficult for them, since they have no personal knowledge of the appearance of information graphics (Elzer et al., 2008). Taking into account this case, it could be more useful to provide the overall meaning of the image. Visual descriptive words are another example of information presented in a text alternative that should be tailored according to the audience: it is often meaningless for individuals with congenital blindness, while it may give important information to someone who has an acquired visual impairment.
    c. Type of image. Text descriptions are considered essential for certain type of images, such as maps, graphs, math and science images, pictures of products for online shopping, icons and on-screen buttons.

Taking into account the previous observations, current proposals from accessibility discipline on how to write efficient text alternatives to images are reviewed in this section. Practical tools for improving the decision process of choosing accessible alternatives to images and for simplifying the creation of text alternatives are also presented. Strong and weak points of the proposals are finally commented.

4.8.2 Current proposals from accessibility discipline on image text description

The tools and resources helping authors in the creation of effective text alternatives include:

- A set of recommendations. The recommendations provide information about which content should be included in a text alternative in order to make it effective. They can be applied to all types of images or targeted to specific types of images.
- A decision tree. The informative process of choosing the appropriate text alternative according to the function of the image and its context can be schematized in a decision tree that can used to orient authors in the choice of the appropriate alternatives.
- Fill-in templates. A template of basic information with some gaps to fill with the particular information of the graphic to describe is offered to the author of the image.
- Image samples. The author receives examples of alternative textual descriptions targeted to specific types of images similar to the one he is currently describing.

4.8.2.1 Guidelines of official accessibility guidelines

WCAG 2.0 techniques provide general recommendations concerning the creation of short and long text alternatives of images. They offer a set of questions addressed to authors for deciding what text to include in the short and long alternatives (W3C, 2014k)(W3C, 2014d): "Why is this non-text content here?; What information is it presenting?; What purpose does it fulfill?; If I could not use the non-text content, what words would I use to convey the same function and/or information?". They also present some general examples of application.

The techniques for providing useful text alternatives in HTML5 (W3C, 2012a) offer general recommendations on good practices:

- "Provide the same informational content as the image.
- Where an image performs a specific function, such as a graphical link, provide information about its functionality.
- Be succinct as possible while still conveying equivalent values. Short text that describes its purpose or gives an overview will often suffice.
- Write suitable *`alt`* text according to context. The same image in a different situation may need very different `alt` text.
- Avoid redundant *`alt`* text. An example of this would be repeating the same text in your document, as well as in the `alt` attribute, and is unnecessary".

Section 508 (US Section 508, 2002b) recommends to provide a text equivalent for every non-text element via "alt", "longdesc", or in element content. It suggests the Social Security administration guidelines (Social Security Administration, 2010) as technical guidance for the creation of text alternative for images (US Section 508, 2011). The contribution of these guidelines will be described below.

Besides these official recommendations, several organizations and institutions in the accessibility area have proposed guidelines to formalize the creation of alternative texts, as those reviewed below.

4.8.2.2 Effective Practices for Description of Science Content within Digital Talking Books

The National Center for Accessible Media (NCAM) is a part of the Boston public broadcaster WGBH's Media Access Group addressing barriers to media and emerging technologies for people with disabilities in their homes, schools, workplaces and communities. "Effective Practices for Description of Science Content within Digital Talking Books" (NCAM, 2009a) are the results of a 4-year effort of NCAM to provide best practices for the description of STEM images. The aim of the project is to improve the access to diagrams, photos and other STEM images in DTBs by student or scientists who are blind or visually impaired. These best practices summed up the opinions collected from 30 STEM professional describers and students/scientists with vision loss. They were also tested with 60 focused adults with blindness or visual impairments.

The guidelines proposed by NCAM are integrated by a set of useful examples that clarify the application of the descriptions to different types of graphics, such as bar chart, scatter plot, standard and complex diagrams, illustrations and math equations. The guidelines offer general recommendations and target specific types of figures also. The general recommendations can be summarized in the following principles:

- Brevity: provide important information with few words as possible.
- Data: description should focus on the data and not on extraneous visual elements.
- Clarity: avoid description poorly written or confusing description.
- Drill-down organization: a description might begin with a title, followed by a brief summary and then continue with details and data.

These high-level principles can be applied to the following types of image:

- Bar charts (vertical and horizontal) and line graphs. They should be presented as accessible tables (properly coded with caption, table headers and table data) with a brief summary, not as narrative descriptions. Title and axis labels should be provided in the description. The description of visual attributes is not necessary (e.g. dark blue), unless there is an explicit need, such as an exam question referring to the colors. The description of simple horizontal bar charts can be alternatively provided with lists.
- Venn diagram. Focus on the data, which should be provided in a brief statement.
- Scatter plot are hard to describe. The image should be identified as a scatter plot and the description should focus on the change of concentration. Title and axis labels should be provided in the description. Optionally, data may be converted into an accessible table if it is necessary to provide detailed information.
- Pie chart should be converted into accessible tables. The description of visual attributes is not necessary (e.g. dark blue), unless there is an explicit need such as an exam question referring to the colors. The description of simple horizontal bar charts can be alternatively

provided with lists. A helpful technique is to list the numbers from smallest to largest, regardless of how they are presented in the image.

- Flow charts are difficult to describe. Simple flow charts can be converted into nested lists with good results. The "boxes" can be presented as numbers and the possible transitions as sub-bullets. For flow charts with multiple starting points, it is useful to begin with a brief overview, explaining the distinct elements of the image and then to describe each section of the flow chart linearly, using nested lists. An alternative approach may be to explicitly state when the different lines merge. A chart with complex paths should be converted into a single linear list, with the possible next steps of each box or item nested below it.
- Standard diagrams and illustrations. If presented with a caption, the caption should be included in the description. Common and useful terminology should be used and information separated in easy to scan bundles by using brief sentences, line breaks, or even bullet points or lists. When the illustration embeds a graph, the use of drill-down organization (brief summary, then details) is recommended. If necessary to provide all the data, the graph can be converted into an accessible table. If the graph includes numbers (for example enclosed in parentheses) that are not explained in the caption or the surrounding text, they should be included into the description but without guessing or assuming their meaning.
- Standard math diagrams (for example, a figure illustrating the reflection and refraction of light). Short sentences that focused on the data organized in a linear fashion are recommended.
- Complex diagrams and illustrations. For this kind of images, two examples are provided:
    - One illustrated chart depicting the characteristics and relationships of some birds. In this case, the recommendation is to replicate the information about the relationships in a table.
    - One colorful illustration representing the carbon cycle. Despite the complexity of the visual representation, its main purpose is to show the storage of carbon and fluxes in Carbon through Earth's atmosphere, oceans and land. The process conveys data that can be schematically represented by tables.
- Math equations: Math equations should be rendered in MathML.

4.8.2.3 Description Image Guidelines for textbooks (Bookshare)

Some recommendations provided by Bookshare description Image Guidelines (Benetech, 2012) are specific for creating descriptions of images to blind students. They provide some suggestions not included in the NCAM guidelines:

- Describe only what you see objectively. Do not see motivations or intentions and do not editorialize, interpret, explain, or analyze the material.
- Go from general to specific, in a similar way to the drill-down organization.
- Color: Since color has emotional connotations even for blind students and many students may be partially or previously sighted, the inclusion of color references enhances their ability to better imagine a scene. Color descriptions should therefore be included for most

images (e.g. the blue sky or the green lizard). The exception is for images such as charts or maps (e.g. the blue bar or the red line) where color is irrelevant to the teaching concept, and can be ignored in most cases.

- Language. Use words that are succinct, vivid, and imaginative to convey visual images.
- Narrative tense. Write in present tense, using active voice (e.g., “Ted breaks the window,” is preferable to, “The window was broken by Ted.”) Use third-person narrative style to show neutrality and noninterference.
- Emotions. Instead of describing emotion, such as suspicious or furious, write what you see such as: “she raises her eyebrows”, or “her fists are clenched”.
- Avoid visually related words, such as “we see”, “shown below”, or “can be viewed”. Preferable words include “depicts” and “indicates”.
- Pronouns. Use pronouns only when it is clear to whom or what the pronoun refers.
- Metaphor/simile. Describe shapes, sizes, and other essential attributes of images by comparison to objects or items familiar to the intended audience. For example, “the room is as big as a bus” and not “the room is as big as a whale”, since a blind student will likely have had first-hand experience with a bus, but probably not a whale.
- Censorship: don’t censor information for personal reasons such as your discomfort with the material or political beliefs. Describers must objectively relay the visual elements of nudity, sexual acts, violence, etc.

4.8.2.4 UKAAF guidelines

UK Association for Accessible Formats (UKAAF) guidelines (UKAAF, 2012a) provides general principles similar to those proposed by NCAM. Relevant suggestions for describing images complementing the guidelines described previously are:

- Take into account the main purpose of the image and the level of previous experience of the reader.
- Decide if it is the data within the image or the appearance of the image which is important.
- If the data within the image is important (as in charts, graphs and diagrams) and the image summarizes data and other essential information, it is this data and essential information that needs to be described. Give the title and caption and provide information about the pattern represented, if there is any. Data can be provided by a table or a list. For figures included in a document, include any image reference, so that the description can be linked to the main text.
- When the appearance of the image is important, begin by providing the title, caption and references, which states what image is. Then briefly describe the general layout, that help the users start to build up an image in their mind. Then describe the relevant details and finally describe the story of the image, which states what the image is about. For example, if an image demonstrates how something works, for example a diagram of an engine, provide this information after the layout of the image. This information will be easier to understand once the layout of the image has been understood.

- Specific cases. For images where the visual appearance is significant – for example, in the case a reader is expected to be able to reproduce the image), then the appearance of the image will need to be carefully described.
- Terminology. Link the description to the text by including details referred to in the main text, if the image caption or title is part of the image file and cannot be accessed by a screen reader; then the caption will need to be repeated. Be aware of using visual concepts, as these may have little meaning for some people without sight.
- Punctuation. Use simple things like commas, inverted commas, quote marks, question marks and so on, as these help the synthesizer to use correct intonation.
- Symbols and numbers. Avoid symbols that may be interpreted differently or ignored completely on different systems such as smart quotes and special characters. Use words for these, for example 'pi' instead of 'π'. Ordinary numbers can be written as digits, for example '10', '435' but avoid combinations of letters and numbers, for example, write 'eighteen hundreds' rather than '1800s'.

Other several specific recommendations and examples are provided by UKAAF for line graphs, scatter graphs, pie charts, bar charts, flow chart, other schematic diagrams and cycle diagrams, photo, paintings and art illustrations, maps, math and chemical notations. They are not repeated her for being similar to those provided by NCAM.

4.8.2.5 Interactive Scientific Graphics - Recommended Practices for Verbal Description

Recommendations provided by Wolfram Research - a company dedicated to the development of software for communication and education in computation and math fields - and DIAGRAM Center are created for enabling the access to interactive scientific graphics to people with blindness and low vision (Keane & Laverentz, 2014). The guidelines are based on a review of the current practices of many organizations in the verbal and textual description of visual information conveyed by scientific static graphics and non-scientific dynamic graphics. Starting from the guidelines offered by six main sources (Section 508, HTML5 specifications, WebAIM, NCAM, *Description Key* and Audio Description Coalition), relevant high-level aspects that should be taken into account in the creation of descriptions to interactive graphics were defined. The author of the thesis considers that the *Description Key* and Audio Description Coalition contributions are included in the suggestions of this study and no further description is required. As the thesis focuses on still images, recommendations related to the interaction with dynamic graphics are not taken into account.

Guidelines included in Wolfram and not mentioned before are:

- Content. It should be accurate (“do not misrepresent information”), equivalent (“describe all information in graphic”), objective (“only describe information in graphic”) and essential (“only represent necessary information”).
- Vocabulary. It should be contextual (“use words from an appropriate STEM discipline”), common (“use common and researchable words”), appropriate (“use words that reflect

the intended audience’s knowledge”), consistent (“do not use multiple words to describe the same thing”) and unambiguous (“do not use one word to describe multiple things”).

- Phrasing. It should be clear (“information should be easy to extract”), concise (“use phrases that are as simple as possible”) and understandable (“repetition should be unnecessary”).

4.8.2.6 ISO/IEC 20071-11:2012 Guidance for alternative text for images

The ISO/IEC 20071-11:2012 standard (ISO/IEC, 2008) establishes a procedure for identifying the information being communicated within an image and provides guidance for the generation of descriptive and informative texts for images.

The procedure is composed of six steps:

- Step 1. Identify the purpose that the image served within the document.
- Step 2. Identify the image components within the image.
- Step 3. Identify the image (or image component) content.
- Step 4. Elaborate on the image (or image component) content.
- Step 5. Organize the identified information into text alternatives.
- Step 6. Evaluate the resulting text alternatives.

Step 5 includes the selection of the possible ways for providing text alternatives according to the purpose and the content of the image (short description, long description, caption, tooltip and document main text) and the formulation of the text alternatives, according to recommendations summarized as follows (Tang, 2012):

- State the purpose of the image.
- Provide general information followed by additional details.
- Describe objects in the image in a logical and consistent manner. For example, left to right or top to bottom.
- Be succinct and concise. Make use of vivid words.
- Identify the most importance piece(s) of information and communicate it first.
- If the description of the image is provided elsewhere, the text alternative should state where it can be found.

The procedure of creation of text alternative is guided by detailed questions that should be asked about an image and its components, according to an extended version of the *W5H* structure currently used by library cataloguing systems (*What*, *Who*, *Where*, *When*, *How Much*, and *How*) and 12 categories of information within images (*causal*, *conceptual*, *continuous action*, *descriptive*, *discrete action*, *event*, *physical*, *procedural*, *relationship*, *spatial*, *state*, and *value*).

Step 6 includes the evaluation of the text alternatives by both sighted and visually impaired users.

Compared to the guidelines previously described, the opinion of the author of the thesis is that this specification has many strong points:

- It is a standard.
- It is based on the contributions of different research fields (Web accessibility, library cataloguing, captioning and audio description, image retrieval and indexing, art description, and tactile representation) for the definition of the requirements for producing text alternatives.
- It is supported by the results of the evaluation of the text alternatives performed with sighted and visually impaired users.
- The guidance is incorporated in a usable online tool for creating text alternative (USERLab, 2012).

However, while the procedures of the guidance have been designed to be applied to a wide range of static image types within all types of documents, they do not provide guidance on specific image types, such as medical images, which presents specific content that cannot be found in generic photographs or illustration and in diagrams. For example, it is hard to identify and include in the alternative description of an X-ray computed tomography of a brain the information related to the acquisition modality by following the questions proposed by the guidance.

4.8.2.7 Guidelines of the Organización Nacional de Ciegos Españoles (ONCE)

The guidelines for the audio narration of DAISY books within the bibliographic service of the ONCE (Hojas, 2008) basically agree with the recommendations previously described, with an exception concerning photographs. Hojas recommends you not to describe them and, in some cases, even not to mention them to not interrupt the reading flow.

4.8.2.8 Social Security Administration (SSA) guidelines

The guidelines proposed by the Social Security Administration (2010) agree with the recommendations previously described. A specific recommendation that is worth to be mentioned is to avoid the use of the word “figure” in the text description since the type of content (in this case “figure”) is announced before by the screen reader. An exception should be made for charts and diagrams and the type of image should be explained at the start of the text alternative (“Pie chart of…”, “Organizational chart of…”).

4.8.2.9 Decision tree for choosing alternatives to images

The informative process of choosing when and what alternative should be provided to images can be schematized in a decision tree. An example of decision tree is reviewed in this section.

The “image sorting tool” developed by Touch Graphics (Landau, 2013) is a decision tree addressed to non-specialists to decide which images in textbooks require to be adapted and which adaptation should be used to make them accessible to visually impaired students. It helps in the tasks of choosing which kind of alternative (textual description or tactile graphic) should be provided to print images, as showed in Figure 30.

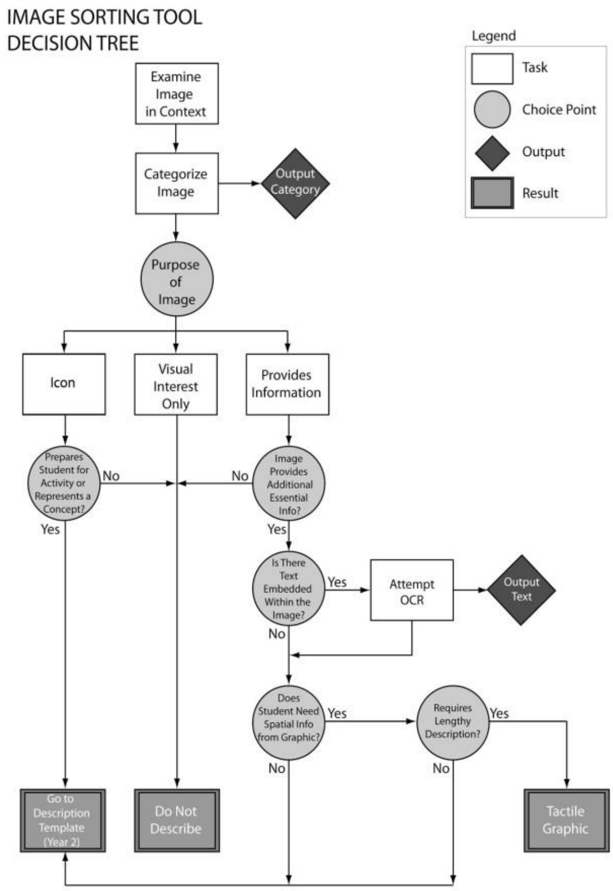


Figure 30 Illustration of the image sorting tool. Extracted from: (Landau, 2013)

The design principles provided by the Guidelines and Standards for tactile graphics of Braille the Authority of North America (Braille Authority of North America, 2011) are addressed to professional designers involved in the creation of tactile graphics from print illustrations. However, some recommendations show interesting parallels with some suggestions on the creation of textual alternatives for complex images previously reviewed. Tactile graphics for complex diagrammatic figures require special treatment before they can be readily understood by the reader. In order to establish the complexity of the diagram and decide what design treatment should be performed, the guidelines recommend designers to follow these decision steps:

1. Determine the main purpose of the diagram.
2. Determine if the diagram is an essential part of the surrounding text.
3. Determine if the original print illustration serves as a visual decoration or provides substantive information needed to understand or appreciate the text.
4. Color in a complex diagram may have contextual significance, but it may be unnecessary to produce it tactually.
5. Determine if the reader is required to retrieve data or make specific observations from the print illustration that would otherwise be revealed when attempting to describe the diagram.

6. Determine if the print graphic is required in a testing situation.

4.8.2.10 Templates

Templates for line graphs, scatter graphs, pie charts, bar charts and flow chart text descriptions are proposed by UKAAF guidelines (2012a). For instance, the structured template for the description of a line chart is the following:

> This line graph shows how {} changes/rises/falls with changes in {}. {omit this if it is obvious from the caption}. (1)
> The x or horizontal axis is labelled {label} and is marked in units of {}, from {} to {}, at intervals of {}.The y or vertical axis is labelled {label} and is marked in units of {}, from {} to {}, at intervals of {}. {If axes have no scale or labels, say so, e.g. The axes are have no scales marked}. (2)
> The line is {straight/curved/irregular} and starts at {describe position and describe movement}. (3)
> This {peak/trough/position} is labelled {label}. (4)
> The general trend is {upwards/downwards}. (5)
> {If the context requires numerical values not given in the main text, read them. If the graph is hard to read accurately, say so.}. Reading the values from the graph at {value of intervals on independent axis, usually x}, estimating to the nearest whole number: {include tabular information as a table in a Word document}. (6)
> {Include any additional information essential for answering SAQs etc.} (7)
> {The values marked on the graph are given in {} in the text.} (8)

The corresponding description elements are:

1. Type of image, brief description of the essential purpose.
2. Horizontal and vertical axis labels, units, highest and lowest value, intervals.
3. Type of line, starting value.
4. Labels.
5. General trend.
6. Numerical values.
7. Additional information.
8. Reference to the main text.

4.8.2.11 Image samples

Examples of accessible alternatives for graphical content in a digital book are presented in the "accessible image sample book" (DIAGRAM Content Working Group, 2014), a free online resource in EPUB3 and HTML. It contains examples of alternatives, including textual descriptions, targeted to specific types of STEM images (such a medical diagrammatic illustration and mathematical equations) and practical recommendations on how to implement alternatives in the EPUB3 code. The opinion of the author of the thesis is that a strong point of the book is the inclusion, among the factors for choosing the best suited alternative, of the characteristics of the intended public. For example, the grade level and the background knowledge and experience of the intended students in STEM disciplines to whom the adaptation is addressed. On the other side, the shortcomings of the book are considered the lack of recommendations about how to describe images and the fact that the use of SVG and image-related texts for the image description is not sufficiently explained.

Load2Learn, a project partnered by RNIB and Dyslexia Action and funded by the UK Department for Education, offers the ImageShare system (RNIB & Dyslexia Action, 2014), an online repository of accessible educational resources, including thousands of downloadable tactile with text descriptions and other alternative representations such as large print. The collection includes simple artworks, mathematics and charts.

4.8.3 Advantages and limitations of current recommendations on image description

Accessibility discipline currently offers many guidelines on how to select suitable adaptations to images and how to write efficient text alternatives to them. They provide general high-level principles on how to textually describe images and also present more detailed guidelines for the textual description of specific types of images, especially for figures displaying data, such as bar charts, pie charts, line charts and scatter plots. For the task of choosing and creating text alternatives, they generally propose an informative process that takes into account several factors, including the purpose of the image, if it provides essential or complementary information, if there is text embedded within the image and the context of the image.

However, the opinion of the author of the thesis is that these guidelines present the following shortcomings:

1. Type of image. Although general solutions have been proposed for the description of images in Science, Technology, Engineering and Mathematics, detailed guidelines on how to textually describe complex images with high information density and technical information, as those commonly used for example in medicine and biology fields, currently do not exist. For example there is no guideline on how to describe an X-Ray image.
2. Context and audience. They are mainly addressed to generic audience and intended for common publications. When recommendations are targeted to the creation of alternative description to STEM images in digital textbooks, such as in the case of the guidelines of NCAM, they are intended for students and scientists in the STEM fields who are blind or have low vision and not for audiences with other disabilities.
3. They are not related to the current practices and purposes of academic authors in the process of creating and selecting images for image submission to academic journals and they do not take into account the different functions of the images defined according to the author's beliefs and aims.
4. They are mostly based on sets of best practices that different institutions have developed in the accessibility field according to their experience, rather than on common theoretical principles on visual information creation and description. Excluding the guidelines proposed by Tang (2012), they do not take into account theoretical principles of visual information representation, such as those provided in the Library and Information Science research discipline.

Considering these limitations, in this thesis further investigation concerning the following issues should be performed:

1. Definition of theoretical models for the analysis and construction of visual information, with the aim of assessing the theoretical principles that regulate the creation of alternative representation for a better description of image content. The Information Visualization and Library and Information Science disciplines propose these models and they will be reviewed in the following section.
2. Review of the current policies and practices in image publishing in STEM academic journals.
3. Assessment of the functions and use of images in biomedical articles according to the behaviors, needs, attitudes and practices of academic researchers, paying particularly attention to their awareness and perception of accessibility.

These analyses have the objective of supporting a proposal for including accessibility criteria into the process of authoring images within academic articles with a minimal impact on the current publishing workflow.

4.9 Theoretical models for image analysis, description and representation in Library and Information Science.

In this chapter a bibliographic review of main principles and theoretical models on the analysis and description of images from the point of view of the Library and Information Science (LIS) discipline is presented. Firstly, the LIS discipline and the role of visual content in this area of study are presented. Secondly, theoretical principles of the representation of information in images and theoretical models used in the cataloguing and indexing of images are detailed. Finally, a summary of the contributions and limitations of the discipline regarding the identified issues is presented.

4.9.1 Library and Information Science (LIS)

4.9.1.1 Introduction to Library and Information Science

Library and Information Science (LIS) is the discipline "devoted to applying theory and technology to the creation, selection, organization, management, preservation, dissemination, and utilization of collections of information in all formats" (Reitz, 2004). An alternative definition is "the professional knowledge and skill with which recorder information is selected, acquired, organized, stored, maintained, retrieved and disseminated" (Enser, 2008).

In LIS a still image is considered as a "prototipo de documento visual, permanente y estable que puede ser explorado con todo grado de libertad. Se trata de un mensaje de código exclusivamente icónico y que por consiguiente puede ser comunicado por un canal visual"[1] (Pinto Molina 2002). As a document, an image is the subject of the processes of formal analysis, content analysis, retrieval and preservation.

[1] "Prototype of visual document, permanent and stable, that can be explored with any degree of freedom. It is an iconic message and therefore can be communicated by a visual channel".

Different treatments of the image can be distinguished in LIS:

- Traditionally, in Library Science an image has been considered as a document itself (a visual document), as in the case of a historical photo in a digital archive. The consequence of this approach was a limitation of studies on the relationships between the image and its context, due to the fact that the image was considered as an isolated entity from the context in which it was set. However, today this traditional approach has evolved to a more flexible approach, which takes into account the user, the domain knowledge and the context of the image. For example, the treatment of the document changes according to whether it belongs to an archive or to a specialized library.
- In some cases, an image is not considered a document, but a simple illustration. Consequently, the image is taken into account only for its practical use, for example an image used as thumbnail that visually support the choice of a reference (decision making process) in a catalogue.
- In other cases, especially with images in scientific domains, the image is treated as an appendix to other database information. For instance, a brain scan in a hospital archive is usually treated as an adjunct to a parent record.

The traditional theory of image in LIS was developed mainly from the scope of the field of research of humanities and the focus of the analysis was on cultural images (e.g. historical pictures) or art images, among others. These types of image have predominated in the literature of representation, indexing and retrieval of images and have been used to frame theoretical backgrounds for image representation studies. Still images from scientific domains, such as medical, architectural and engineering, have not traditionally been subject of investigation on their own, due to the fact that these types of images has been usually treated as adjuncts to parent record. Only recently scientific images have been recognized as an important information objects in their own right. This change has led to the creation of specialized collections for research and training purposes (Enser, 2008).

4.9.1.2 The process of information representation in images

The documentary process for the information representation in images systematizes and organizes the reading of image elements in successive phases, in order to uncover the different layers of their meaning and purpose. In LIS discipline, the analysis and interpretation of images (such as other documents) is divided into two operations:

- Formal analysis of data

  The process for creating entries for a catalogue is currently called bibliographic description. The bibliographic description deals with the identification and the description of the physical and bibliographic characteristics of the image: title and statement of responsibility (author, editor, composer, etc.), details of publication and distribution (place of creation, date of creation, etc.), physical description (format of image, type of image, etc.), series, notes and standard number and terms of availability (e.g. DOI, price). This formal element description determines the name(s) and title(s) to be used as access points in the catalogue, but not the assignment of subject headings. This formal analysis depends on the rules of cataloguing, according to the field of application (ISBD and Anglo-American Cataloguing Rules, Second Edition (AACR2) for libraries, ISAD G for archives, CDWA and VRA for museum, etc.).

- Content analysis

  The analysis of content is “El proceso cognitivo de reconocimiento, descripción y representación del contenido documental”[2](Pinto Molina, 2002). It determines the meaning of the image and the effect on its audience. Several models have been proposed for content analysis of documents (both visual and textual). For example, in the model of Valle Gastaminza (Valle Gastaminza 2001), the analysis of photos generally begins with the identification of the visual elements of the image (e.g. color, texture, spatial distribution and regions) and continues with the study of the meaning of the image in the context of social and other contextual data, such as philosophical, religious or aesthetic beliefs, references to politic and cultural codes.

After this analysis, a set of notions and concepts representative of the content of the image has been acquired. The process of *materialization* of these notions and concepts is called *representation*: this is the content description, the textual synthesis of the information transmitted through the image. The representation of the documents’ subject involves a set of operations that generates secondary information reflecting the most substantive elements of the documents’ content. These operations are indexing, classification and abstracting.

- Indexing is the process of extracting from the document a series of terms especially expressive and revealing of the subject, called descriptors. There are different types of descriptors, such as onomastic descriptors (individuals and legal entities), geographic descriptors and subject descriptors (abstract concepts, objects, and attitudes). In the process of image indexing, these descriptors are usually located in different fields.
  The process of indexing begins with an analysis of the subject of the document (for example, an image) performed by a specialist to determine the most specific subject heading(s) or descriptor(s) that fully describe the content of the document. The indexer selects those terms that identify the subject either by extracting words directly from the document (keywords) or assigning words from an indexing language (a controlled vocabulary, an established list of preferred terms). Afterwards, the terms in the index are presented in a systematic order allowing quick access to information. The descriptors serve as access points to the bibliographic record in a subject search of a library catalogue, index, abstracting service or bibliographic database.

- Classification is the process of dividing objects or concepts into logically hierarchical classes and subclasses based on the characteristics they have in common and those that distinguish them. It usually consists of assigning a code to the image from an existing

[2] “The cognitive process of recognition, description and representation of the document’s content”.

classification, a system with a methodical and systematic structure of classes related to each other according to a set of common characteristics. Examples of classification systems for libraries are the Dewey Decimal Classification (DDC) or the Library of Congress Classification (LCC).

- The content element description is summed up in an abstract. The abstract is the description of the image in text format and in natural language at different levels of completeness, according to the user's profile and needs. Abstract is very relevant in the documentary process of describing images, since the visual content of the image needs to be translated into words for improving retrieval. Despite this, currently there is no standard model for the creation of abstracts for images, unlike the abstracts that represent the content of texts, regulated by an ISO standard (NISO, 2015). A model for the image abstract has been proposed by Pinto Molina (2002), where the abstract of a photo is described by a set of features according to an established order (color, shot angle, type of lighting, element in first shot, objects represented, cultural framework of the photo, etc.).
  The definition of an ISO standard that regulates the creation of abstracts for different types of images could be also useful to improve digital STEM image accessibility. A source for this standard could be the ISO/IEC 20071-11:2012 standard (ISO/IEC, 2008), which provides guidance for the generation of alternative texts for images and which will be reviewed in the section 4.8.2.6.

4.9.1.3 Theoretical models for the image analysis

There are several theoretical models that propose a systematic approach to the analysis of the image content. They represent the formal background for the image description and indexing. In this study, four of the most representative models are presented, due to their relationship with the objective of the thesis and the number of citations received in the LIS academic field.

The "Iconographic model" (Panovsky)

The model of the art historian Panofsky (Panofsky, 1993), which has figured quite prominently in the literature, has the objective of formally analyzing Renaissance art images. It classifies the content of art images according to three different levels of description: a primary subject matter ("pre-iconography"), a second subject matter ("iconography") and a tertiary subject matter ("iconology").

- The primary subject matter is the description of the thematic or natural content of the image. It includes the identification of the visual primitives, like color, texture and shapes. It deals with the generic elements of the image and it does not require interpretative skills.
- The second subject matter is the conventional content in which specific themes and concepts (logical or derived features, objects, activities, and events) are placed. It requires the viewer to interpret the image.
- The tertiary subject matter embraces the intrinsic meaning of the image (inductive interpretation, abstract features) and it demands high-level semantic inference by the

viewer. At this level, the viewer is able to determinate the principles of a nation, historical period, a class, religious or philosophical beliefs.

One example of this approach applied to the picture of La Gioconda is shown in Figure 31:

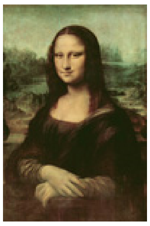


Image: Wikipedia Commons

Pre-iconographic	woman, landscape
Iconographic	Mona Lisa, Italy, three-quarter portrait
Iconological	beauty, nature, contentment

Figure 31 Layers of analysis of the picture *La Gioconda* Extracted from: (JISC Digital Media, 2014d)

Shatford model

Shatford (Shatford Layne, 1986) applies the Panofsky's model to image indexing, relabeling Panofsky's terms as *Generic* (pre-iconographic), *Specific* (iconographic) and *Abstract* (iconological). Each of these three levels has four facets: *Who*, *What*, *Where* and *When*. The matrix of 3x4 generated in this approach is referred to as the Panofsky/Shatford model and it has been used as a model for still and moving images analysis in several studies. Additionally, the model includes a distinction between the things an image is *Of* (objective things, either generic or specific) and the things an image is *About* (more subjective or abstract meanings).

Besides the attributes describing the "subject" of the image (e.g., protagonist, place, action and situation), Shatford considers other attributes conveying *externa*l or *non-visual* information an image might have:

- *Biographical*: the attributes resulting from the analysis of the history or life of the image: date and place of creation, title, localization, price, etc.
- *Exemplified*: the attributes resulting from the analysis of the images considered as a particular kind of object. For example, an image may be a photograph or a poster.
- *Relationship*: the attributes resulting from the analysis of how the image relates to other images or text. For instance,
    - Relationship of membership to the same collection: for instance, between photos from the same report, or from the same series or collection.
    - Intrinsic relationship: the direct relationship between two works. For instance, between a press photo and the correspondent news in text format. They have the same origin and are complementary.
    - Extrinsic relationship: the relation that is informative but not essential. For instance, the relations between the text and the illustration in a child book.

This approach postulates that the indexing of images should provide access to images based on the attributes of those images and should provide access to useful groupings of images, not simply to individual images.

Jaimes and Chang model (The "Syntactic and semantic model")

The work of Panofsky, Shatford and others set the theoretical background for the creation of the pyramid model for image description (Figure 32) proposed by Jaimes & Chang (1999). The objective of this conceptual framework is the indexing of different aspects of visual information.

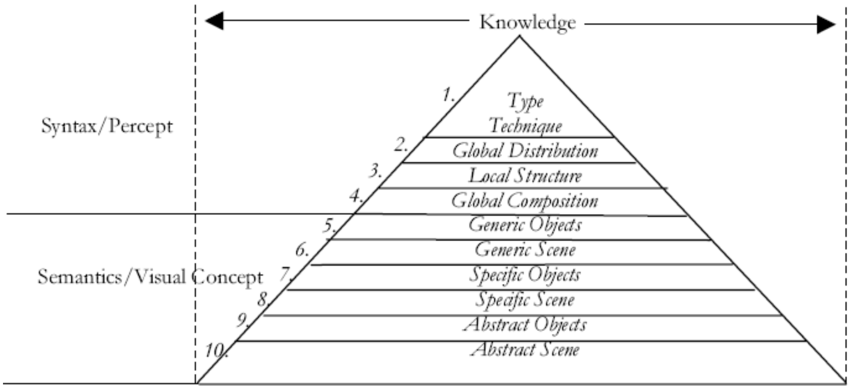


Figure 32 Pyramid model for images description Extracted from: (Jaimes & Chang, 1999)

The pyramid has ten layers classed in two big groups:

1. Syntactic/perceptual category (upper levels). Visual attributes associated with perception:
   a. Type and technique: general visual characteristics of the image. These attributes describe what sort of image it is (e.g. x-ray or diagram; B/W or color).
   b. Global distribution: global content of the image according to low-level perceptual features, i.e. the overall color or texture of the image (e.g. dominant color, contrast, density).
   c. Local structure: characterization of particular features within the image such dot, line, texture and other basic syntax symbols of graphics (e.g. dark spots in x-ray, lines in microscopic images).
   d. Global composition: the arrangement or layout of basic elements in the image (e.g. centered objects, diagonal leading line, and symmetry).

2. Semantic/conceptual category (lower levels). Visual attributes associated with the visual content and divided into descriptions as a whole (scene) and particular element descriptions (objects):
   a. Generic objects. The most general level of object description, which requires general knowledge to recognize these objects (e.g. a girl, a motorbike in a general context; a lumbar spine in a medical context).
   b. Specific objects: more concrete objects that can be identified and named (e.g. an ECG Monitoring Machine in a medical context).
   c. Generic scene: the description of a scene, environment, landscape (e.g. oblique/rotated view of a scanned brain).
   d. Specific scene: the description of a specific scene (e.g. para-sagittal MRI of a brain).

e. Abstract object: the specialized or interpretative knowledge about what the objects represent (e.g. beta-amyloid peptides in a medical context).
f. Abstract scene: what the image as a whole represents. This is the subjective interpretation of the scene (e.g. sadness, happiness, power, heaven and paradise in a general context; pathology judgments in a medical context).

In addition to the categories included in the pyramid, the authors acknowledge that there is some external or non-visual information related to an image, such as its history and context (this corresponds to Shatford's *Biographical* information). This information is the kind of information that can be usually found in metadata, such as date, location, right, source (Burford et al., 2003) (Hollink et al., 2004).

Another interesting aspect of this model is the definition of the *quantity of knowledge* required to describe the different layers. The upper layers require very little knowledge and can be safely left to a computer program for its analysis. The lower layers must be described by humans and require increasing amounts of knowledge (e.g. it takes more knowledge to recognize that the image is of the Mona Lisa rather than a woman). The lowest layers require the cataloguer to interpret the *meaning* of the image, which is obviously a very difficult and subjective task.

This description model has been used by Wang (Wang et al., 2011) to categorize the textual information provided by experts and students of medicine describing medical images.

Eakins/Graham Model

The model of Eakins & Graham (1999) proposes a distinction similar to the previous model, but focuses on queries rather than on indexes and its specific goal is to improve the retrieval of visual documents.

It identifies three levels of image queries, which reflect different levels of information's needs and correspond to different features of the image:

- Level 1: Queries based on primitive features, such as color or shape.
  The primitive (low level) queries are color, texture, shape or the spatial location of image elements. They are both objective and derivable from the images themselves and they do not need references to any external knowledge base for interpreting. Examples of such queries might include "find pictures with long thin dark objects in the top left-hand corner," "find images containing yellow stars arranged in a ring" - or most commonly "find me more pictures that look like this"(J. P. Eakins, 2001). These queries correspond to the four upper levels in the model of Jaimes and Chang.

- Level 2: Queries based on *logical* (sometimes known as *derived*) features, such as the objects depicted.
  Logical queries involve some degree of logical inference about the identity of the objects depicted in the image and normally require a reference to some external source of knowledge. Furthermore, they can be divided into queries of objects of a given type (e.g.

"find pictures of a double-decker bus") and queries of individual objects or persons ("find a picture of the Eiffel tower"). In the first example above, some prior understanding is necessary to identify an object as a bus rather than a lorry; in the second example, one needs the knowledge that a given individual structure has been given the name *the Eiffel tower*. These queries incorporate both general and specific descriptions in the model of Jaimes and Chang.

- Level 3: Queries based on abstract features, such as the interpretation of the scene depicted.
  Abstract queries need a significant amount of high-level reasoning about the meaning and purpose of the objects or scenes depicted and often a subjective judgment. (Eakins & Graham, 1999). Again, this level of retrieval can be subdivided into retrieval of named events or types of activity (e.g. "find pictures of Scottish folk dancing") and retrieval of pictures with emotional or religious significance ("find a picture depicting suffering"). These queries are equal to the abstract descriptions as described by Jaimes and Chang.

Synthesis of the models and opportunities for image description

The four models propose an image analysis divided in different levels or *layers* according to different aspects of the visual information. They agree in distinguishing the visual information into two main levels, based on the inherent features of the image:

- A syntactic or perceptual level, which includes primitive visual features, attributes associated with the perception of the image, objective and directly derivable from the images themselves.
- A semantic or conceptual level, which includes:
  - logical features: attributes associated with the content of the image requiring some degree of logical inference;
  - conceptual features: attributes associated with the abstract content of the image demanding high-level semantic reasoning.

Apart from these two levels, some authors (Hollink et al., 2004) have proposed descriptions of the external information of the image. This information "about" images, such as date, location, rights and source of the image, has traditionally defined as "metadata" (Hollink et al., 2004). The difference between this information and the one provided in the syntactic and semantic levels is that it cannot directly be derived from the content of the visual resource and it is objective, not affected by any interpretation.

The LIS models of image analysis reviewed in this section could offer a conceptual framework for analyzing the content of image-related texts (especially captions and mentions in the main text of the article) according to the "syntactic", "semantic" and "external" levels of information they convey. A lexical analysis of their content based on these categories could help in assessing the extent to which they could be potentially appropriate as alternative descriptions of the images.

4.9.2 Information Visualization (InfoVis)

Information Visualization (InfoVis), a relatively new discipline concerning "the depiction of information using spatial or graphical representations, to facilitate comparison, pattern recognition, change detection, and other cognitive skills by making use of the visual system" (Hearst, 2003). In other words, Infovis aims to develop graphical representations to help people in the comprehension and interpretation of data using special properties of visual perception.

It is a branch of Human-Computer Interaction field with origins in late 80s, but actually it involves more research fields, such as computer design, semiotic, cognitive psychology and visual perception.

The main contribution of InfoVis is the theoretical definition of the process of creating visualizations in relation to the process of knowledge construction (Card, Mackinlay, & Shneiderman, 1999)(Chi & Riedl, 2002). In this model, the image is intended as a part (or result) of a process of creation from data and information to a visual representation. The theoretical definition of the stages of the process could help to reverse it from the visual representation to the information behind it, to generate alternative representations.

Several models of image description have been proposed from the InfoVis discipline (Fathulla & Basden, 2007). Their common theoretical framework offers a syntax of the visual representation - visual variables and rules of construction of the components (visualization elements, patterns and structures) - and classifies images according to taxonomies based on this syntax (Engelhardt, 2002). Many authors (Blackwell & Engelhardt, 2002) (Takagi & Ishihara, 2007) (Duke, 2004) have noted that this framework could be a reference in the study of application of alternative representations to images, according to the principles of accessibility.

Furthermore, InfoVis takes into account that the designer and the user are the main agents involved in the process of knowledge creation by visualization. Some lines of research investigate the mental models and barriers on creating visual representation by users who are not familiar with information visualization and visual data analysis (Grammel, Tory, & Storey, 2010). This approach could be a reference on how users with disabilities explore alternative representations of graphics.

However, beyond the possible contributions of the discipline to the issue of how to textually describe images, InfoVis deals with highly-structured graphics representing data. As described in section 4.8, the content of text alternatives for this type of images has been defined by current accessibility guidelines and templates for creating text alternative of them are currently available. Finally, the application of the description models offered by InfoVis is beyond the scope of this thesis.

4.10 Conclusions of the literature review

This chapter has presented a review of the principles, methods and techniques provided by the accessibility discipline for creating accessible images according to the needs of specific categories of users with disabilities. In particular, text alternatives have been identified as the main solution for making images accessible in STEM academic articles. Advantages and limitations of the main mechanisms for applying them have been assessed. In order to overcome these limitations, other solutions that might be integrated into the current image publishing workflow without requiring severe changes in the current practices of authors and publishers have been explored. Image-related texts, such as captions, mentions, text inside images and metadata, have been recognized as potential mechanisms for providing text alternatives in STEM articles, due to their importance and use in the academic publishing. Other solutions offered by the Semantic Web technologies have been also presented and their possible integration in the publishing workflow has been evaluated.

Practical accessibility guidelines on how to textually describe visual information have been also reviewed and their advantages and limitations have been assessed. The main limitations identified by the author are:

- The lack of detailed guidelines on how to textually describe complex images as those commonly used for example in biomedical field, such as an X-Ray.
- To be mainly general and not focused on covering the needs of the audience of STEM academic publications.
- To be disconnected from the current practices and purposes of academic authors in the process of creating and selecting images for image submission
- The functions of the images defined according to academic author's beliefs and aims are not taken into account in the selection and creation of text alternatives.
- The lack of connection to theoretical principles on the textual description of visual information, such as those provided by the Library and Information Science research disciplines.

These observations have driven the author of the thesis to plan the following actions:

- To analyze the current policies and practices in image publishing in STEM journals, paying particular attention to the correct application of universally design principles.
- To study the function and content of image-related texts in STEM articles with the aim to explore their potential as mechanisms for conveying text alternatives.
- To assess the functions and use of images in biomedical articles according to the attitudes and practices of academic researchers, paying particularly attention to their awareness and perception of accessibility.

These analyses have the objective of supporting a proposal for including accessibility criteria into the process of authoring images within academic articles with a minimal impact on the current publishing workflow.

Chapter 5

Analysis of published articles and authors' practices and attitudes

5 Analysis of published articles and authors’ practices and attitudes

In this chapter, the results of the analysis performed are presented and commented. In the first section, the results of the audit of the current policies and practices in the submission of figures in STEM academic articles are detailed, with particular emphasis on how they revert on the accessibility of the figures. The second section presents the results of the analysis of authors’ practices and attitudes, with the final aim of assessing solutions for the inclusion of accessible recommendations integrated in the current author’s workflow.

5.1 Audit of the current policies and practices in STEM article publishing

This section describes the results of the analysis of the current policies and implementation of visual content in highly cited articles in the STEM research fields of biomedicine, computer science and mathematics. In particular, the results of the following analysis are detailed:

- Analysis of the accessibility of images published in the articles in biomedicine, computer science and mathematics.
- Analysis of the use of figures in articles in biomedicine, computer science and mathematics.
- Audit of the submission policies for authors provided by the publishers of the journals where the STEM articles were published.
- Analysis of the function and use of image-related texts (captions, mentions, text in the figures and metadata) currently used in biomedical academic articles.

Each analysis answers one of the following research questions:

- Are figures in STEM academic articles and in particular in biomedical articles currently accessible? (Q3).
- Are images in STEM academic articles and in particular in biomedical articles a relevant source of information? (Q4).
- Are there effective instructions for authors to make their images accessible in STEM and in particular in biomedical academic articles? (Q5).
- Are current image-related texts appropriate as alternative descriptions of the images in STEM and in particular in biomedical academic articles? (Q6)

Table XII shows the journals and articles selected for the analysis, the impact factor, the JCR subject category of the journals and when the Journal Citation Reports was consulted.

Table XII Articles journals selected for the analysis. Impact factor and JCR Subject Category of the journals are detailed.

Article ID	Selected Articles	Selected Journals	Impact Factor	JCR Subject Category	Accessed
MED1	Siegel, R., Naishadham, D., & Jemal, A. (2012). Cancer statistics, 2012. CA: A Cancer Journal for Clinicians, 62 (1), 10–29.	CA: A Cancer Journal for Clinicians	101.780	Oncology	26 September, 2012
MED2	Mega, J. L., Braunwald, E., Wiviott, S. D., Bassand, J.-P., Bhatt, D. L., Bode, C., … Gibson, C. M. (2012). Rivaroxaban in Patients with a Recent Acute Coronary Syndrome. New England Journal of Medicine, 366 (1), 9-19.	New England Journal of Medicine	53,298	Medicine, General and Internal	26 September, 2012
MED3	Steinman, R. M. (2012). Decisions About Dendritic Cells: Past, Present, and Future. Annual Review of Immunology, 30 (1), 1-22.	Annual Review of Immunology	52,761	Immunology	26 September, 2012
MED4	Codogno, P., Mehrpour, M., & Proikas-Cezanne, T. (2012). Canonical and non-canonical autophagy: variations on a common theme of self-eating? Nature Reviews Molecular Cell Biology, 13 (1), 7-12.	Nature Reviews Molecular Cell Biology	39,123	Cell Biology	26 September, 2012
MED5	Enden, T., Haig, Y., Kløw, N.-E., Slagsvold, C.-E., Sandvik, L., Ghanima, W., … Sandset, P. M. (7). Long-term outcome after additional catheter-directed thrombolysis versus standard treatment for acute iliofemoral deep vein thrombosis (the CaVenT study): a randomised controlled trial. The Lancet, 379 (9810), 31-38.	Lancet	38,278	Medicine, General and Internal	26 September, 2012

Article ID	Selected Articles	Selected Journals	Impact Factor	JCR Subject Category	Accessed
MED6	Branco, M. R., Ficz, G., & Reik, W. (2012). Uncovering the role of 5-hydroxymethylcytosine in the epigenome. Nature Reviews Genetics, 13 (1), 7-13.	Nature Reviews Genetics	38,075	Genetics and Heredity	26 September, 2012
MED7	Keith, B., Johnson, R. S., & Simon, M. C. (2012). HIF1α and HIF2α: sibling rivalry in hypoxic tumour growth and progression. Nature Reviews Cancer, 12 (1), 9-22.	Nature Reviews Cancer	37,545	Oncology	26 September, 2012
MED8	Prindle, A., Samayoa, P., Razinkov, I., Danino, T., Tsimring, L. S., & Hasty, J. (2012). A sensing array of radically coupled genetic `biopixels’. Nature, 481 (7379), 39-44.	Nature	36,28	Multidisciplinary science	26 September, 2012
MED9	Challen, G. A., Sun, D., Jeong, M., Luo, M., Jelinek, J., Berg, J. S., … Goodell, M. A. (2012). Dnmt3a is essential for hematopoietic stem cell differentiation. Nature Genetics, 44 (1), 23-31.	Nature Genetics	35,532	Genetics and Heredity	26 September, 2012
MED10	Renz, H., Brandtzaeg, P., & Hornef, M. (2012). The impact of perinatal immune development on mucosal homeostasis and chronic inflammation. Nature Reviews Immunology, 12 (1), 9-23.	Nature Reviews Immunology	33,287	Immunology	26 September, 2012
MED11	Rausch, T., Jones, D.T.W., Zapatka, M., Stütz, A. M., Zichner, T., Weischenfeldt, J., Jäger, N., …. Korbel, J.O. (2012) Genome Sequencing of Pediatric Medulloblastoma Links Catastrophic DNA Rearrangements with TP53 Mutations. Cell 148 (1–2), 59-71.	Cell	32403	Biochemistry, Cell Biology	26 September, 2012

Article ID	Selected Articles	Selected Journals	Impact Factor	JCR Subject Category	Accessed
MED12	Aldridge, B. B., Fernandez-Suarez, M., Heller, D., Ambravaneswaran, V. , Irimia, D., Toner, M., & Fortune., S. M. (2012) Asymmetry and Aging of Mycobacterial Cells Lead to Variable Growth and Antibiotic Susceptibility. Science, 335 (6064), 100-104.	Science	31.201	Multidisciplinary Science	26 September, 2012
CS1	Zargar, S.T.; Joshi, J.; Tipper, D. (2013). A Survey of Defense Mechanisms Against Distributed Denial of Service (DDoS) Flooding Attacks. IEEE Communications Surveys and Tutorials, 15 (4), 2046-2069.	IEEE Communications Surveys and Tutorials	6.311	Computer Science, Information Systems	23 April, 2013
CS2	Audhkhasi, K., & Narayanan, S. (2013). A Globally-Variant Locally-Constant Model for Fusion of Labels from Multiple Diverse Experts without Using Reference Labels. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35 (4), 769-783.	IEEE Transactions on Pattern Analysis and Machine Intelligence	4.908	Multidisciplinary science	23 April, 2013
CS3	Anighoro, A., & Rastelli, G. (2013). Enrichment Factor Analyses on G-Protein Coupled Receptors with Known Crystal Structure. Journal of Chemical Information and Modeling, 53 (4),739-743.	Journal of Chemical Information and Modeling	4.675	Computer Science, Information Systems	23 April, 2013
CS4	DeVore, R., Petrova, G., Hielsberg, M., Owens, L., Clack, B., & Sood, A. (2013). Processing Terrain Point Cloud Data. SIAM Journal on Imaging Sciences, 6 (1), 1-31.	SIAM Journal on Imaging Sciences	4.656	Computer Science, Artificial Intelligence	23 April, 2013

Article ID	Selected Articles	Selected Journals	Impact Factor	JCR Subject Category	Accessed
CS5	Faro, S. & Lecroq, T. (2013). The exact online string matching problem: A review of the most recent results. ACM Computing Surveys, 45 (2), Article 13.	ACM Computing Surveys	4.529	Computer Science, Theory & Methods	23 April, 2013
CS6	Venkatesh, V., Brown, S.A., & Bala, H. (2013). Bridging the Qualitative-Quantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems, 37 (1), 21-54.	MIS Quarterly	4.447	Computer Science, Information Systems	23 April, 2013
CS7	Toews, M., & Wells, W.M.III. (2013) Efficient and robust model-to-image alignment using 3D scale-invariant features. Medical Image Analysis, 17 (3), 271-282.	Medical Image Analysis	4.424	Computer Science, Artificial Intelligence	23 April, 2013
CS8	Wu, G.D., & Huang, P.H. (2013). A Vectorization-Optimization-Method-Based Type-2 Fuzzy Neural Network for Noisy Data Classification. IEEE Transactions on Fuzzy Systems, 21 (1), 1-15.	IEEE Transactions on Fuzzy Systems	4.260	Computer Science, Artificial Intelligence	23 April, 2013
CS9	Xie, Y. (2013). animation: An R Package for Creating Animations and Demonstrating Statistical Methods. Journal of Statistical Software, 53 (1).	Journal of Statistical Software	4.010	Computer Science, Interdisciplinary Applications	23 April, 2013
CS10	Chen, Z.,Wong, KY. K., Matsushita, Y., & Zhu, X. (2013) Depth from Refraction Using a Transparent Medium with Unknown Pose and Refractive Index. International Journal of Computer Vision, 102 (1-3), 3-17.	International Journal of Computer Vision	3.741	Computer Science, Artificial Intelligence	23 April, 2013

Article ID	Selected Articles	Selected Journals	Impact Factor	JCR Subject Category	Accessed
MATH1	Nazarov, F., Ryabogin, D. & Zvavitch, A. (2014). An asymmetric convex body with maximal sections of constant volume. Journal of The American Mathematical Society, 27, 43-68.	Journal of The American Mathematical Society	3.567	Mathematics	01 October, 2013
MATH2	Bogomolov, F., & Tschinkel, Y. (2013). Galois Theory and Projective Geometry. Communications on Pure and Applied Mathematics, 66 (9), 1335–1359.	Communications on Pure and Applied Mathematics	3.337	Mathematics, Applied	01 October, 2013
MATH3	Avila, A., & Gouëzel, S. (2013) Small eigenvalues of the Laplacian for algebraic measures in moduli space, and mixing properties of the Teichmüller flow. Annals of Mathematics, 178 (2), 385-442.	Annals of Mathematics	3.027	Mathematics	01 October, 2013
MATH4	Miermont, G. (2013). The Brownian map is the scaling limit of uniform random plane quadrangulations. Acta Mathematica, 10 (2), 319-401.	Acta Mathematica	2.714	Mathematics	01 October, 2013
MATH5	Bestvina, M. (2014). Geometric group theory and 3 manifolds hand in hand: the fulfillment of Thurston's vision. Bulletin of The American Mathematical Society, 51, 53-70.	Bulletin of The American Mathematical Society	2.556	Mathematics	01 October, 2013
MATH6	Chiodo, A., Eisenbud, D., Farkas, G., & Schreyer, F.O. (2013). Syzygies of torsion bundles and the geometry of the level ℓ modular variety over Mg. Inventiones Mathematicae, 194 (1).	Inventiones Mathematicae	2.259	Mathematics	01 October, 2013

Article ID	Selected Articles	Selected Journals	Impact Factor	JCR Subject Category	Accessed
MATH7	Hauenstein, J.D. & Wampler, C. W. (2013). Isosingular Sets and Deflation. Foundations of Computational Mathematics, 13 (3), 371-403.	Foundations of Computational Mathematics	1.918	Mathematics	01 October, 2013
MATH8	Figalli, A., Kim, Y.H., & McCann, R.J. (2013). Regularity of optimal transport maps on multiple products of spheres. Journal of The European Mathematical Society, 15 (4), 1131–1166.	Journal of The European Mathematical Society	1.880	Mathematics, Applied	01 October, 2013
MATH9	Kato, K., & Saito, T. (2013) Ramification Theory for Varieties over a Local Field. Publications Mathématiques de L'IHÉS, 117 (1), 1-178.	Publications Mathématiques de L'IHÉS	1.867	Mathematics	01 October, 2013
MATH10	Mackowiak, P. (2013). The existence of equilibrium in a simple exchange model. Fixed Point Theory and Applications, 104.	Fixed Point Theory and Applications	1.886	Mathematics, Applied	01 October, 2013

5.1.1 Analysis of the journals

This section presents the results of the analysis of the journals in biomedicine, computer science and mathematics. In particular, the results of the following issues are presented and discussed:

- Accessibility policy, as stated in the journal policies on the publisher's website (where available).
- Submission policy addressed to authors, retrieved from the journal instructions for authors on the publisher's website, including specific recommendations related to the image accessibility (where available) and concerning image-related texts in the article (where available).
- Options for image retrieval offered by the retrieval system on the journal's website and options of image presentation offered on the journal's website (only for biomedical articles).

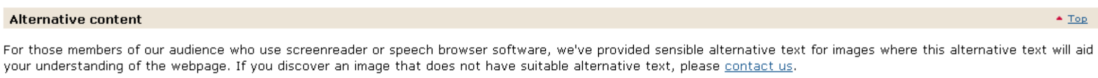
The main aim of this analysis is to answer to the research question Q5: Are there effective instructions for authors to make their images accessible in STEM and in particular in biomedical academic articles?.

5.1.1.1 Accessibility policy

Accessibility policy results

Among the 12 biomedical journals analyzed, eight had an accessibility policy statement on their web site. In all of them, accessibility statements were presented as a publisher statement and they were not mentioned in the submission guidelines for authors. In the accessibility statement of *CA: A Cancer Journal For Clinicians* and *The Lancet*, a reference to compliance with official accessibility legislation in the USA (508 Guidelines) and references to WCAG 2.0 were found.

In relation to the use of alternative content for images, the accessibility statement of six journals (all of them published by the Nature Publishing Group) provided an explicit reference to it (Figure 33).


Alternative content ▲ Top

For those members of our audience who use screenreader or speech browser software, we've provided sensible alternative text for images where this alternative text will aid your understanding of the webpage. If you discover an image that does not have suitable alternative text, please contact us.

Figure 33 Example of reference to the application of alternative content in Nature's Accessibility Policy. The paragraph says: "For those members of our audience who use screen reader or speech browser software, we've provided sensible alternative text for images where this alternative text will aid your understanding of the webpage. If you discover an image that does not have suitable alternative text, please contact us". Extracted from: (Codogno, Mehrpour, & Proikas-Cezanne, 2012).

We also found an invitation to request help to access images without suitable alternative text in the alt text tag of the images (Figure 34).

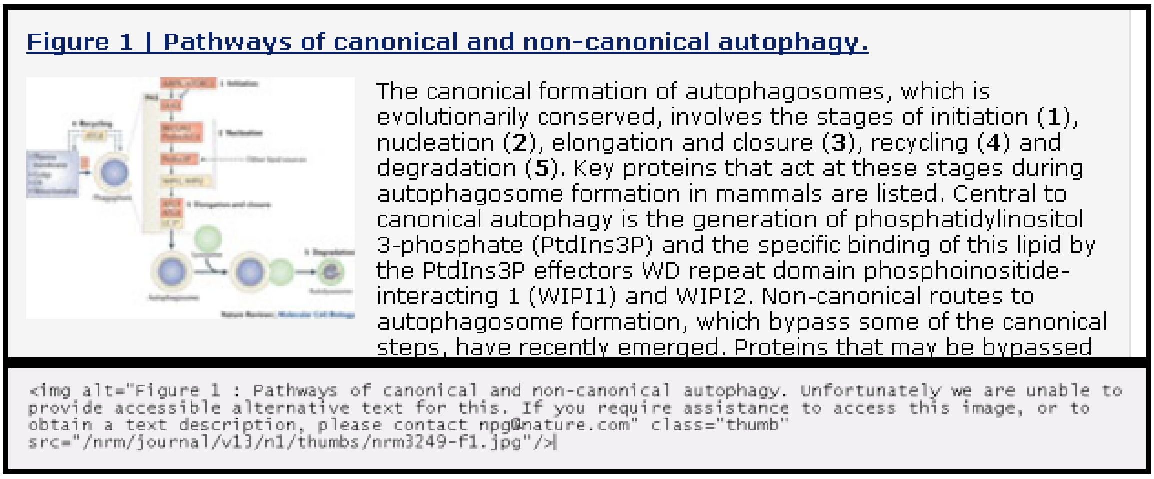


Figure 34 Example of alt text including an invitation to request an accessible description for an image in a paper published by a Nature's journal. The alt text says: "Unfortunately we are unable to provide accessible alternative text for this. If you require assistance to access this image, or to obtain a text description, please contact npg@nature.com". ". Extracted from: (Codogno et al., 2012).

Five journals out of ten analyzed in computer science had an accessibility policy statement either on their website or on the site of their publishers, four of them referring to WCAG 2.0 and one of them to WCAG 1.0. Only the Accessibility Policy of the *International Journal of Computer Vision* made a specific reference to the application of textual alternatives "in order to give people of all abilities and disabilities access to the content of your figures".

No journal in Mathematics had an accessibility policy statement.

Conclusions about accessibility policy

In synthesis, the results showed that only 13 journals out of 32 had an accessibility policy statement in their web sites, while only nine out of 32 provided a reference to the use of the alternative to users with special needs in their submission guidelines (Table XIII). Regarding the accessibility of images, only nine journals made explicit reference to the use of alternative descriptions for images on their website.

Table XIII Results of the analysis of the presence of an accessibility policy statement in the web sites of the journals and the reference to the use of text alternatives

	Accessibility Policy	
Journal	**web accessibility**	**reference to text alternatives**
MED1	x	
MED2		
MED3		
MED4	x	x
MED5	X	

	Accessibility Policy	
Journal	web accessibility	reference to text alternatives
MED6	X	x
MED7	X	x
MED8	X	x
MED9	X	x
MED10	x	x
MED11		
MED12		
CS1	x	
CS2	x	
CS3		
CS4		
CS5	x	
CS6		
CS7	x	
CS8	x	
CS9		
CS10		x
MATH1		
MATH2		
MATH3		
MATH4		
MATH5		
MATH6		x
MATH7		x
MATH8		
MATH9		
MATH10		

The subscription of an accessibility policy statement implies a commitment to make the content of the journal accessible to people with disabilities. As commented in section 1.2, several legal mandates over the world require public and private organizations to provide accessible services and products, including the United States Section 508 Amendment to the Rehabilitation Act and the Real Decreto Legislativo 1/2013 in Spain. Even publishers of journals that do not make such a statement should bear in mind that they are used in government-funded organizations such as hospitals, universities and libraries, which are subject to national accessibility laws.

The main difference among the journals in the three disciplines analyzed was that while seven journal's web sites in biomedicine and five in computer science presented an accessibility statement, no web site of the mathematical journals did so. The analysis suggested that publishers in the research field of mathematics seem to have less awareness of the needs for providing universally accessible contents, in particular images, compared to the publishers in the fields of biomedicine and computer science.

5.1.1.2 Submission policy

General observations on submission policy

All the biomedical journal websites offered a section specifying their submission policy for papers, including a sub-section explaining how to submit images in papers. All the biomedical journals provided specific guidelines for the preparation of figures for articles. These guidelines were provided by specific documentation that included different kind of information, such as detailed image specifications required for submission, examples of good and bad graphics, FAQs about the image submission process and a glossary of basic concepts about images. The accessibility issues related to images (alternative descriptions, color and contrast, font type and size of text within the labels in images, file format, dimensions and resolution of the image) were hardly mentioned explicitly in the image submission guidelines for authors, with the exception of occasional references to specific reader barriers and the legibility of the image in three journals of biomedicine, in one journal of computer science and in two journals of mathematics. *CA: Cancer Journal For The Clinicians*, The *New England Journal of Medicine* and *The Lancet* were listed as following the "Uniform Requirements for Manuscripts Submitted to Biomedical Journals" (URM) issued by the International Committee of Medical Journal Editors (ICMJE) (International Committee of Medical Journal Editors, 2014). These requirements include specific points for figure submission, such as image quality and size, use of symbols and colors in specific illustrations, and general guidelines concerning the captions of figures.

Concerning the journals in computer science, all the journal websites offered a section specifying their submission policy for papers, including a sub-section explaining how to submit artwork in papers.

All the journals provided specific guidelines for the preparation of figures for articles, as explained before for images in biomedicine. No journal declared to follow any guideline issued by an international committee in the preparation and formatting of manuscripts. Accessibility issues related to images were not mentioned explicitly in the image submission guidelines for authors. The only exception was a reference provided by the *International Journal of Computer Vision* on the use of patterns in addition to colors for conveying information, in order to make color-blind users able to distinguish visual elements. The same journal recommended providing a minimum color contrast ratio between letters within images and their background.

All the websites of mathematical journals offered a section specifying their submission policy for papers, including a sub-section explaining how to submit artwork in papers. Five journals offered a detailed documentation about image specifications for submission. Accessibility issues related to images were not mentioned explicitly in the image submission guidelines for authors. The only exceptions were the references provided by the *Inventiones Mathematicae* and F*oundations of Computational Mathematics* on the use of patterns in addition to colors for conveying information and the application of a minimum color contrast ratio between letters and their background. Only one journal (*Fixed Point Theory and Applications*) declared that it followed the standard guidelines issued by an international committee, the *Committee on Publications ethics* (Committee on Publication Ethics, n. d.). However, these guidelines did not make specific suggestions on the preparation and formatting of figures.

Table XIV shows a summary of the results of the analysis of the submission policies.

Table XIV Results of the analysis of the submission policies of the journals.

Journal	submission policies and guidelines for authors (including guidelines for artwork)	specific documentation for artwork submission	follow international guidelines for submission	follow international guidelines covering some accessibility-related issues	accessibility issues explicitly mentioned in the image submission guidelines
MED1	x	x	x	x	
MED2	x	x	x	x	
MED3	x	x			
MED4	x	x			
MED5	x	x	x	x	
MED6	x	x			x
MED7	x	x			x

Journal	submission policies and guidelines for authors (including guidelines for artwork)	specific documentation for artwork submission	follow international guidelines for submission	follow international guidelines covering some accessibility-related issues	accessibility issues explicitly mentioned in the image submission guidelines
MED8	x	x			
MED9	x	x			x
MED10	x	x			
MED11	x	x			
MED12	x	x			
CS1	x	x			
CS2	x	x			
CS3	x	x			
CS4	x	x			
CS5	x	x			
CS6	x	x			
CS7	x	x			
CS8	x	x			
CS9	x	x			
CS10	x	x			x
MATH1	x	x			
MATH2	x	x			
MATH3	x				
MATH4	x				
MATH5	x	x			
MATH6	x	x			x
MATH7	x	x			x
MATH8	x				
MATH9	x				

Journal	submission policies and guidelines for authors (including guidelines for artwork)	specific documentation for artwork submission	follow international guidelines for submission	follow international guidelines covering some accessibility-related issues	accessibility issues explicitly mentioned in the image submission guidelines
MATH10	x		x		

All journals offered guidelines to authors on how to prepare and submit a manuscript. 27 journals out of 32 provided specific and detailed documentation for artwork submission. The five journals that do not offered this documentation were from the mathematics discipline. This result suggested that figures currently have a relevant role in academic articles, especially in journals of biomedicine and computer science.

Only four journals out of 32 explicitly declared to follow international guidelines for the submission of academic articles. ICMJE guidelines seemed to be a relevant international guideline for preparation and submission of articles in biomedical journals (as it was cited by three journals), while one journal in mathematics presented a reference to this kind of guidelines. Journals in computer science did not present any reference to international guidelines for submission.

Only six journals out of 32 explicitly mentioned image accessibility issues and their impact on readers with special needs. This result, together with the lack of a web accessibility statement in the journals website, suggested a possible lack of awareness of the image accessibility requirements among publishers.

Technical requirements in image submission

In image submission we analyzed technical requirements such as format, dimensions, color and resolution that can affect their accessibility, comparing them to specific treatment of figures targeting users with special needs and highlighting the possible consequences of stated requirements regarding accessibility/readability for readers with special needs.

Format of the image

All the biomedical journals accepted images in *PostScript* (PS) format, *Encapsulated PostScript* (EPS) format and in *Tagged Image File format* (TIFF). It seemed that the high acceptance of PostScript was due to the fact that it supports vector graphics (graphs, chart, etc.), allows text resize without losing quality and preserves the editability of the figure from the original version. Only *The Lancet* accepted SVG as a submission format for figures. Many articles accepted also JPEG and Photoshop format (PSD). TIFF, JPEG and PSD were generally preferred for photographic images and illustrations (raster graphics). Six journals accepted PDF, eight accepted PPT and seven accepted the Adobe Illustrator format (AI). Other accepted formats were GIF, BMP, XLS, CRD and DOC(X), although these formats were restricted to specific types of images, and they were not considered as preferred formats for final submission.

Concerning the journals in computer science, all the journal submission guidelines specified recommended formats for image submission. Five journals only accepted images embedded in MS Word files and/or PDF converted from LaTex format. For raster images, five journals recommended TIFF format, four JPEG and two PNG. For vector images, the most accepted formats were EPS (six journals). One journal accepted the native Chemdraw file format (CDX). The *Journal of Chemical Information and Modeling* also accepted AI format. Other accepted formats in two journals were XLS, PPT and DOC, but only under specific conditions (e.g. "Only when the graphic was originally drawn in the program, is it acceptable to submit Microsoft Office Documents").

All the mathematical journal submission guidelines specified recommended formats for image submission. Two journals only accepted images embedded in PDF files and two journals only in PDF/EPS files. For raster images, six journals recommended TIFF format, three JPEG format and two GIF format. For vector images, the only accepted format was EPS (eight journals). One journal accepted CDX format.

Other accepted formats were PPT (three journals), DOC(X) (three journals) and XLS (two journals). Two journals accepted PICT.

Table XV shows a summary of the formats recommended by the submission policies.

Table XV Digital formats accepted for images by the journals, categorized according to the format type: raster images, vector images and other document formats that can embed images. PSD format (1) is primarily a raster image format, but it also can include vector information. EPS (2) is a metafile commonly used to encapsulate vector graphics. SVG (3) is an XML markup language commonly used for describing vector graphics. PDF (4) is a metafile commonly used for encapsulating both raster and vector graphics.

Journal	Digital Image formats accepted																	
	Raster Image						Vector Image							Document formats (encapsulating both vector and raster images)				
	TIFF	JPEG	PSD(1)	GIF	BMP	PNG	EPS(2)	PS	AI	CRD	PICT	SVG(3)	CDX	PDF(4)	PPT	XSL	DOC(X)	LaTeX
MED1	x						x											
MED2	x	x	x	x	x		x		x	x				x	x	x	x	
MED3	x		x		x		x		x	x		x		x	x	x		
MED4	x	x	x				x		x	x					x	x		
MED5	x	x	x				x	x	x			x		x	x			
MED6	x	x	x				x		x	x					x	x		
MED7	x	x	x				x		x	x				x	x	x		
MED8	x	x		x			x	x						x	x			
MED9	x						x	x	x									
MED10	x	x	x				x		x	x					x	x		
MED11	x	x					x							x				
MED12	x						x	x						x			x	
CS1	x	x				x	x	x		x				x	x	x		
CS2	x	x		x	x	x	x							x				
CS3	x	x	x				x		x				x	x				

Journal	Digital Image formats accepted																	
	Raster Image						Vector Image							Document formats (encapsulating both vector and raster images)				
	TIFF	JPEG	PSD(1)	GIF	BMP	PNG	EPS(2)	PS	AI	CRD	PICT	SVG(3)	CDX	PDF(4)	PPT	XSL	DOC(X)	LaTeX
CS4														x				
CS5																	x	x
CS6																	x	
CS7	x	x					x							x	x	x	x	
CS8	x						x	x						x	x	x	x	
CS9																		x
CS10							x								x	x	x	x
MATH1	x	x		x			x				x							
MATH2	x						x							x				
MATH3							x							x				
MATH4														x				
MATH5	x	x		x			x				x							
MATH6	x						x								x	x	x	
MATH7	x						x								x	x	x	
MATH8							x							x				
MATH9														x				
MATH10	x	x			x	x	x						x	x	x		x	
Total	23	15	8	5	4	3	26	6	9	7	2	2	2	19	15	12	10	3

The most accepted formats were TIFF (23 journals) and JPEG (15 journals) for raster images and EPS (26 journals) for vector images. The high acceptance of TIFF and EPS as preferred figure file formats showed a great affinity with the findings of Jackson (Jackson et al., 2001) in Radiology Journals. This suggests that the image file formats accepted by health science journals have remained generally the same for over a decade. The inclusion of images in a PDF file is another common accepted option for the submission of articles in the three areas of research.

From the point of view of accessibility, vector images are preferred to raster images because they maintain their resolution despite zooming and resizing. Under accessibility concerns, the formats most suitable for image publication are those which do not lose quality due to compression and retain all information that is created by the capture device or the author. A better resolution of the image and the ability to explore the figure at increased magnification levels would benefit many readers and allow them to focus on small details of the image, as in the case of detailed *anatomic* images.

The results showed that the most accepted formats for submitting vector images were the EPS and PDF metafile. They are usually recommended for line art and combinations of photos and labeling and they also support the inclusion of information in the form of metadata. PDF supports mechanisms for the inclusion of text alternatives, as described in section 4.3.4. The SVG format, accepted only by two journals, is an interesting alternative format for encapsulating vector images very suitable for accessibility concerns. It allows the inclusion of text descriptions for each logical component of the image and for the image as a whole, which may be rendered on screen as speech or as Braille using assistive technologies.

Among raster formats, TIFF was the most accepted. It is a lossless compression format and it does not lose quality due to compression. Therefore, it is suitable for accessibility concerns. JPEG was the second most accepted format of raster graphic. Although the most commonly used version of JPEG allows a smaller image file size than TIFF, it applies a compression process that causes the loss of some visual quality which cannot be restored. Therefore, it is not considered the best solution for conveying visual information to people with visual disabilities.

The inclusion of images in document files accepted by publishers as PPT, XLS and DOC, is not recommended, because it can causes a loss of quality when images with high resolution are exported to other formats. For example, although PowerPoint offers the possibility of exporting raster figures to other image formats in a specified resolution, the maximum resolution allowed is 307 dpi (Microsoft, 2015).

Image dimensions

Small dimensions could negatively affect the ability of low-vision readers, and all readers in general, to properly understand the content of the image, especially in the printed version of the paper. Authors are therefore asked to submit images in a size accessible to all readers.

In biomedicine journals, the majority of the guidelines required a minimum final size for image submission, starting from a minimum final width of approximately 3.5” for half-page figures and 8.5” for full-page figures. The stated image height ranged from 7.9” to 11”.

The majority of the guidelines provided by journals in computer science required a minimum size for image submission, starting from a minimum final width of approximately 1.18/3.39” to a maximum of 6.9/7.48”. The stated maximum image height (four journals) ranged from 8.8” to 9.21”.

Concerning the mathematical journals, the guidelines of three journals required a minimum final width, ranging from 1.53” to 3.34”. Four journals stated a maximum final width, ranging from 4.92” to 6.85”. The stated maximum image height (three journals) ranged from 8.85” to 9.21”.

Table XVI shows a summary of the results of the analysis of the image dimensions recommended in the submission policies. For biomedical articles, the maximum length of the horizontal dimension (width) of an image included in the article and the columns of the articles were reported.

Table XVI Image dimensions recommended by the journals and the actual width of print (excluding the left and right indentation) manually measured on the articles. The actual width of print have been measured only for biomedical articles.

Journal	Recommended image dimensions			Total print dimensions	
	Minimum width	**Maximum width**	**Maximum height**	**Width (cm)**	**Columns**
MED1	6.75" (17.14 cm)			7.19” (18.25 cm)	1
MED2	4" (10.26 cm)	6.5" (16.51 cm)		6.60” (16.77 cm)	2
MED3		6.33" (16.07 cm)	7.9" (20.06 cm)	6.33” (16.06 cm)	2
MED4	3.93" (10 cm)			7.07” (17.95 cm)	2
MED5	3.93" (10.7 cm)			7.38” (18.75 cm)	2
MED6	3.93" (10 cm)			7.07” (17.95 cm)	3 (2)
MED7	3.93" (10 cm)			7.10” (10.03 cm)	2
MED8	3.50" (8.9 cm)	7.2" (18.3 cm)	9.72" (24.7 cm)	7.18” (18.24 cm)	2
MED9				7.06” (17.93 cm)	2
MED10	3.93" (10 cm)			7.06” (17.93 cm)	2
MED11		8.5" (21.59 cm)	11" (27.94 cm)	6.84” (17.37 cm)	2
MED12		8.5" (21.59 cm)	11" (27.94 cm)	7.27” (18.46 cm)	3 (2)
CS1	3.50" (8.89 cm)	7.16" (18.2 cm)	8.8" (22 cm)		
CS2	3.39" (8.61 cm)	6.9" (17.53 cm)	8" (20.32 cm)		
CS3	3.33" (8.45 cm)	7" (17.78 cm)	9.17" (23.29 cm)		

Journal	Recommended image dimensions			Total print dimensions	
	Minimum width	Maximum width	Maximum height	Width (cm)	Columns
CS4					
CS5					
CS6					
CS7	1.18" (3.3 cm)	7,48" (19 cm)			
CS8	3.20" (8.12 cm)	7.16" (18.18 cm)			
CS9					
CS10	1.53" (3.9 cm)	6.85" (17.4 cm)	9,21" (23.4 cm)		
MATH1					
MATH2					
MATH3					
MATH4					
MATH5					
MATH6	1.53" (3.9 cm)	6.85" (17.4 cm)	9,21" (23.4 cm)		
MATH7	1.53" (3.9 cm)	6.85" (17.4 cm)	9,21" (23.4 cm)		
MATH8		4.92 (12.5 cm)			
MATH9					
MATH10	3.34" (8.5 cm)	6.69" (17 cm)	8.85" (22.5 cm)		

The results suggested that there is no agreement among publishers about the recommended dimensions of the images.

The biomedical guidelines reviewed seemed generally to relate the suggested dimensions to the standard PDF format used in the publication of the manuscript, in order to fit the image in the width of one or two columns.

Color of the image

The proper use of color and contrast in images allows users with colorblindness to correctly perceive images and improves access to images for every user in poor lightning conditions and on mobile devices with a limited color palette and ones that use grey scales (such as some types of e-reader devices). From the accessibility perspective, it is also recommended not to rely on colors as the only means to differentiate information, and supplement the color variations with pattern changes. It is recommended that publishers include references and examples on how to use color for properly conveying information to readers with special needs to inform and help authors in their task.

The biomedical journals generally expressed the preference for figures to be submitted in color, especially if colors were useful to identify specific parts of the image. In relation to accessibility, the guidelines of *Nature Genetics* suggested providing optimal contrast with the use of color. In particular, their recommendation of using blue and yellow colors rather than red and green is based on the most common color blindness: deuteranopia. The guidelines of *Nature Genetics* also made a specific reference to readers with color blindness ("Authors are encouraged to consider the needs of color-blind readers") and suggested that primary data should be recolored using "color-safe combinations such as green and magenta, turquoise and red, yellow and blue or other accessible color palettes" (Nature Publishing Group, 2012).

Among the journals in computer science, only the *SIAM Journal on Imaging Sciences* explicitly encouraged the use of color. The *International Journal of Computer Vision* recommended the use of patterns in addition to colors for conveying information in order to allow color-blind users to distinguish visual elements and recommended a minimum color contrast ratio between letters and their background of 4.5:1. This recommendation corresponds to the guideline 1.4.3 of WCAG 2.0. *IEEE Transactions on Pattern Analysis and Machine Intelligence* suggested "avoiding bright, light, or neon colors for images that will be in print".

No mathematical journal encouraged the use of color. *Inventiones Mathematicae* and *Foundations of Computational Mathematics* recommended the use of patterns in addition to colors for conveying information in order to allow color-blind users to distinguish visual elements and recommended a minimum color contrast ratio between letters and their background of 4.5:1.

Table XVII sums up the results of the recommendations on the use of color.

Table XVII The table shows whether the journal specifies any preference on the submission of color images and recommendations on the use of color and contrast.

Journal	Color and contrast		
	Color image preferred	**Recommendations on the use of color**	**Accessibility related recommendation**
MED1	x	x	
MED2			
MED3	x	x	
MED4			
MED5	x	x	x
MED6			
MED7			
MED8	x	x	
MED9	x	x	x
MED10			

Journal	Color and contrast		
	Color image preferred	Recommendations on the use of color	Accessibility related recommendation
MED11		x	
MED12		x	
CS1			
CS2		x	
CS3		x	
CS4	x		
CS5			
CS6			
CS7			
CS8			
CS9			
CS10		x	x
MATH1		x	
MATH2			
MATH3			
MATH4			
MATH5		x	
MATH6		x	x
MATH7		x	x
MATH8			
MATH9			
MATH10			

The results show a preference of journals in biomedicine for submitting images in color, compared to the journals in the other disciplines. Specific recommendations on the use of color are not provided by 14 journals out of 32. However, only five journals made explicit references to the specific use of color for readers with visual limitations.

Resolution of the image

A low resolution could negatively affect access to the image by low-vision readers. Particularly at high levels of magnification, the readability of raster images can be compromised by “pixilation”. Vector formats could be a solution for avoiding the loss of resolution when images are displayed at the magnification rates of x3 or higher. In the publishing industry, a resolution of 300 dpi is commonly considered the minimum acceptable standard for ensuring quality for the print reproduction of figures (University of Chicago Press, 2015).

In most biomedical journals the minimum resolution required was 300 dots per inch (in *Nature* the minimum was 150 dpi) (Table XVIII). In 10 journals, the resolution requirements depended on the type of images; in the others, it depended on the submission step, as higher resolution was requested for revised figures submitted after peer review. The preferred resolution ranged from 266 to 600 dpi for color images, from 266 to 600 dpi for grayscale images and from 600 to 1200 dpi for line-art images.

Six journals in computer science required a minimum resolution for all types of image submission. In five journals, the resolution requirements depended on the type of image: for example, the preferred resolution for line-art images differed for color and grayscale images and ranged from 300 to 1200 dpi.

Six mathematical journals required a minimum resolution for all types of image submission. In five journals, the resolution requirements depended on the type of image: for example, the preferred resolution for line-art images differed for color and grayscale images and ranged from 300 to 1200 dpi.

Table XVIII shows a summary of the image resolution recommended by the guidelines analyzed.

Table XVIII Recommended resolutions for the submission of different image types: color, grayscale images and line art.

Journal	Recommended Image resolution		
	Color	**Grayscale**	**Line-art**
MED1	300 dpi	300 dpi	1200 dpi
MED2	266 dpi	266 dpi	1200 dpi
MED3	300 dpi	300 dpi	300 dpi
MED4	300 dpi	600 dpi	1200 dpi
MED5	300 dpi	300 dpi	300 dpi
MED6	300 dpi	600 dpi	1200 dpi
MED7	300 dpi	600 dpi	1200 dpi
MED8	300 dpi	600 dpi	1200 dpi
MED9	300 dpi	600 dpi	1200 dpi
MED10	300 dpi	600 dpi	1200 dpi
MED11	300 dpi	500 dpi	1000 dpi
MED12	400 dpi	400 dpi	600 - 1200 dpi
CS1	300 dpi	300 dpi	600 dpi
CS2	300 dpi	300 dpi	300 dpi

Journal	Recommended Image resolution		
	Color	**Grayscale**	**Line-art**
CS3	300 dpi	600 dpi	1100 dpi
CS4			
CS5			
CS6			
CS7	300 dpi	300 dpi	1000 dpi
CS8	400 dpi	300 dpi	600 dpi
CS9			
CS10	300 dpi	300 dpi	1200 dpi
MATH1	300 dpi	300 dpi - 600 dpi	600 - 1200 dpi
MATH2	300 dpi	300 dpi	600 dpi
MATH3			
MATH4			
MATH5	300 dpi	300 dpi - 600 dpi	600 - 1200 dpi
MATH6	300 dpi	300 dpi	1200 dpi
MATH7	300 dpi	300 dpi	1200 dpi
MATH8			
MATH9			
MATH10	300 dpi	300 dpi	300 dpi

The results show a similar trend among the journals in different disciplines: the preferred resolution ranged from 266 to 600 dpi for color images, from 300 to 600 dpi for grayscale images and from 300 to 1200 dpi for line-art images.

These differences could be related to the overall file size, as color images need 32 bits for each pixel or color definition, grayscale images only 8 bits, and line-art images only 1 bit.

Size and name of the image file

File size is directly proportional to image quality and the balance between the file size and quality depends on the intended use of the image. If the figure is for publication, limitations on file size could force the author to compress high-quality images, causing a loss of clarity.

The file name is the simplest textual data describing the image. Accurate file naming can provide valuable information to be read by screen readers and retrieved by search engines to identify a figure and establish its importance for retrieval.

The size of image files was specified in the guidelines of six biomedical journals. Small files of less than 2-3 MB were generally recommended for submission. *Cell* indicated 20 MB as the maximum size for the image file. Five guidelines required figure numbers, based on the order of appearance in the text, to be included in the name of the image file. *Nature* and *Science* recommended including the author surname in the file name, as for example “smithfig1”. *The Lancet, Nature* and *Science* recommended including the file extension. *Science* did not accept figures broken into parts, and suggested not dividing files into parts such as “smithfig1a.ps”.

Concerning the journals of computer science, only *Medical Image Analysis* specified a preferred file size (7-10 MB maximum). In this discipline four journals provided recommendations on image file naming. All of them suggested including the number that represents the sequential location of the image in the article. *IEEE Communications Surveys and Tutorials* suggested the inclusion of the author’s last name in the file name. *Medical Image Analysis* recommended the inclusion of the file extension in the file name “to enable quick and easy format identification”.

Concerning the mathematical journals, only *Fixed Point Theory and Applications* specified a preferred file size (10 MB maximum). In this discipline, five journals provided recommendations on image file naming. One of them suggested 20 characters as its maximum length. Three journals recommended the inclusion of the author’s last name in the file name.

Table XIX sums up the recommendations concerning the size and the name of the image file to be submitted as stated in the guidelines.

Table XIX Preferences of the journals about the size and the name of the image file to be submitted. The citations of the recommendations about the image file name were extracted from the submission guidelines for authors of each journal and accessed on the date indicated in Table XII.

Journal	Image File Size	Image File Name
MED1		“figure # (ie, figure 1, figure 2, etc.), numbered consecutively according to the order in which they are cited in the text”
MED2		
MED3		“Ensure the file name includes the correct figure number”
MED4	less than 2-3 MB	
MED5		“Examples: FIG1.TIF = figure 1 in TIFF format - SC4.EPS = scheme 4 in EPS format - PL2.TIF = plate 2 in TIFF format”
MED6	less than 2-3 MB	
MED7	less than 2-3 MB	
MED8		“CorrespondingAuthorSurname_fig1.jpg”
MED9	high-resolution file (not specified)	
MED10	less than 2-3 MB	
MED11	1 - 2 MB - 20 MB (max)	

Journal	Image File Size	Image File Name
MED12		“Examples of acceptable file names: smithtext.doc -smithtextfigs.prn - smithfig1.eps - smithtextfigs.ps - smithms.ps - smithsupp.pdf”
CS1		“Starting with the first 5 letters of the author’s last name, then the number that represents the sequential location of the image in the article. Example: in author “Anderson’s” paper, the first three figures would be named ander1.tif, ander2.tif, and ander3.ps. Please do not use descriptive names”
CS2		2Examples: figure1, figure2, figure3 or fig1, fig2, fig3”
CS3		
CS4		
CS5		
CS6		
CS7	7 - 10 MB max	“Ensure numbering, type and format is reflected in the file name. Ensure that the file extension is present to enable quick and easy format identification”
CS8		
CS9		
CS10		“Name your figure files with "Fig" and the figure number, e.g., Fig1.eps”
MATH1		“no longer than 20 characters, including an extension. File names should be specific and descriptive, not generic: smith-fig-3.eps is better than fig3.eps”
MATH2		“filenames combine a manuscript identification number, the figure number and part, and a suffix representing the filetype, e.g.: author name>_<figure number>.<suffix>. Example: “darimont_f1a.eps” would be Figure 1, part (a) accompanying manuscript by Darimont. Try to avoid generic names like Fig1.eps”
MATH3		
MATH4		
MATH5		“filenames combine a manuscript identification number, the figure number and part, and a suffix representing the filetype, e.g.: author name>_<figure number>.<suffix>. Example: “darimont_f1a.eps” would be Figure 1, part (a) accompanying manuscript by Darimont. Try to avoid generic names like Fig1.eps”
MATH6		“Name your figure files with "Fig" and the figure number, e.g., Fig1.eps”
MATH7		“Name your figure files with "Fig" and the figure number, e.g., Fig1.eps”
MATH8		
MATH9		
MATH10	10 MB max	

Six biomedical journals presented guidelines about the preferred image file size, while this information was not provided in journals in computer science and mathematics (besides one case in each one). There was no agreement concerning the file size, ranging from less than 3 MB to 20 MB.

Concerning the name of the file, all the guidelines agreed in recommending the inclusion of the figure number. Five journals distributed among the three disciplines suggested including the author surname; three (two in computer science, one in mathematics) recommended the inclusion of the image file format in the file name. Any of these data was related to the image content.

Multi-panel images

Multi-panel images usually have a high density of semantic content and require specific treatment for tagging and referencing them in the caption in order to make clear the localization, order and relations between different images included in the same figure. They create additional barriers to access and interpretation by people with visual or cognitive impairments because of the complexity of visual representation. A good accessibility recommendation would be to minimize their use to clearly defined cases.

Seven of the reviewed guidelines in biomedical articles offered specifications concerning multi-panel figures. Generally they suggested not submitting single panels as individual files. Many journals suggested marking the sequence of the panels with a capital letter, avoiding if possible the use of subpart letters (e.g., A, B, C, D, E instead of A, B, C (a), C(b), C(c)). Some journals recommended organizing panels so that the different parts were clearly recognizable and the essential details of the figure are visible on the printed page at the smallest size.

Two guidelines of journals in computer science (*IEEE Communications Surveys and Tutorials* and *IEEE Transactions on Pattern Analysis and Machine Intelligence*) referred to the inclusion of specific text labels in the image. For example, it was recommended to identify parts of figures by labels (e.g. (a), (b), (c)) within the figure by a Times Roman font with an 8-point size.

Four guidelines of mathematical journals referred to specific use of text labels in the image and referencing multi-panel images in captions.

Table XX Specifications on the submission of multi-panel images details the recommendations about multi-panel images.

Table XX Specifications on the submission of multi-panel images. The citations of the recommendations about the multi-panel images were extracted from the submission guidelines for authors of each journal and accessed on the date indicated in Table XII.

Journal	Multi-panel image
MED1	"Files with individual panels not accepted. Figures with multiple parts should be labeled and referred to as (a), (b), (c), etc. If there are several parts to a figure, label them as 3.4(a), 3.4(b), 3.4(c), etc."
MED2	"Multipanel ICMs labelled as Panel A, Panel B, etc."

Journal	Multi-panel image
MED3	"Refer to parts of the figure as (a), (b), (c), etc. If further distinction is needed, subparts can be described as (left), (right), (top), (middle), and (bottom). In addition to individual figure files, provide a PDF file containing all figures."
MED4	
MED5	
MED6	
MED7	
MED8	"Unnecessary figures and parts (panels) of figures and tables should be avoided. Not more than one panel unless the parts are logically connected; each panel of a multipart figure should be sized so that the whole figure can be reduced by the same amount and reproduced on the printed page at the smallest size at which essential details are visible."
MED9	"Figures should not contain more than one panel unless the parts are logically connected; each panel of a multipart figure should be sized so that the whole figure can be reduced by the same amount and reproduced on the printed page at the smallest size at which essential details are visible."
MED10	
MED11	"Don't send figure panels as individual files. Each figure legend should have a brief title that describes the entire figure without citing specific panels followed by a description of each panel."
MED12	"For figures with more than one part, label parts ""a"", ""b"" etc, and create a PDF scan of the whole figure to show preferred layout. On each composite include the corresponding author's name, Nature reference number, when known, and the figure number."
CS1	"If your figure has multiple parts, please embed callouts to identify parts of figures, i.e., (a), (b), (c), within the figure. Callouts should be in a Times Roman font with a consistent 8-point size."
CS2	"Figures with multi-image parts are preferred to be submitted as a single file. If possible please avoid submitting your figures in multiple parts such as fig1a, fig1b, fig1c, etc."
CS3	
CS4	
CS5	
CS6	
CS7	
CS8	
CS9	
CS10	
MATH1	"Multiple-part figures should be configured as one figure in a graphics program, not in TeX. Aligning multiple-part figures is very difficult in TeX. It is easier and more cost-effective to do so in a graphics program."
MATH2	"If there are several parts to a figure, label them as 3.4(a), 3.4(b), 3.4(c), etc." "If a figure has several parts, ensure that all parts are explained in the caption."
MATH3	
MATH4	

Journal	Multi-panel image
MATH5	"Multiple-part figures should be configured as one figure in a graphics program, not in TeX." "Aligning multiple-part figures is very difficult in TeX. It is easier and more cost-effective to do so in a graphics program."
MATH6	
MATH7	
MATH8	
MATH9	
MATH10	"Multi-panel figures (those with parts a, b, c, d etc.) should be submitted as a single composite file that contains all parts of the figure. (see the example)".

Conclusion about the submission policy

The results about the submission policies showed that currently all the journals analyzed among the three fields of research provided guidelines explaining how to submit images in articles. Especially journals in biomedicine and computer science, offered detailed guidelines specific for the preparation of figures for articles. These results seem to indicate that currently the use of images in academic articles in all the three STEM fields analyzed is important.

However, there was no agreement among publishers of journals in the same research field and among different research fields concerning image submission requirements. Only few journals explicitly declared to follow international guidelines for the submission of academic articles, suggesting that currently there is no standard for the image submission requirements.

Concerning the technical requirements of the images, some similarities have been detected in the submission format, being the most accepted formats TIFF, JPEG, EPS and PDF among the journals in the three fields.

Other similarities have been found in the dimensions of the image, since they seem to be related to the width of one or two columns of the standard page size used in the publication; in the resolution, which showed similar trends among the three disciplines; in the file naming, where all the journals recommended including the figure number.

Only a little minority of guidelines of the journals in the three research fields mentioned image accessibility issues and their impact on readers with special needs. For example, only two journals (in medicine) accepted SVG format, which supports accessibility features. As noted before, this result suggests a possible lack of awareness of the image accessibility among STEM publishers.

According to the results of the analysis, the answer to the research question “Are there effective instructions for authors to make their images accessible in STEM and in particular in biomedical academic articles?” (Q5) is negative.

5.1.1.3 Instructions concerning image-related texts in the article

Image title and caption

As commented in section 3.5.1, the title and the caption of an image convey visual information in a textual way and can help blind people to locate the image and access its content. They can also be useful for interpreting the figures for readers with cognitive impairments or even readers with low visual literacy (people who have limited skills in making meaningful interpretations of visual stimuli). As specified in the section dedicated to the use of color, it is important to avoid referring to image features only by specifying color or position, as people with color blindness, blind people and readers using black and white print versions of the image may not be able to understand information provided in this way.

Concerning the tile among biomedical journals, six out of twelve journals recommended the inclusion of a concise title for the images, and for example *the New England Journal of Medicine* suggested a title of eight words maximum. The guidelines of *Nature* and *Cell* suggested including the image title at the beginning of the caption.

All the other guidelines talked about caption. All guidelines encouraged authors to apply a more descriptive caption of the image, complementing its title. The guidelines of *Nature* and *Cell* suggested making the caption brief. The specification for the length of the caption varied greatly, ranging from the general recommendation to keep it succinct (and give more detail discussion for the text) to keep it to a maximum of 200 words. *Nature* recommended a different number of words depending on whether the papers include details of methods (in which case the caption could use up to 100 words, and must not relate to methods) or not (in which case the caption could use up to 300 words). Some guidelines offered advices about the content of the caption: for example, the *New England Journal of Medicine* suggested including “relevant clinical information, including a short description of the patient's history, relevant physical and laboratory findings, clinical course, response to treatment (if any), and condition at last follow-up” (The New England Journal of Medicine, 2012). This journal also asked authors to “describe and clearly indicate all modifications, selective digital adjustments or electronic enhancements in all digital images” (Massachusetts Medical Society, 2015), in order to ensure veracity of submitted images. Four guidelines recommended describing and explaining in the caption any text label indicating some specific structure in the figure. The same four guidelines recommended including callouts corresponding to each panel (A, B, C, etc.) for multi-panel images. *Cell* specified that “each figure caption should have a brief title that describes the entire figure without citing specific panels followed by a description of each panel” (see the section “Multi-panel images”). Four guidelines specified that keys and additional explanations should be added in the caption for the nomenclature, abbreviations, arrows, symbols, letters and units used in a figure.

Two guidelines offered an advice on captions about images presenting specific data types, such as pooled data and diagrams. The *Annual Review of Immunology* and *Science* referred to specific types of image such as schemes (e.g., structural chemical formulas) or complex equations. For example, the *Annual Review of Immunology* specified that "equations and chemical structures that cannot be typeset in one or two lines are considered art. If possible, such equations should be submitted in math-friendly software applications such as LaTeX or MathType" (Annual Reviews, 2014).

Concerning the journals in computer science, seven journals explicitly required to include a figure caption. Only two journals (*the Journal of Statistical Software* and *the International Journal of Computer Vision*) provided suggestions on caption content. The guidelines of *The International Journal of Computer Vision* recommended authors "to identify all elements found in the figure caption".

Concerning the mathematical journals, one journal recommended the inclusion of a title of 15 words maximum. Six journals explicitly required to include an image caption. Three of them specified the length of the caption. The guidelines of five journals offered advices about the content of the caption. *Communications on Pure and Applied Mathematics* recommended adding a key to the caption if there are arrows, symbols, letters or abbreviations not used elsewhere. *Inventiones Mathematicae* and *Foundations of Computational Mathematics* recommended identifying all elements found in the figure in the figure caption, while *Fixed Point Theory and Applications* recommended incorporating figure text labels directly into the figure.

Table XXI sums up the results related to title and caption requirements and their suggested length.

Table XXI Recommendations related to title and caption requirements and their suggested length.

Journal	Image title		Image Caption	
	Required	Length	Required	Length
MED1		Not specified	x	Succinct
MED2	x	8 words max.	x	150 words max.
MED3	x	concise	x	
MED4			x	200 words max.
MED5			x	
MED6	x	concise	x	200 words max.
MED7	x	concise	x	200 words max.
MED8		brief	x	100 - 300 words max.
MED9			x	
MED10	x	concise	x	200 words max.
MED11	x	brief	x	
MED12	x		x	200 words max.

Journal	Image title		Image Caption	
	Required	Length	Required	Length
CS1			x	
CS2			x	
CS3			x	
CS4				
CS5				
CS6				
CS7			x	
CS8			x	
CS9			x	A sentence
CS10			x	Concise
MATH1			x	
MATH2			x	Succinct
MATH3				
MATH4				
MATH5			x	
MATH6			x	Concise
MATH7			x	Concise
MATH8				
MATH9				
MATH10	x	15 words max.		300 words max.

Table XXII presents a summary of the specific information concerning the content of the image that should be included in the figure caption according to the guidelines analyzed.

Table XXII Specific information requested in the figure captions: references to labeled structures; references to nomenclature, abbreviations, arrows and other symbols and letters; specifications of particular data types.

Journal	Specific information concerning elements of the figure included in caption		
	labeled structures	nomenclature, abbreviations, arrows, symbols, letters and units	specific data types
MED1	x	x	
MED2	x		
MED3			x

Journal	Specific information concerning elements of the figure included in caption		
	labeled structures	nomenclature, abbreviations, arrows, symbols, letters and units	specific data types
MED4			
MED5			
MED6			
MED7			
MED8		x	
MED9		x	
MED10			
MED11	x		
MED12	x	x	x
CS1			
CS2			
CS3			
CS4			
CS5			
CS6			
CS7			
CS8			
CS9			
CS10	x	x	
MATH1			
MATH2		x	
MATH3			
MATH4			
MATH5			
MATH6	x	x	
MATH7	x	x	
MATH8			
MATH9			
MATH10		x	

The results show that the guidelines provided by biomedical journals about titles and captions of the figures are generally more detailed than those provided by the guidelines of journals in computer science and mathematics. However, the majority of guidelines agree in recommending captions as a brief text with a maximum length of 100 to 300 words. The most cited information that should be included in captions is the description of structures in the image marked by labels and the explanation (a key) of the labels found in the figure, such as arrows, symbols, letters or abbreviations.

Labeling

From the point of view of accessibility, labels are a source of extra information related to images, so they can be used as a support for interpreting them. Font size is a central accessibility issue for people with low vision and the WCAG 2.0 guidelines and other guidelines for clear print (UKAAF, 2012b) encourage the use of a minimum 12 pt size for body text or text in images. To ensure best readability for readers with low vision or cognitive impairments and readers using small screens, sans-serif fonts such as Verdana and Arial are better considered than serif fonts such as Times New Roman or Courier (Ranaldi & Nisbet, 2010)(Rello & Baeza-Yates, 2013).The application of standard abbreviations and units in images has a positive effect on the readability of labels and other text associated with images, since it standardizes the conventions in labeling and makes the images comprehensible for the whole community of researchers.

In the guidelines of biomedical journals analyzed, an important recommendation about image labels was to maintain the labels separated from the figure, in order to make them easier to edit during the publishing workflow. Three guidelines explicitly recommended ensuring that text and other elements in the figures remain editable. The guidelines of the *New England Journal of Medicine,* the *Annual Review of Immunology* and *The Lancet* suggested providing two versions of the figure, one with the appropriate labels (text, arrows, etc.) and one without them, or alternatively creating the figure with layers, one with visuals and one with labels. The visual would be used as it is, while the publisher would make label adjustments to adapt to house style or sizing changes. The *New England Journal of Medicine* asked for the inclusion of an arrow indicating the top of the image in each original figure. *Science* recommended avoiding labels that are “not absolutely necessary for understanding the figure” and explaining them in the caption. In the case of labels with very small typography (e.g., units for scale bars) or data presented in small tables or histograms, *Nature* and *Nature Genetics* recommended presenting the information briefly in the text of the caption. For letters and labels, a sans-serif font (Arial or Helvetica) and font sizes between 7 and 9 pt were recommended by four guidelines. The guidelines of the *Annual Review of Immunology* and *Nature* recommended a font size no smaller than 5-6 pt in order to ensure readability in print. Three journals expressed preferences to capitalize the first label letter (in particular for the labeling of panels). Six journals referred to the use of standard abbreviations and units in the image and to spell out unusual units or abbreviations.

Science referred to some general principles of information visualization for creating figures ("In laying out information in a figure, the objective is to maximize the space given to presentation of the data" and "Avoid wasted white space and clutter") (Science, 2012).

Concerning the journals in computer science, *IEEE Transactions on Pattern Analysis and Machine Intelligence* and the *International Journal on Computer vision* suggested providing the labels separated from the figure, in order to make them easier to edit. For letters and labels, a sans-serif font (Arial or Helvetica) was recommended by five journals. Three journals recommended Times New Roman font (serif), and two suggested Courier (serif). The label font size recommended by six guidelines ranged between 7 and 12 pt. Four journals recommended the use of standard abbreviations and units and? to spell out unusual abbreviations and units.

Among the mathematical journals, only *Inventionae Mathematicae* and *Foundations of Computational Mathematics* recommended providing the labels separated from the figure. The font families suggested were Helvetica and Arial (three journals), Times New-Roman (two journals) and AMS fonts (two journals). The label font size recommended by five guidelines ranged between 10 and 12 pt. Three journals recommended the use of standard abbreviations and units and to spell out unusual abbreviations and units.

Table XXIII shows the preferences for labels specified by the guidelines analyzed.

Table XXIII The table shows whether the journal specifies the preference for labels separated from the figure submitted, the recommended font size and font family of the image's labels.

Journal	Image Labeling			
	Editable label	**font size**	**font family (according to the citation order)**	**standard abbrev. and units**
MED1		6- 9 pt	Arial, Helvetica	x
MED2	x			
MED3	x	7 - 9 pt. (6 pt min.)	Sans-serif fonts (Arial, Helvetica, Myriad Pro)	x
MED4				
MED5	x			x
MED6				
MED7				
MED8		7 -8 pt. (5 pt min.)		x
MED9			Arial, Helvetica	x
MED10				
MED11			Helvetica	

Journal	Image Labeling			
	Editable label	font size	font family (according to the citation order)	standard abbrev. and units
MED12				x
CS1		9-10 pt	Arial, Helvetica, Times New Roman, Cambria	x
CS2	x	8 pt	Arial, Helvetica, Times New Roman, Courier, Symbol, ITC Zapf Dingbats	
CS3		10 pt	Arial, Helvetica	x
CS4				
CS5				
CS6				
CS7		7 - 10 pt	Arial, Times New Roman, Courier, Symbol (or similars)	x
CS8				x
CS9		about the size of the caption		
CS10	x	8 - 12 pt	Sans-serif fonts (Arial, Helvetica)	
MATH1		10/12 pt (min.)	match those used in text (Modern Modern, AMS fonts)	
MATH2			Helvetica, Arial, Times, Times New Roman	
MATH3				
MATH4				
MATH5		10/12 pt (min.)	match those used in text (Modern Modern, AMS fonts)	
MATH6	x	8 - 12 pt	Sans-serif fonts (Arial, Helvetica)	x
MATH7	x	8 - 12 pt	Sans-serif fonts (Arial, Helvetica)	x
MATH8				
MATH9				
MATH10		8pt (min.)		x

In synthesis, only 7 journals out of 32 specified the preference for providing labels separated from the figure submitted. No agreement was found about the recommended font size of the labels, ranging from a minimum of 5 pt to a maximum of 12 pt. Most recommended font types among the journals in the three disciplines were Helvetica and Arial (sans-serif fonts) and Times New Roman (serif font).

Mentions

As mentioned in the introduction of the thesis, references to the image from the article text could include relevant information related to the image content, which may be useful for all kinds of readers to locate, read and understand it.

The guidelines of two biomedical journals (*A Cancer Journal for Clinicians* and the *Annual Review of Immunology*), of three journals in computer science (*IEEE Communications Surveys and Tutorials, MIS Quarterly* and *Journal of Statistical Software*) and of two mathematical journals (*Communications on Pure and Applied Mathematics* and *Annals of Mathematics*) explicitly required authors to reference the figures in the article text (Table XXIV). Despite these results, the great majority of the images in the articles reviewed were referred in the main text (Table XXIV). This result suggests that the use of mentions is currently a common practice adopted by authors of scientific articles.

Table XXIV Recommendations in journal guideline to call out figures in the article text and images referred to in the text of the articles analyzed.

Journal	Guidelines specifying to call out all figures in the text in article text	Images referred in the text of the articles analyzed	
		number out the total number of figures	percentage
MED1	x	8 out of 8	100%
MED2		3 out of 3	100%
MED3	x	3 out of 3	100%
MED4		2 out of 2	100%
MED5		0 out of 1	0%
MED6		2 out of 2	100%
MED7		3 out of 4	75%
MED8		4 out of 4	100%
MED9		8 out of 8	100%
MED10		3 out of 3	100%
MED11		7 out of 7	100%
MED12		4 out of 4	100%
CS1	x	7 out of 10	70%
CS2		7 out of 9	78%
CS3		1 out of 2	50%
CS4		19 out of 24	79%
CS5		1 out of 1	100%
CS6	x	1 out of 1	100%

Journal	Guidelines specifying to call out all figures in the text in article text	Images referred in the text of the articles analyzed	
		number out the total number of figures	percentage
CS7		7 out of 10	70%
CS8		8 out of 10	80%
CS9	x	11 out of 13	85%
CS10		16 out of 16	100%
MATH1		5 out of 5	100%
MATH2	x	12 out of 14	86%
MATH3	x	1 out of 1	100%
MATH4		9 out of 10	90%
MATH5		2 out of 12	17%
MATH6		1 out of 1	100%
MATH7		0 out of 2	0%
MATH8		4 out of 5	80%
MATH9		59 out of 59	100%
MATH10		3 out of 3	100%

Metadata

As commented in section 3.5.4 metadata provides descriptive, administrative and structural information to images in academic articles. They may also serve as a means to provide text alternatives and contribute on the improvement of image accessibility in digital documents.

No guideline of the journals analyzed was found to provide specific reference to the inclusion of metadata to the images for submission. A review of the metadata included in the figures of the articles analyzed will be reported in section 5.1.3.

Conclusions on guidelines about image-related texts

The results showed that the guidelines provided by STEM journals did not follow common requirements about image-related texts, neither among journals in the same fields nor among journals in different fields.

However, some similarities were found. For example, the inclusion of an image caption was explicitly required in the majority of the guidelines, supporting the observation of the importance of the use of image captions in STEM articles, as commented in section 4.4.1. Furthermore, the majority of guidelines agreed in identifying captions as a brief text with a maximum length of 100 to 300 words. Concerning the labels and text in the image, the journals in the three disciplines generally agreed in recommending Helvetica, Arial and Times New Roman font types

Concerning the mentions in the main text, while few journals explicitly required authors to reference the figures in the article text, the great majority of the images in the articles reviewed were referenced in the main text, suggesting that this is a common practice adopted by authors of scientific articles.

The lack of references to the inclusion of metadata to images suggested a general lack of awareness about the use and function of metadata among the publishers in the three fields analyzed.

The characteristics of the image-related text included in the articles will be reviewed in detail in section 5.1.3.

5.1.1.4 Image presentation and retrieval (only for biomedical articles)

As a complementary analysis, the options of image presentation and retrieval were analyzed in biomedical articles. Although the thesis aims to study STEM articles in general, due to time constraints this analysis was focused on biomedical publications.

Image presentation options

In all of the biomedical journals' websites, images were presented within the article to which they belong. The images were offered to readers in HTML and PDF version in all the papers analyzed. While images in the PDF version were always presented without further options for visualization, variations were identified in the options of presentation of images in the HTML versions (Table XXV).

Table XXV Alternative image presentations in HTML articles. The zoom option refers to the option of viewing an high-resolution version of the figure by zoom.

Journal	Alternative image presentations in HTML articles				Download options	
	Enlarged version		Visualization options			
	Pop up with caption	On another page with caption	Contextual menu for image navigation	Zoom	Download as PPT	Download as PDF file
MED1	x	x			x	x
MED2	x	x		x	x	x
MED3	x	x	x	x	x	x
MED4	x	x			x	x
MED5	x	x	x		x	x
MED6		x			x	x

Journal	Alternative image presentations in HTML articles				Download options	
	Enlarged version		Visualization options			
	Pop up with caption	On another page with caption	Contextual menu for image navigation	Zoom	Download as PPT	Download as PDF file
MED7		x			x	x
MED8		x			x	x
MED9		x			x	x
MED10		x			x	x
MED11	x		x			x
MED12		x			x	x

Five journals offered an enlarged version of the images in a pop-up window. In one case (*New England Journal of Medicine*) the pop-up was written in Flash: this technology introduces specific accessibility barriers and requires several techniques that are not easy to implement for making image content readable by assistive technologies. All journals showed the enlarged version including the caption of the image and two journals had a zoom option to enlarge the images, but not always reaching the 200% recommended zoom scale on accessibility guidelines. The *Annual Review of Immunology* presented a contextual menu showing a summary of images in the article. *Cell* and *The Lancet* accessed through the ScienceDirect interface also offered a similar menu for images.

Eleven journals offered the option to download the image directly in a Power Point version (Table XXV). Different image presentation options have consequences on different kinds of readers accessing the content of the image. Blind people or people with low vision may find it difficult to recognize a visual context change, such as a new window popping up; on contrast the zoom option is a useful tool for low-vision readers. Downloading the image as a PPT or PDF could allow its visualization with specific tools owned by readers that enhance/change contrast or zoom and facilitate access.

The options reviewed did not seem to particularly address the needs of readers with visual impairments. However, they could be useful for some of them. The existence of options for customization could be a pathway to enriched accessible presentation options.

Image retrieval options

The analysis of the retrieval system of the journals' website showed that only the *Annual Review of Immunology* explicitly offered a search option addressed to images using words in figure captions. This option was integrated in the advanced search interface. The *New England*

Journal of Medicine search interface retrieved keywords in the whole article (captions included), and offered the option of restricting the results to images. The ScienceDirect portal enriched the retrieval system of *The Lancet* and *Cell* providing searches for images by caption text and whole article text (Table XXVI).

Outside the publishing world, the Google image interface allows queries to be made using images as examples, but this option was not offered by any of the retrieval systems considered nor even in Google Scholar.

Table XXVI Search options of the retrieval systems offered by the online version of the journals analyzed. None of the journals offers the search by figure labels or by example image.

Journal	Explicit scope of searches offered by search interfaces							
	by keywords						by figure type/color	by example figure
	in article title	in article text	in article abstract	in article keywords	in figure caption	in figure labels		
MED1	x	x	x	x				
MED2	x	x	x		x			
MED3	x	x	x	x	x			
MED4	x	x			x			
MED5	x	x	x	x	x			
MED6	x	x			x			
MED7	x	x			x			
MED8	x	x			x			
MED9	x	x			x			
MED10	x	x			x			
MED11	x	x	x		x			
MED12	x	x	x		x			

In addition to the explicit options for retrieval offered in the search interface, tests were performed to detect the actual ability to retrieve the information included in captions, image labels and alternative texts through global searches (i.e., in article text or keywords). The results, also showed in Table XXVI, emphasized that while all the journals retrieved words in captions, none of them retrieved words in either alternative text or labels of figures, which are texts that tend to be semantically rich.

5.1.2 Analysis of the articles

This section presents the results of the analysis of the articles in biomedicine, computer science and mathematics. In particular, the results of the analysis of the following issues are presented and discussed:

- General information about the PDF document of the whole article: total number of pages; figure count per document and page; inclusion of structural tags, level of encryption; version; application or software used to create/edit and convert the original document to PDF.
- General information about figures in the PDF and HTML versions of the articles: total number of figures in the articles and image type (raster/vector).
- Specific elements of the figures that determine the perception and understanding of the images by users with disabilities, as defined according to the WCAG 2.0 (see 4.1.4):
  - Text alternatives of figures in the PDF and HTML versions of the articles.
  - Technical features of the image: use of color, minimum color contrast, resolution and dimensions of the figure, font size and family of the textual labels in the image and in caption/mentions, the presence of acronyms and abbreviations in the image and in caption/mentions.

The main aim of this analysis is to verify the assumption that images in biomedical and other academic publications are not accessible and, in particular, images don't have text alternatives associated to them and answer to the following questions:

Q3. Are figures in STEM academic articles and in particular in biomedical articles currently accessible?

Q4. Are images in STEM academic articles and in particular in biomedical articles a relevant source of information?

5.1.2.1 General information about PDF documents

The results show that only four articles were tagged (14%), although the PDF version of 29 articles (97%) could support structured tags (Table XXVII).

Table XXVII Percentage of PDFs with a version supporting tags and tagged PDFs per discipline.

	Discipline			**Total**
	BioMedicine	**Computer Science**	**Mathematics**	
PDF version supporting tags	100%	90%	100%	97%
Tagged PDFs	40%	0%	0%	14%

Two articles in computer science were encrypted with specific permissions for screen readers.

The software applications used for PDF document creation and production in academic articles are shown in Table XXVIII.

Table XXVIII Tools used to create tagged PDF articles in different disciplines.

Creator tool	**Producer tool**	**article discipline**	**Tagged PDF**

Creator tool	Producer tool	article discipline			Tagged PDF
		MED	CS	MATH	
Adobe InDesign CS2	Adobe PDF library	1			no
Adobe InDesign CS4	Adobe PDF library	1			no
Adobe InDesign CS5	Adobe PDF library	4			yes
Adobe InDesign CS6	Acrobat Distiller			1	no
Arbortext Advanced Print Publisher	PDFlib PLOP	1			no
	Acrobat Distiller	1	2		no
LaTeX with hyperref package	Acrobat Distiller		1	2	no
	PDFlib PLOP			1	no
	pdfTeX		1	2	no
VteX PDF Tools	Acrobat Distiller		1	3	no
PScript5.dll	Acrobat Distiller		1		no
dvips	Acrobat Distiller		2		no
LaTex2e	Acrobat Distiller			1	no
not available	Acrobat Distiller	1	2		no
	iText 2.1.5	1			no

Four articles were created by *ArborTExt Advanced Print Publisher* and none of them was tagged. Starting with version 10, released in 2009, the software is supposed to support PDF tags. Seven articles were created in LaTeX format with the hyperref package and converted to PDF by different producers. Four articles were created by the VTeX software, a TeX system that offers native support for PostScript, SVG, PDF and HTML. Two articles were converted from DVI file to PostScript by the dvips program and then converted to PDF by Acrobat Distiller, and none of them was tagged. One article was created using PScript5.dll and it was not tagged. One article was created in LaTex2e and converted to PDF by Acrobat distiller. None of the PDF documents converted by Acrobat Distiller was tagged. Currently, the conversion from LaTex to PDF does not generate tags automatically (Moore, 2009). Possible solutions have been proposed by Hagen (2010), Schalitz (2007) and Strübing (2012), but they are neither widely adopted nor straightforward.

All of the four tagged articles were published by the Nature publishing group (biomedical field) and were created by Adobe InDesign software and converted by the Adobe PDF Library.

These results show that most PDF articles are not correctly tagged. Therefore, figures in PDFs cannot be identified by assistive technologies and they are not ready to support the inclusion of alternative text.

5.1.2.2 General information about images

Figures were analyzed in quantitative terms. The analysis of the appropriateness and quality of the figures were not taken into account in this part of the study.

Quantitatively speaking, figures were found to be an important part of the articles quantitatively speaking, although the total number of figures in the articles revealed important differences among publications in different disciplines (Table XXIX).

Table XXIX Average number of figures per article and page in different disciplines.

	Figure count	Page count	Figures per article	Figures per page
Biomedicine	38	97	4	0,3
Computer Science	95	225	10	0,42
Mathematics	112	514	11	0,22
Total	245	836	8	0,31

Biomedical articles included fewer figures per article than articles in computer science and mathematics. However, page count of biomedical articles was also far shorter. On the contrary, the average number of figures per page was one per 3.33 pages in biomedical articles, one per 2.4 pages in computer science and one per 4.5 pages in mathematics.These results showed that figures are currently an important element of information commonly used in STEM articles.They answer affirmatively to the research question Q4 (Are images in STEM academic articles and in particular in biomedical articles a relevant source of information?).

The use of vector vs raster graphics in the analyzed articles was quantified. Figure 35 shows that vector graphics accounted for more than 70% of the images in medical articles and 96% of the images in mathematical articles. Images in computer science articles were 41% raster graphics, 34% vector graphics and 23% a combination of both.

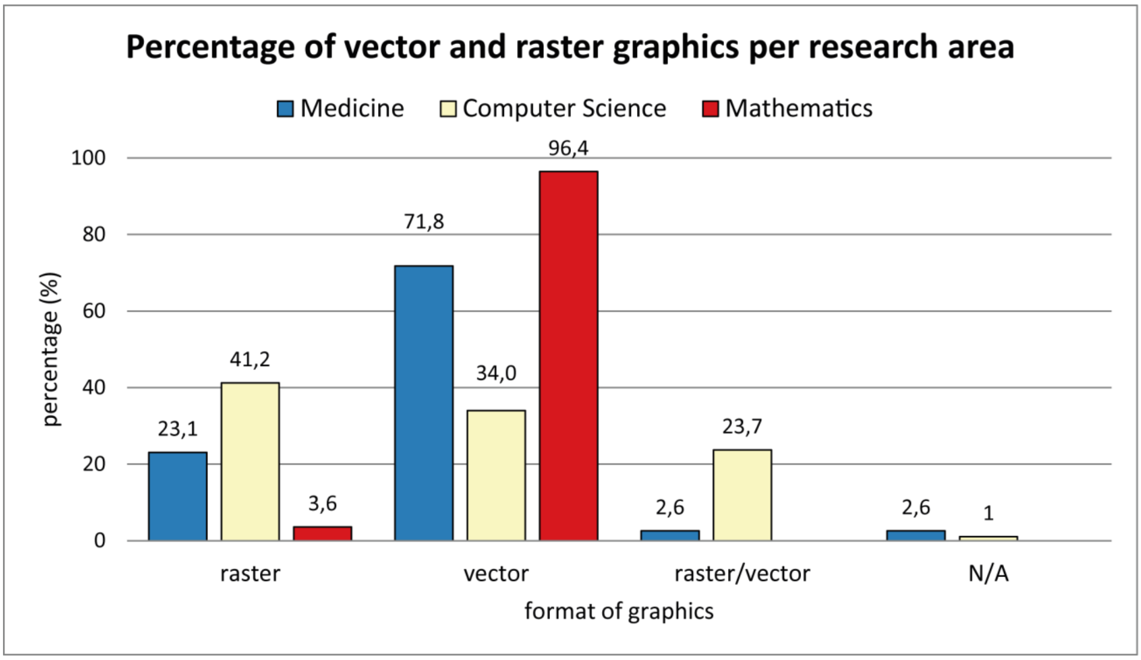


Figure 35 Percentage of vector and raster graphics among articles in different disciplines. Some images marked as figures in the articles were considered as tables and their format was not evaluated (N/A).

It was not possible to retrieve the original format of the images, since raster images are converted to the PDF imaging model (except in the case of JPEG images) and all of them are exported as JFIF.

5.1.2.3 Application of text alternatives in HTML and PDF

The images contained in the 30 articles selected were analyzed in detail to review the actual implementation of textual alternatives to the figures. Specifically, the analysis of articles was focused on the proper application of the *alt* and *longdesc* attributes of the <img> tag (in the HTML version of the paper) and the *alternate text* and *actual text* attributes of the <figure> tag (in the PDF version).

In the biomedical articles reviewed, it was detected a general incorrect use and even the absence of the *alt* attribute and the absence of the *longdesc* attribute in the HTML versions of the articles and the absence of *alternate text* and the *actual text* attributes in PDF versions (Table XXX). In nine papers the *alt* attribute was redundant (five of them used the title of the figure as *alt* text) or uninformative (e.g., "thumbnail image", "full-size image", "20 K", "Fig. 1", etc.). In the articles of the *New England Journal of Medicine* and the *Annual Review of Immunology* the value of the *alt* attribute was empty. In the articles of *Nature Reviews Molecular Cell Biology* and *Nature Reviews Genetics* the *alt* attribute included apologies for the lack of application of a suitable alt text, accompanied by the invitation for users to contact the publisher if they are unable to access the image ("Unfortunately we are unable to provide accessible alternative text for this. If you require assistance to access this image, or to obtain a text description, please contact npg@nature.com"). None of the figure in biomedical articles had a *longdesc* attribute in the HTML version and the *alternate text* and the *actual text* attributes in the PDF version (Table XXX).

None of the articles in computer science contained images with appropriate alternative text descriptions (Table XXX). While only three articles were offered in HTML version (from the *Journal of Chemical Information and Modeling*, *Medical Image Analysis* and *International Journal of Computer Vision*), all of them presented a redundant or uninformative *alt* attribute (e.g., abstract image" and "figure") in all the figures. Only the PDF version of article of the *International Journal of Computer Vision* contained images with *alternate* and *actual* text attributes, but they were uninformative (e.g., "image of fig.1").

None of the articles in mathematics contained images with appropriate alternative text descriptions (Table XXX). While only three articles were offered in HTML version (from the *Inventiones Mathematicae*, *Foundations of Computational Mathematics* and *Fixed Point Theory and Applications)*, all of them included uninformative *alt* attribute (e.g. "/static-content/images/393/art%_Fig1_HTML.gif") for all images. None of the articles in PDF version contained images with *alternate* and *actual* text attributes.

Table XXX Application of the *alt* and *longdesc* attributes for the images in the HTML version of the articles and application of *alternate* and *actual* attributes for the images in the PDF version of the articles.

Article	*Alt* attribute (HTML version)	*Longdesc* attribute (HTML version)	*Alternate text* attribute (PDF version)	*Actual text* attribute (PDF version)
MED1	Incorrect			
MED2				
MED3				
MED4	Incorrect			
MED5	Incorrect			
MED6	Incorrect			
MED7	Incorrect			
MED8	Incorrect			
MED9	Incorrect			
MED10	Incorrect			
CS1	N/A	N/A		
CS2	N/A	N/A		
CS3	Incorrect			
CS4	N/A	N/A		
CS5	N/A	N/A		

Article	*Alt* attribute (HTML version)	*Longdesc* attribute (HTML version)	*Alternate text* attribute (PDF version)	*Actual text* attribute (PDF version)
CS6	N/A	N/A		
CS7	Incorrect		Incorrect	Incorrect
CS8	N/A	N/A		
CS9	N/A	N/A		
CS10	Incorrect			
MATH1	N/A	N/A		
MATH2	N/A	N/A		
MATH3	N/A	N/A		
MATH4	N/A	N/A		
MATH5	N/A	N/A		
MATH6	Incorrect			
MATH7	Incorrect			
MATH8	N/A	N/A		
MATH9	N/A	N/A		
MATH10	Incorrect			

The actual application of text alternatives was also compared to the accessibility policies of the journals. Table XXXI shows that even when articles are published in journals with a specific accessibility policy, alternative descriptions for images were not used properly (uninformative, redundant or empty alt text) or were not applied at all (lack of *alt* and *longdesc* attributes in HTML versions or lack of *alternate text* and *actual text* PDF versions).

Table XXXI Comparison between the subscription to accessibility policies and the actual application of alternative text to images

Journal	Accessibility policy	reference to the use of alternative text for images	*Alt* attribute (HTML version)	*Longdesc* attribute (HTML version)	*Alternate text* attribute (PDF version)	*Actual text* attribute (PDF version)
MED1	x		Incorrect			
MED2						
MED3						
MED4	X	x	Incorrect			
MED5	X		Incorrect			

Journal	Accessibility policy	reference to the use of alternative text for images	***Alt*** **attribute (HTML version)**	***Longdesc*** **attribute (HTML version)**	***Alternate text*** **attribute (PDF version)**	***Actual text*** **attribute (PDF version)**
MED6	X	x	Incorrect			
MED7	X	x	Incorrect			
MED8	X	x	Incorrect			
MED9	X	x	Incorrect			
MED10	x	x	Incorrect			
CS1	x		N/A	N/A		
CS2	x		N/A	N/A		
CS3			Incorrect			
CS4			N/A	N/A		
CS5	x		N/A	N/A		
CS6			N/A	N/A		
CS7	x		Incorrect		Incorrect	Incorrect
CS8	x		N/A	N/A		
CS9			N/A	N/A		
CS10		x	Incorrect			
MATH1			N/A	N/A		
MATH2			N/A	N/A		
MATH3			N/A	N/A		
MATH4			N/A	N/A		
MATH5			N/A	N/A		
MATH6		x	Incorrect			
MATH7		x	Incorrect			
MATH8			N/A	N/A		
MATH9			N/A	N/A		
MATH10			Incorrect			

The results showed a widespread lack of application of different types of text alternatives in the articles in all the three fields analyzed. Even when a text alternative was applied (especially in the *alt* attribute of *img* element in HTML version), the content of the attribute was incorrect.

5.1.2.4 Analysis of accessibility issues applied in the articles

Besides text alternatives, other visual issues affecting the accessibility of the images were reviewed in the articles, according to the review of accessibility guidelines detailed in section 4.1.4:

1. the use of color as the sole mean to distinguish two or more different information items inside the image.
2. The minimum color contrast between all visual information (including text inside images) and its background. The analysis of the color contrast was performed by the Colour Contrast Analyser software v 2.2 (Vision Australia, 2014) by checking the minimum contrast ratio (4.5:1) recommended by the level AA of conformance of the WCAG 2.0.
3. The use of sans-serif font family for the text inside figures. The analysis did not apply to images without text.
4. A minimum of 12 pt of font size for text inside figures. The analysis did not apply to images without text.
5. Abbreviations in the text inside figures and in captions in articles with one or more figures with abbreviations.
6. The complexity of the text in captions and mentions, which will be discussed thoroughly in section 5.1.3.

The results of the analysis of these visual issues related to accessibility are showed in Table XXXII.

Table XXXII Accessibility issues related to figures applied in PDF articles. (*)The complexity of the text in captions and mentions will be analyzed in the next section 5.1.3.

Accessibility issue	Applied in PDF article		
	MED	CS	MATH
Use of color	8 articles out of 10	8 articles out of 10	2 articles out of 10
Minimum color contrast	1 article out of 8	4 articles out of 8	1 article out of 3
Sans-serif font type	10 articles out of 10	6 articles out of 10	0 articles out of 7
Label font size 12 pt minimum	0 articles out of 10	0 articles out of 10	0 articles out of 7
Abbreviations	9 articles out of 10	8 articles out of 10	0 articles out of 10
Complex text*			

The main differences among articles from the three research fields were:

- The mathematical articles included a reduced number of color images compared to the articles in the other disciplines.
- No abbreviations were found in the text of figures and captions in mathematical articles, while the majority of articles in biomedicine and computer science presented abbreviations.
- No mathematical article presented images with text in sans-serif font, while all the biomedical articles and more than half of articles in computer science had them.

The main similarities among articles from the three research fields were that very few images had a minimum color contrast and no image presented text with a minimum 12 pt font size.

In synthesis, the results showed a consistent lack of accessibility of the images within articles of the three STEM fields analyzed.

5.1.2.5 Conclusions concerning the images in STEM articles

In this section, the results of the analysis of STEM academic articles have been presented.

Firstly, the current accessibility of the images they include has been assessed. The results have showed that figures in scientific academic publishing are currently not accessible in three fields analyzed: they lack of appropriate text alternatives and they partially conform to the accessibility recommendations concerning the visual features of the figures. This conclusion answers negatively to the question Q3 ("Are figures in scientific academic publishing currently accessible?"). It validates the assumption that images are currently not accessible in academic articles in STEM fields and, in particular, the assumption that images do not have text alternatives associated to them.

Secondly, the results have showed that figures are currently an important element of information commonly used in the articles of the three research fields. They answer affirmatively the research question Q4 ("Are images in STEM academic articles and in particular in biomedical articles a relevant source of information?").

5.1.3 Analysis of image-related texts

This section presents the results of the analysis of the image-related texts in the biomedical articles. This additional analysis of the articles was performed to assess the function and use of image-related texts (captions, mentions, text in the figures and metadata) currently used in biomedical academic articles, with the aim of evaluating the extent to which they could be potentially appropriate as alternative descriptions of the images.

The analysis was structured in two parts:

- A quantitative analysis of the captions, mentions in the main text, text inside the image and metadata. In particular, the following information was extracted and analyzed: text transcriptions of the image-related texts, length in characters, position in the document and page number and section.

- A lexical analysis (quantitative and qualitative) of the content of captions and mentions.

This analysis aimed to answer question Q6: “Are current image-related texts appropriate as alternative descriptions of the images in STEM and in particular in biomedical academic articles?”.

5.1.3.1 General characteristics of image-related texts (quantitative analysis)

Once stated that the *alt* attribute and other text alternatives were missing or incorrect, the use of image-related texts in the academic articles in PDF was explored.

Firstly, the characteristics and use of the image-related texts were explored. The length of the text included in captions and mentions (Figure 36) and the number of images described by captions, mentions, text labels and metadata (Figure 37) was calculated. Figure 36 shows the average number of characters used for each figure in the three disciplines. The number of characters in related texts (captions and mentions) in biomedical articles was much higher than the number used in related texts in computer science and mathematics. All of them were far longer than 125 characters, a recommended length by some sources (DO-IT knowledge database, 2013).

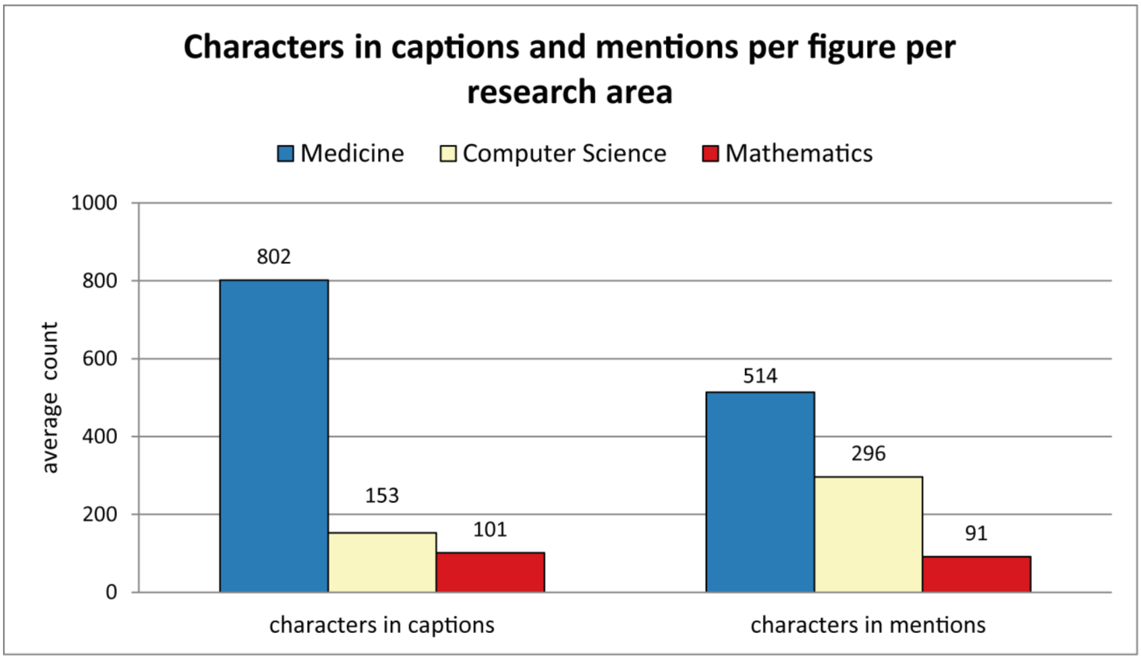


Figure 36 Average count of characters in captions and mentions per figure among articles in different disciplines.

As shown in Figure 37, captions are widely used in all three areas (describing between 95% and 86% of the images). Mentions were also very frequent in articles in medicine and computer science (describing between 88% and 82% of the images), while they were less frequent in mathematical articles (describing 65% of the images).

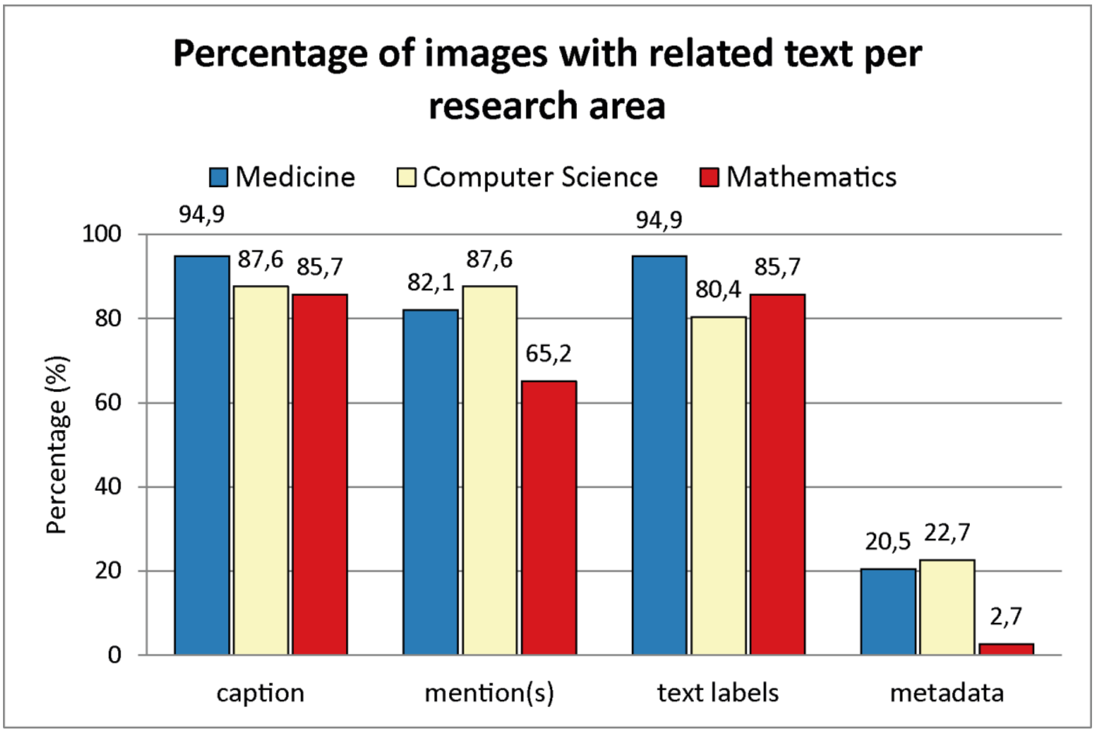


Figure 37 Percentage of images with related text in articles in different research areas.

The percentage of images with text labels was calculated for both raster and vector graphics. They were very frequent in medicine (95%), mathematics (86%) and computer science (80%). However, it is worth noting that text labels and other text in figures have some limitations from the perspective of accessibility:

- In vector graphics label text is available to technical aids but its meaning is often determined by its spatial position in the visual context of the image and, when isolated from the figure, the text results in a sequence of characters that is useless for understanding the image, and it often is written with incorrect spacing.
- In raster images the text is codified as a bitmap and only available through optical character recognition software. This text would also be meaningless for the same reasons as in vector graphics.

Metadata were applied in 20% of the figures in medicine, in 23% of the figures in computer science and in only 2% of the figures in mathematics. The metadata associated with the paths composing the vector images were not taken into account, since they are not associated with the vector graphics as descriptors of the whole content of the image. The metadata properties retrieved were mainly technical features such as dimensions, color space, size in bytes and creator tool. The only metadata describing the semantic content were uninformative *author* and *title*, with values such as “Figure 2-2012”, “figs_lineart_template”, “AuthornameFig##” “Temp#1B8AFAC”, etc.

The position of images (Figure 38) and captions in the articles was considered another important element for assessing the function of the images.

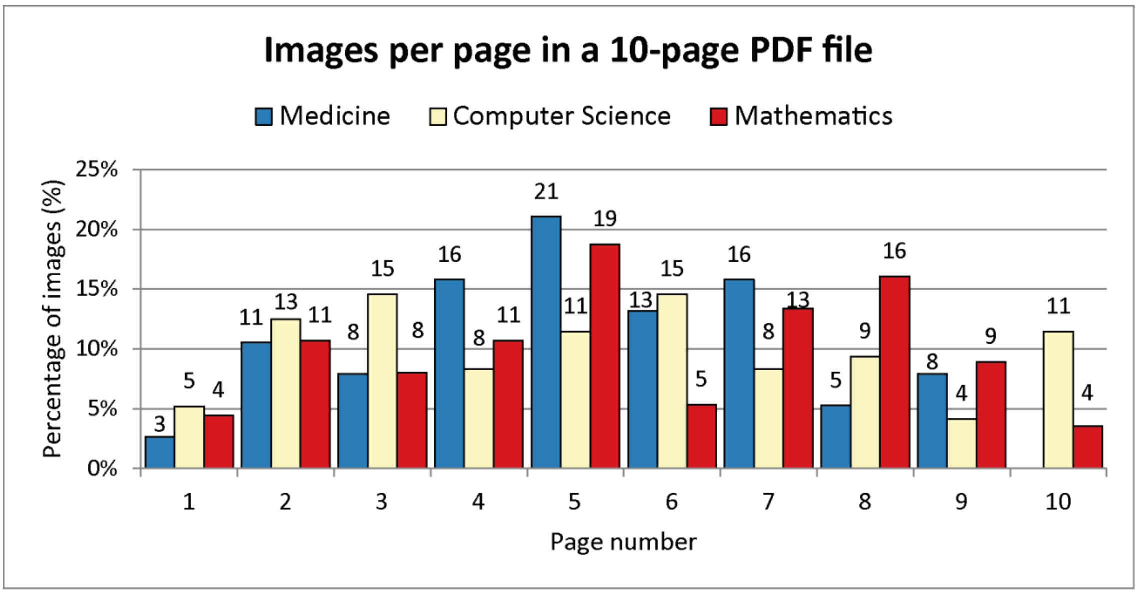


Figure 38 Frequency of images (as a percentage) per page in a PDF normalized to 10 pages in different disciplines.

We found that all the captions were located next to the images they referred to, either below or beside them. The number of mentions on the same page as the image they refer to (Figure 39) was greater than the number of mentions on other pages in medicine articles. The trend was the opposite in computer science and mathematical articles. Mentions which started on the same page as the figure and continued on the next were counted in both “same page” and “other page” sums.

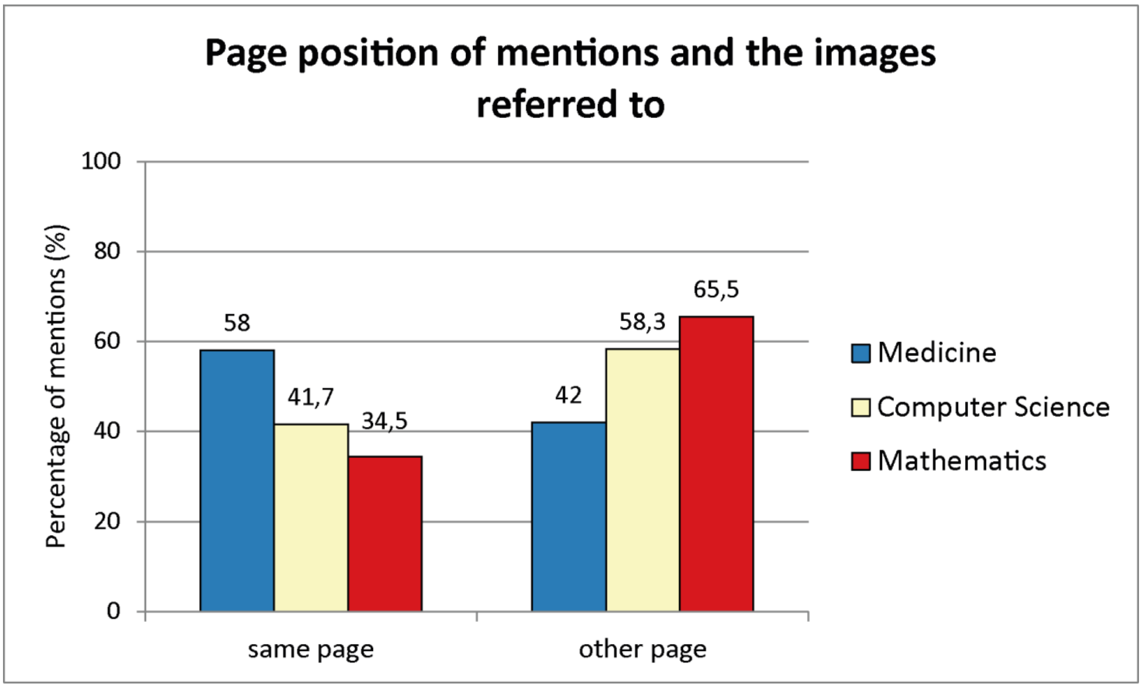


Figure 39 Page position of mentions (as a percentage) and the images referred to.

In all three areas, most mentions appear in the same section as the figure they refer to (Figure 40).

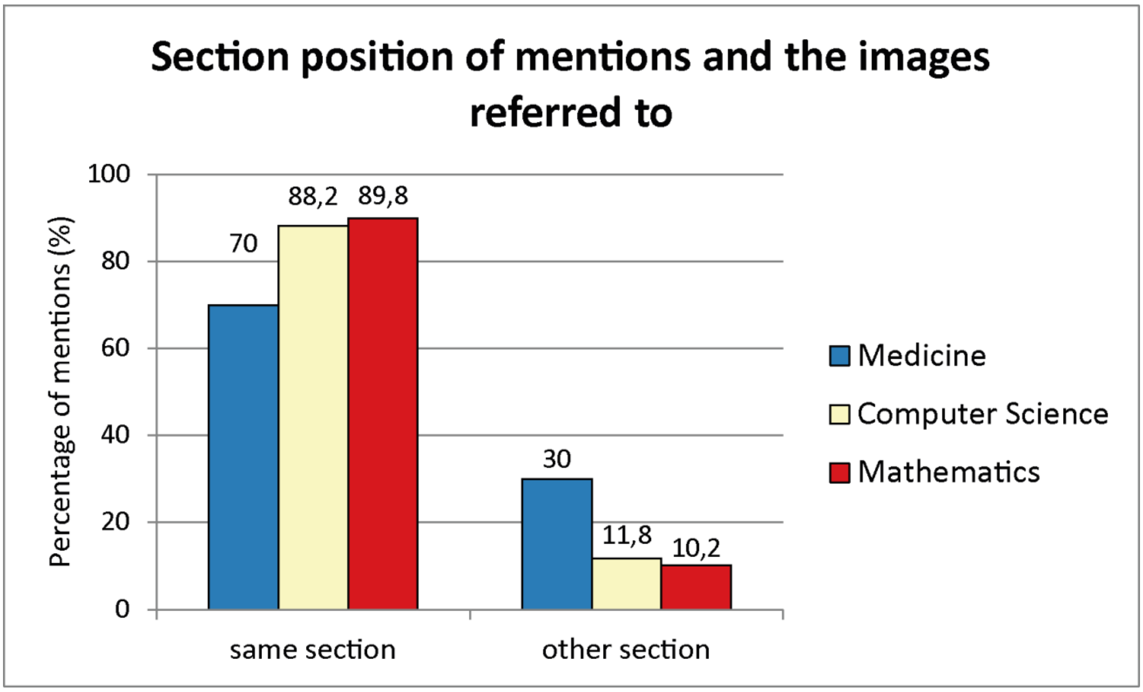


Figure 40 Section position of mentions (as a percentage) and the image referred to.

The distribution of mentions is shown in Figure 41.

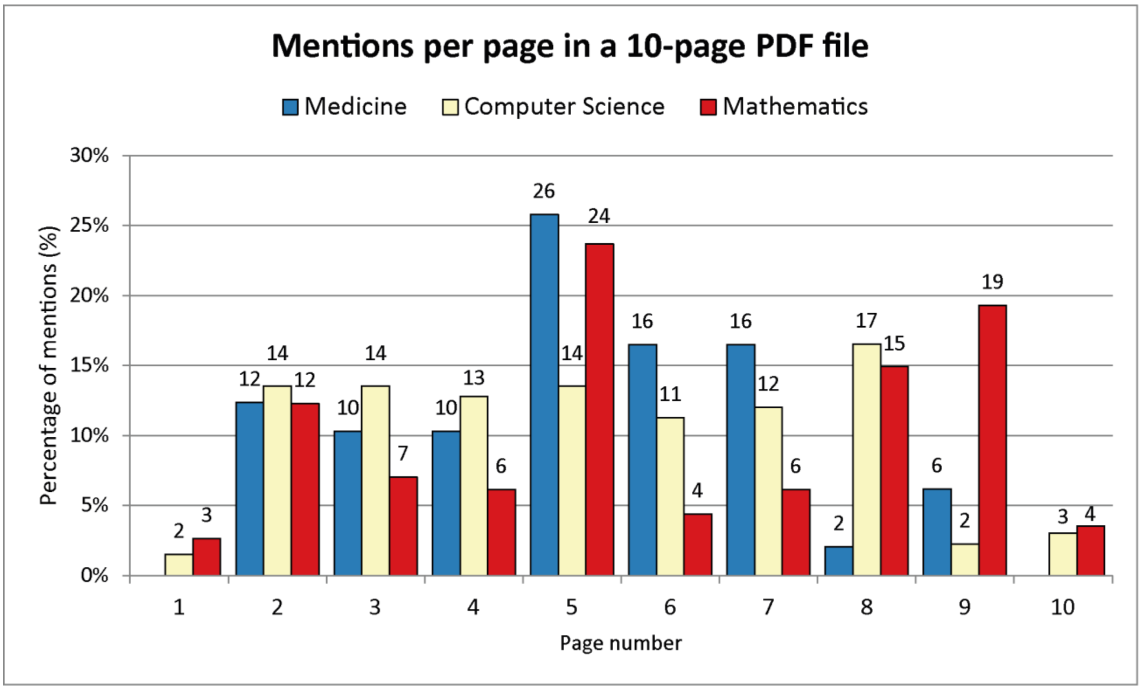


Figure 41 Frequency of mentions (as a percentage) per page in a PDF normalized to 10 pages in the different disciplines.

All three disciplines showed a similar pattern in the position of images and mentions, with both appearing more frequently in the middle of the article. According to the IMRaD (Introduction, Methodology, Results and Discussion) model, the results suggest that the images and their mentions tend to be located more in the "methodology" and "results" sections than in the “Introduction” and “Discussion” sections. The tendency seems to be a little more evident in medical articles, where the IMRaD structure is recommended by the submission guidelines of the International Committee of Medical Journal Editors (ICMJE) (ICMJE, 2013).

5.1.3.2 Function and use of captions and mentions (lexical analysis)

Besides the general analysis of captions and mentions, which measured their application in PDF articles and their length, specific quantitative and qualitative content analyses of the text included in these image-related texts were performed to analyze their lexical content and compare their functions, identify complementarities and overlaps in their use and assess the extent to which they could be potentially appropriate as alternative descriptions of the images.

In this section, the results of the content analysis of captions and mentions are reported. Since the manual qualitative analysis of natural text is very time-consuming, a limited set of journals and articles was selected. The analysis was finally performed on a subsample of nine articles selected in order to reflect the variety of types of images presented in the three disciplines (Table XXXIII).

Table XXXIII List of the selected articles and the types of figures found in each article: statistical graphic (A); schematic illustration (B); diagram without illustrations (C); photograph (D), including medical images; line art (E).

Article ID	Article Title	Image type				
		A	B	C	D	E
MED1	Cancer statistics, 2012	✓				
MED8	A sensing array of radically coupled genetic ‘biopixels’	✓	✓			
MED9	Dnmt3a is essential for hematopoietic stem cell differentiation	✓	✓			
CS1	A Survey of Defense Mechanisms Against Distributed Denial of Service (DDoS) Flooding Attacks	✓	✓	✓	✓	
CS7	Efficient and robust model-to-image alignment using 3D scale-invariant features				✓	✓
CS8	A Vectorization-Optimization-Method-Based Type-2 Fuzzy Neural Network for Noisy Data Classification	✓		✓	✓	
MATH1	An asymmetric convex body with maximal sections of constant volume					✓
MATH4	The Brownian map is the scaling limit of uniform random plane quadrangulations					✓
MATH5	Geometric group theory and 3 manifolds hand in hand: the fulfillment of Thurston's vision					✓

All nine articles reviewed included in their captions and mentions an explicit reference to their related image by using the words *figure* or *fig* (Table XXXIV*).* This consistent reference suggests the possibility of programmatically associating the image to its related text, although the exact scope of the mention is difficult to discern by automatic methods.

Table XXXIV Words used in captions and mentions to identify the images they refer to.

Article ID		**Words that identify images in PDF articles**										
		figure	**fig**	**image**	**diagram**	**scan**	**chart**	**plot**	**graph**	**illustration**	**schema**	**picture**
MED1	caption	✓										
	mention	✓	✓									
MED8	caption	✓		✓	✓							
	mention		✓									
MED9	caption	✓			✓							
	mention		✓									
CS1	caption		✓									
	mention	✓	✓		✓							
CS7	caption		✓	✓		✓						
	mention		✓									
CS8	caption		✓									
	mention		✓				✓	✓				
MATH1	caption	✓							✓			
	mention	✓							✓			
MATH4	caption	✓								✓	✓	
	mention	✓								✓	✓	✓

Article ID		Words that identify images in PDF articles										
		figure	**fig**	**image**	**diagram**	**scan**	**chart**	**plot**	**graph**	**illustration**	**schema**	**picture**
MATH5	caption	✓	✓									✓
	mention	✓										

As described in the “methodology” section, the content analysis explored the presence of words belonging to the “syntactic”, “semantic” and “external” levels of information conveyed by the figures.

The analysis of words belonging to the syntactic description level among the documents is detailed in Table XXXV.

Table XXXV Presence of words belonging to the syntactic level of image description in captions and mentions.

Article ID	Caption/ Mentions	Primitive visual features	Spatial references	Type and technique	The verbs show, display, represent, indicate	Caption as a brief title?
MED1	caption	✓				no
	mention				✓	
MED8	caption	✓	✓	✓	✓	no
	mention		✓			
MED9	caption	✓	✓	✓	✓	no
	mention				✓	
CS1	caption					yes
	mention			✓	✓	
CS7	caption	✓	✓	✓	✓	no
	mention		✓	✓	✓	
CS8	caption					yes
	mention			✓	✓	
MATH1	caption			✓		yes
	mention	✓		✓	✓	
MATH4	caption	✓	✓	✓		no
	mention	✓	✓	✓		
MATH5	caption	✓		✓		yes
	mention				✓	

In biomedicine, the syntactic description of images was mainly accomplished through the text included in the caption. References to visual primitives (e.g. *line*, *circle*, *blue*) and spatial localization of elements in the figure (e.g. *right*, *top*, *middle*) were primarily found in the captions. Also, the prevailing use of the verbs *show*, *display*, *represent* and *indicate* in captions (as compared to mentions) confirms the usefulness of the caption for a text description of the image, including what the image looks like, rather than what the image “is about” (its intended message). Mentions in biomedical articles do not provide information about salient visual features of the figure and are mostly focused on what the image ”is about”. In this discipline, the type of image (for example ”diagram”) is also indicated in captions.

In four out of six articles in the fields of computer science and mathematics, the caption served as a concise title of the image and was limited to a brief sentence. In these cases, the mention(s) were mainly used as longer explanations of the figure covering both syntactic and semantic levels.

Non-external or non-visual information that relates to the image, such as its history and context (date, location, rights, source, etc.) were detected in the captions and mentions. Metadata are the most common way of providing this kind of information. However, as discussed in the previous section, the application of semantic metadata to the figures analyzed is minimal in medical and computer science articles, and is almost absent in figures in mathematical articles.

On another level, the presence and the frequency of technical vocabulary in captions and mentions was detected, in order to assess the complexity of the texts, since the level of complexity of the image-related texts has an impact on the access of people with cognitive impairments. Through the use of the “vocabulary profiler”, a tool that categorizes all the words in the text by their frequency, texts in captions/mentions were segmented into four categories: words from a list of the 1000 most frequent; words from a list of the second 1000 most frequent; words from the *Academic Word List;* and less common words (technical terms, names, etc.). Figure 42 shows the frequency of words in captions and mentions, from the most general words (1-1000 words) to the most technical words (off list words) in the nine articles analyzed.

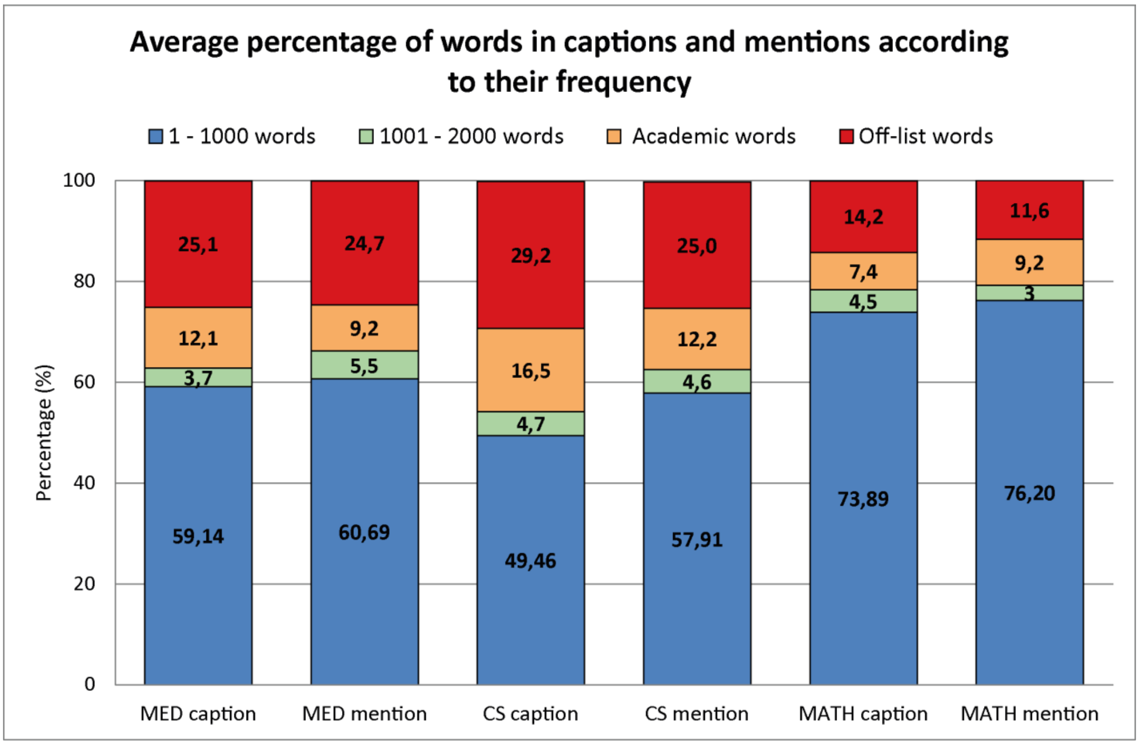


Figure 42 Average percentage of words in captions and mentions according to their frequency in documents from different research areas.

The trend among captions and mentions of the different research fields is similar: most words belong to the 1000 most frequent words (from 76% of mathematics mentions to 49% of computer science captions); words from the list of the second most frequent 1000 range from 3% (mathematics mentions) to 5.50% (medical mentions); words from the *Academic WordList* range from 7% (mathematics captions) to 17% (computer science captions) and subject specific and technical words range from 12% (mathematics mentions) to 29% (computer science captions). In all the cases analyzed the percentage of subject-specific and technical words by far exceeds the common 5% of "off-list words" in generic academic texts (Hyland & Tse, 2007) , and therefore we can state that the vocabulary used in captions and mentions is very semantically dense.

5.1.3.3 Conclusions about image-related texts

As commented in section 5.1.2, the assumption that images do not have text alternatives associated to them has been confirmed. For this reason, the current function and use of image-related texts in STEM articles has been analyzed and discussed in this section, in order to evaluate the use of captions, mentions, text labels and metadata as effective text alternatives for visual content.

Although this analysis has been performed with a limited number of academic articles and a deep analysis requires further investigation, the analysis of the articles led to some conclusions:

- The use of captions and mentions of images in the main text of the article and text labels in figures is widespread in biomedicine, computer science and mathematics.

- Text labels can be automatically recognized and extracted easily only from vector images, but optical character recognition is required for raster images. However, the result of the extraction of this text is a chain of meaningless characters when taken out of its visual context.

- Metadata are sparingly used in figures in biomedical and computer science articles and almost not at all in mathematics. The metadata mostly concern the technical features of the image, while the descriptive metadata are limited to the *author* and *title* fields, and even so, the actual use of the title is uninformative. Current image editing software discourages the use of metadata. Improvements in authoring tools are required for automatically creating meaningful information for images.

- Concerning the function and content of captions and mentions vary in different research areas, some "implicit patterns" seem to appear by type of image according to its function and its intended message. Biomedical figures present illustrations of complex biological processes that in general require detailed textual descriptions. On the other hand, mathematical images are mainly schematic line graphs representing math expressions and abstract concepts which seem to require a reduced number of words for their description.

- Vector graphics present serious accessibility barriers even in the most accessible PDFs. Some tests performed by the author with SVG vector graphics demonstrated that they lose their original structure and the accessible text alternatives when included in PDF documents.

In conclusion, the results suggest that currently image-related texts in STEM academic articles have different characteristics and use and they are appropriate as alternative descriptions in different degrees. The answer to the Q6 question ("Are current image-related texts appropriate as alternative descriptions of the images?") depends on the type of image-related text.

Text labels and other text inside figures cannot be used as an effective mechanism for providing text alternatives to images as they are meaningless out of context. At present, metadata cannot be considered a feasible mechanism to provide a text alternative to images, as their use is hindered by imaging software and authors do not use them.

On the other hand, captions and mentions have promising features for serving as image alternatives: they are semantically dense and often they can be programmatically associated with the images, and their use is widespread. In particular, captions and mentions in biomedical articles have characteristics compatible with long text alternatives: they are used in almost all the articles, they are long enough to be recognized as a detailed description of the figure's content, and they are often located on the same page as the figure they refer to. In order to support these observations, the authors' practices and attitudes about how and why they use captions and mentions in their submitted articles will be explored and the results of this analysis will be detailed in the next section.

5.2 Analysis of authors' practices and attitudes

This section presents the results of the analysis of authors' practices, needs and attitudes concerning the use and function of images and image-related texts in academic articles in medicine and biology, paying particular attention to the awareness and perception of accessibility.

Once confirmed the lack of accessibility of images in scientific journals to people with visual impairments and the almost null use of textual alternatives in academic articles, this analysis explores the current practices of researchers in medicine and biology in the selection and inclusion of images and descriptive text in academic articles, with the aim of introducing improvements in the process. The knowledge of these practices is considered a necessary foundation for assessing practical recommendations for the inclusion of accessible images in academic articles, enabling the support of the authors and the widely adoption into the academic publishing.

In synthesis, the aim of this analysis is to answer the questions Q4 ("Are images in STEM academic articles and in particular in biomedical articles a relevant source of information"?), Q6 (""Are current image-related texts appropriate as alternative descriptions of the images in STEM and in particular in biomedical academic articles?) and Q7 ("How do authors create and submit images in STEM academic articles?").

The section is divided into two subsections: the first reports the results of an online survey answered by 198 researchers and the second the results of the personal interview with 22 researchers. The results are evaluated and discussed according to the research questions.

5.2.1 Online survey

As described in the methodology section, an online transverse survey was created with the aim to study the practices, thoughts and beliefs of the researchers in order to understand why and how academic researchers in medicine and biology research fields use, edit and describe images and their related texts for publishing purposes. The survey was conceived for gaining a "panoramic view" of the issue and as a previous step for better orienting the following interviews.

5.2.1.1 Inclusion of images in scientific publications

The main results of the survey are reported in this section. The complete results of the survey are detailed in Appendix F (in English). A total of 179 participants completed the entire questionnaire and 19 completed the questionnaire partially.

The results of the survey showed that 177 participants (89% of the answers) regularly include images in their scientific publications (Figure 43).

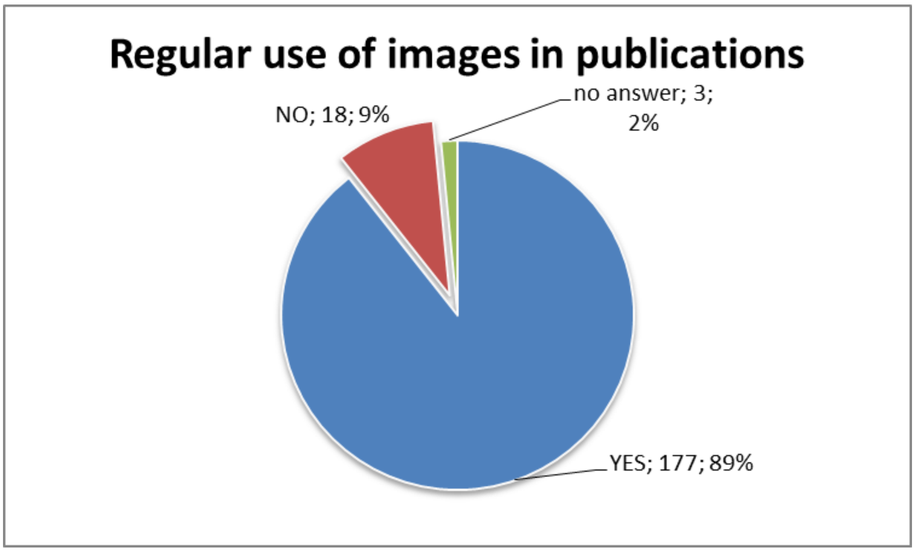


Figure 43 Percentage of participants that regularly use figures in their publications

5.2.1.2 Use of images in publications

Among the 177 participants who regularly include imagines in their publications, 166 (93,7%) indicated to include images in academic articles (Figure 44).

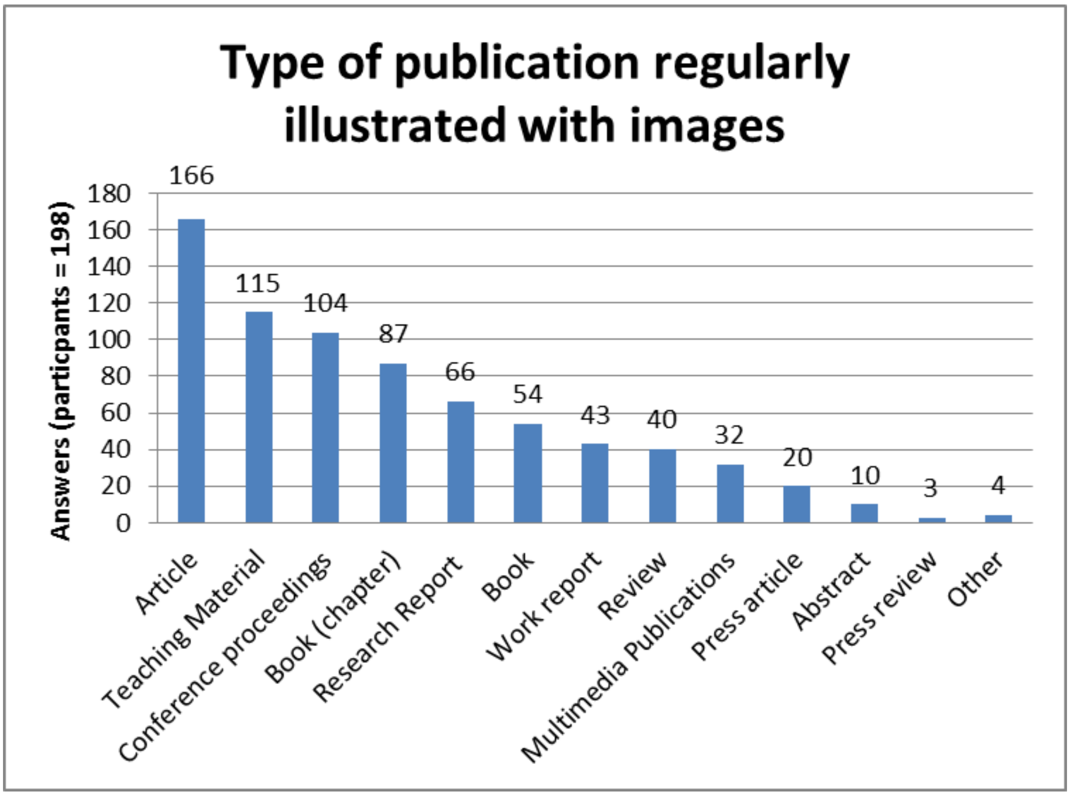


Figure 44 Type of publication regularly illustrated with images by the participants. "Other" includes two posters, one web and one clinical session.

The motivations that justify this inclusion were mainly to illustrate the content of the article (121 participants) and to synthesize information (112 participants) (Figure 45).

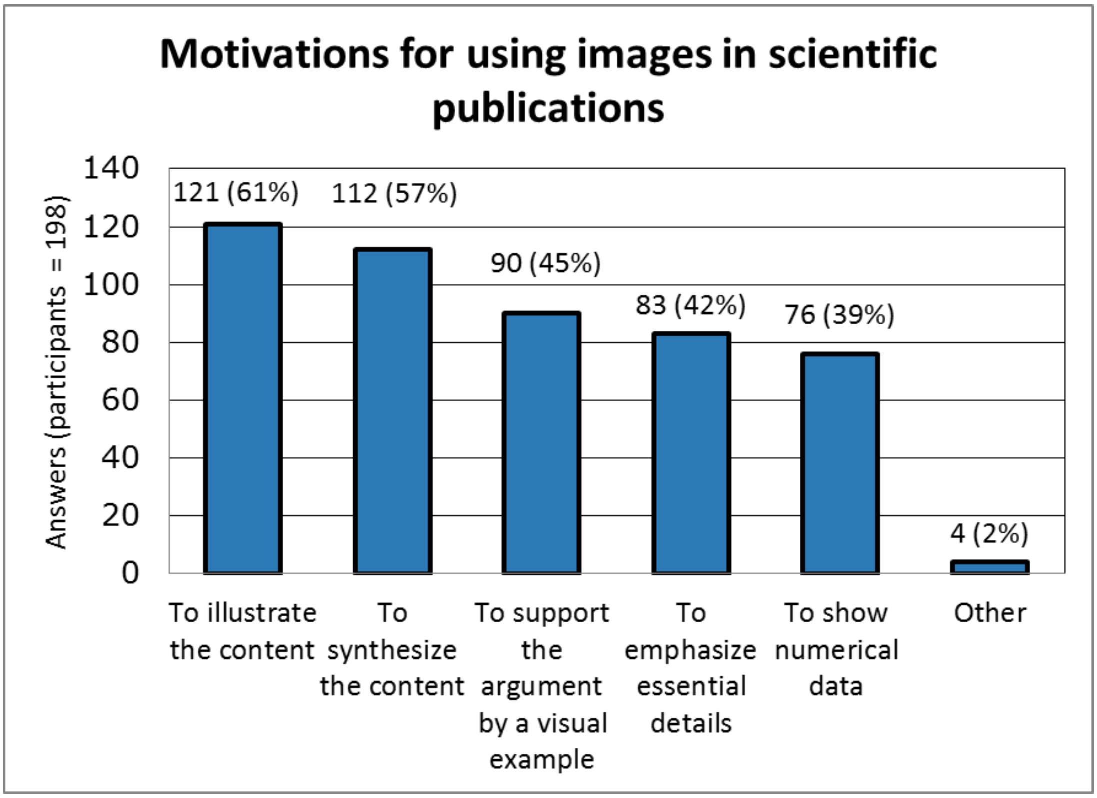


Figure 45 Function of the images in academic articles according to the participants' answers.

The most used image formats were JPEG and TIFF (Figure 46). "Other" category included Adobe Illustrator Format (AI) and other formats not specified.

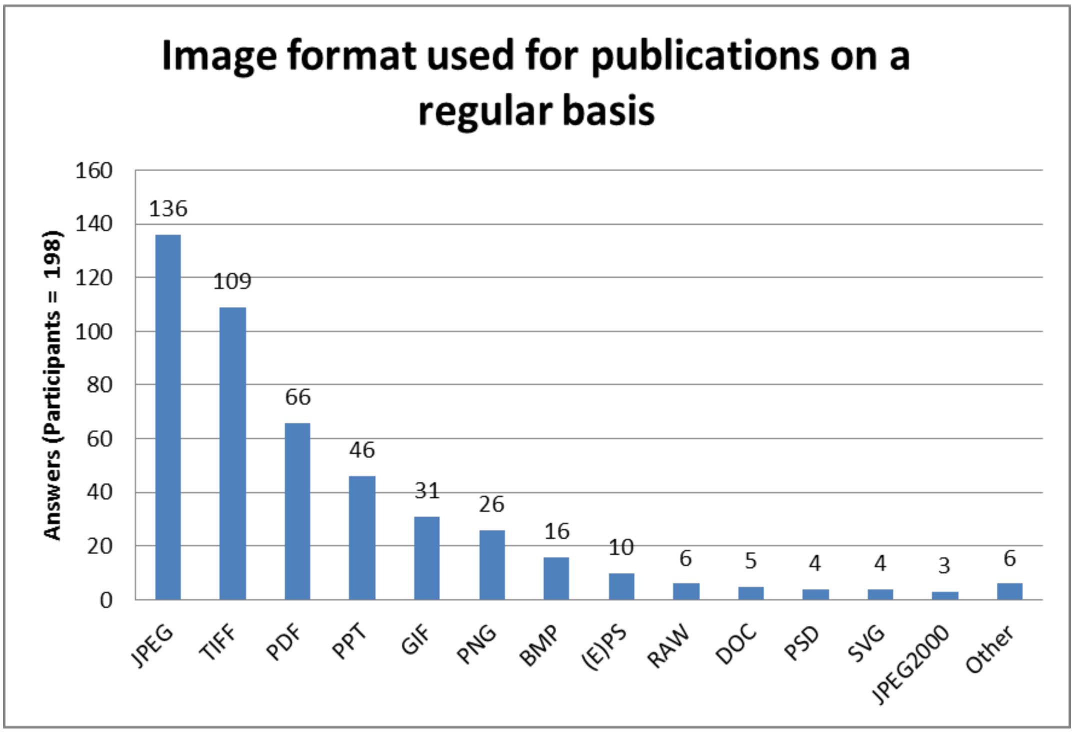


Figure 46 Most used image formats on a regular basis in publications according to the participants.

Participants were asked to specify the reason for using each format. The most cited reasons for choosing a specific format were:

- It is useful for a specific type of image: photo, diagram, vector graphic, drawings, chemical structures, math notations, microscope images, etc. (39 answers out of 150).
- It is required/recommended by publishers for the academic paper submission (35 answers)
- It is commonly adopted, generated by different software and it is portable (23 answers).
- It is useful for a specific action: image manipulation, sharing, reviewing, etc. (19 answers).
- It maintains high quality (such as TIFF) (14 answers) or it has good balance between quality and compression (such as JPEG) (10 answers).
- It is useful in a specific context: in a paper for an academic publication, in a presentation as teaching material, etc. (10 answers).

The most used tools for image editing selected by the participants were Microsoft PowerPoint and Adobe Photoshop (Figure 47).

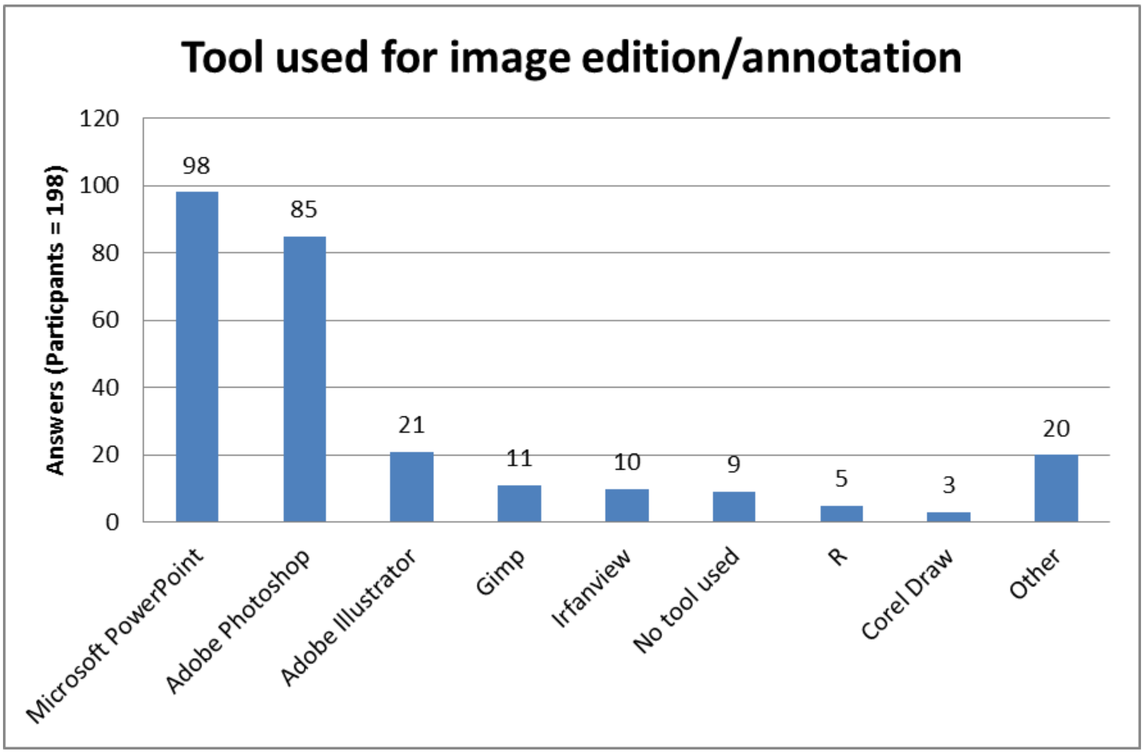


Figure 47 Tools used by participants for image edition.

"Other" tools included Corel PhotoPaint, Inkscape, Jeynote, Excel and Prim, among others.

5.2.1.3 Use of textual descriptions of the images

The visual content is often associated to a descriptive text. 142 participants (72% of the answers) declared to describe images used in academic article by a related textual description (Figure 48).

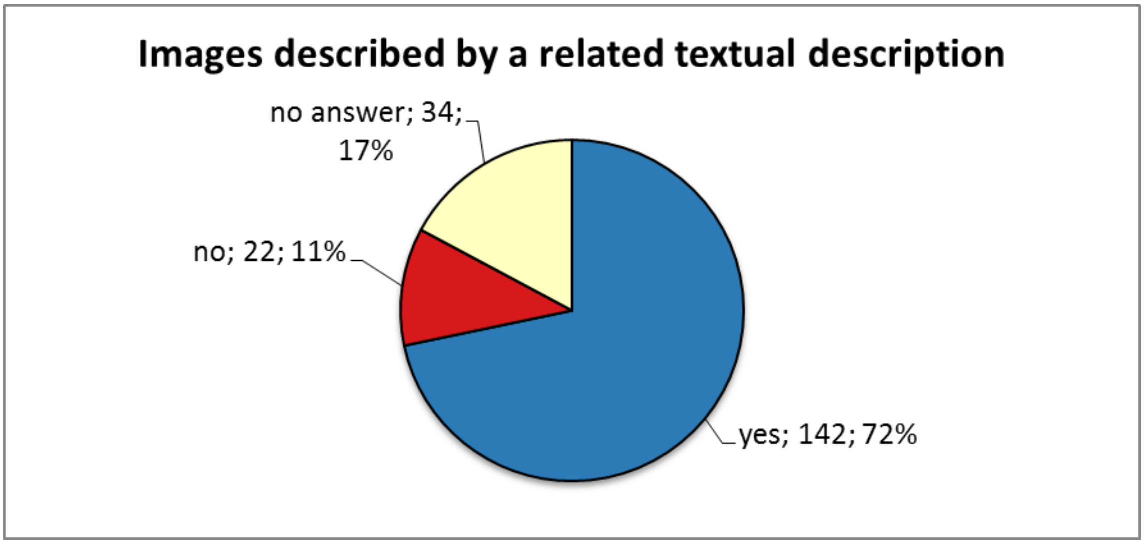


Figure 48 Percentage of participants who associate a description text to the images they use in academic articles.

Among the different types of textual descriptions, 117 participants (59%) declared to regularly associate textual annotations to images in their publications (Figure 49). This information confirmed the important role of textual annotations presented in scientific figures.

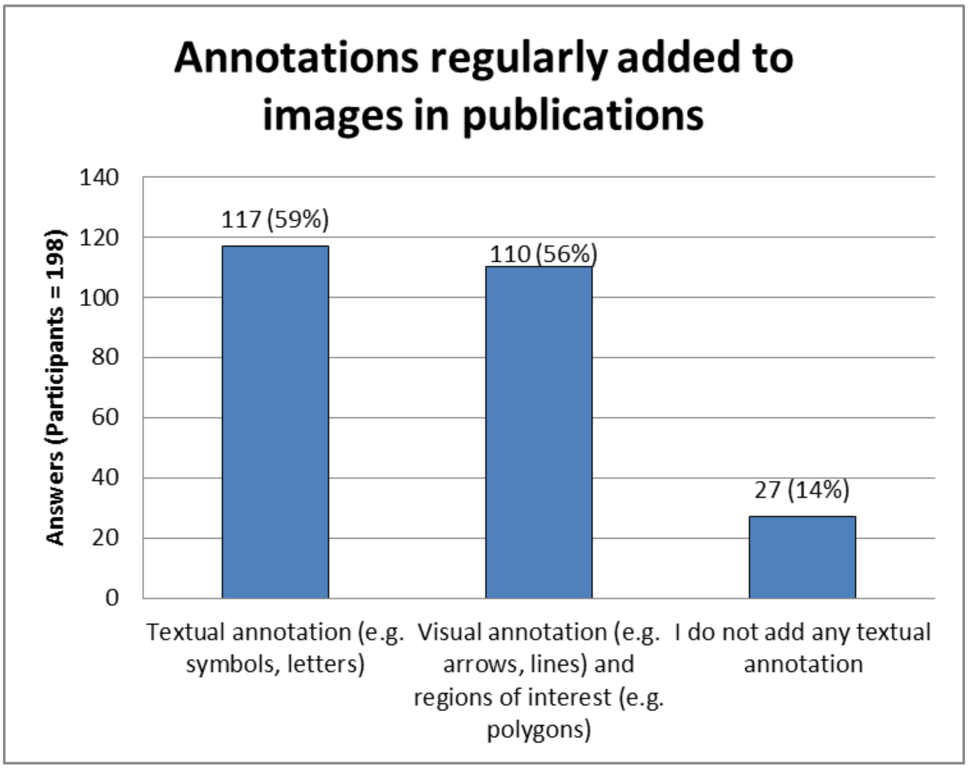


Figure 49 Annotations regularly added by the participants to the images used in publications

In order to evaluate the relevance of the textual descriptions from the perspecitve of the author as reader of academic articles, common practices in reading were explored (Figure 50). The results emphasized that the caption of the image was the element that contains the most relevant information for 171 participants out of 198. On the opposite side, only 23 participants used to extract relevant information from metadata.

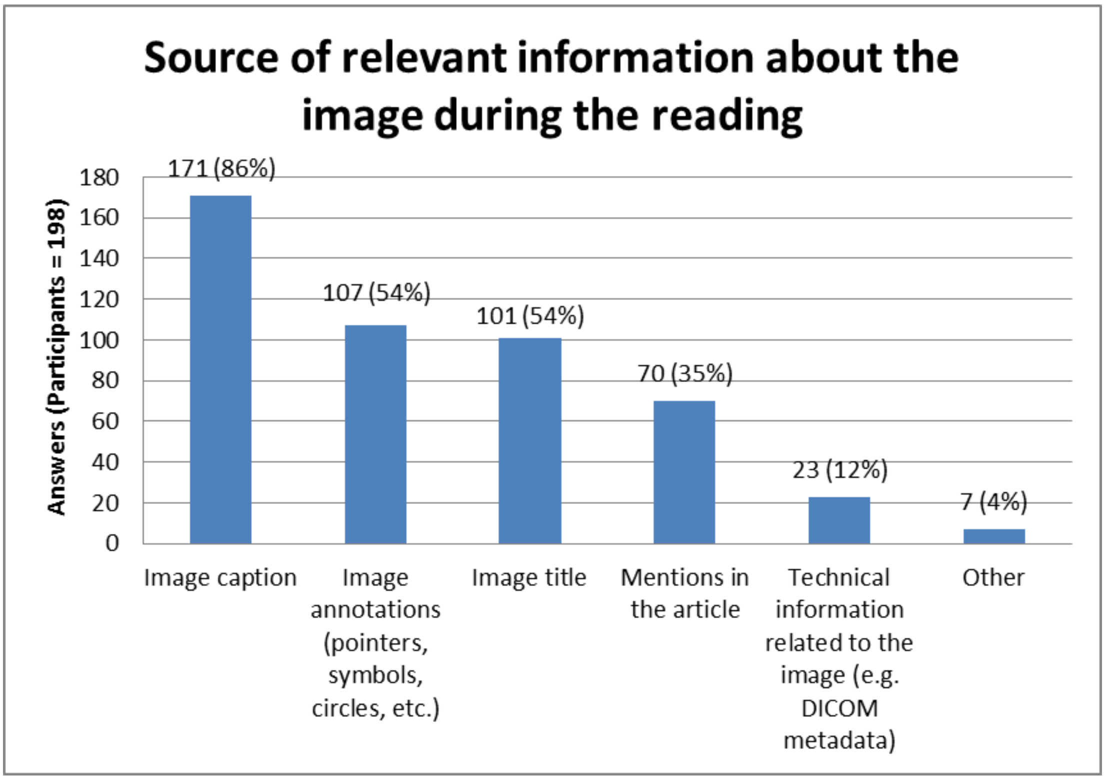


Figure 50 Textual elements related to images in scientific articles considered a source of relevant information of the image content during the reading of academic articles by the participants.

5.2.1.4 Knowledge and awareness of image accessibility

Regarding the image accessibility, the results of the survey showed that only 11 participants (5% of the total answers) were aware of the accessibility policies of scientific journals (Figure 51).

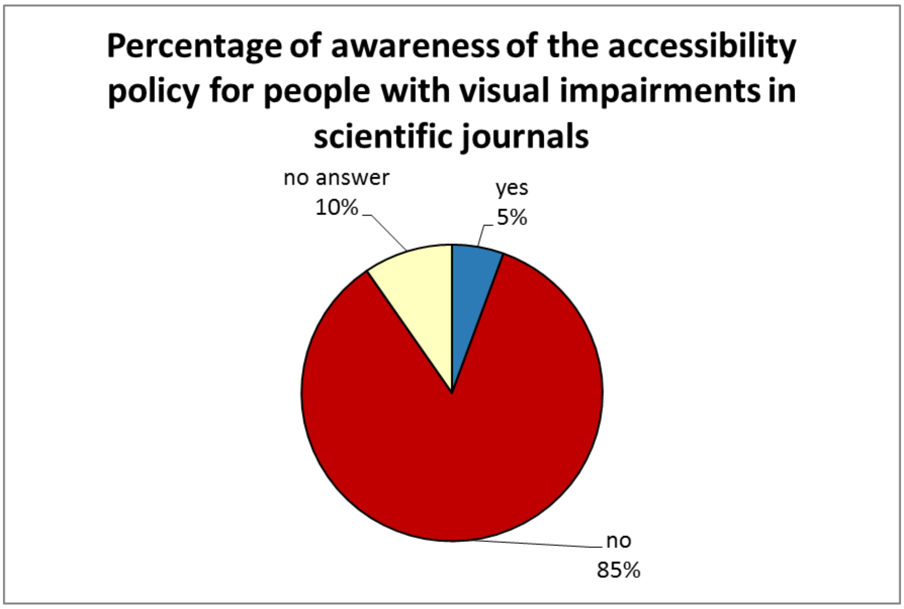


Figure 51 Percentage of participants that declared to be aware of the accessibility policy for people with visual impairments in scientific journals.

Concerning the accessibility of visual content, 15% of the participants (answers) declared to know the meaning of “alternative text” of images (Figure 52).

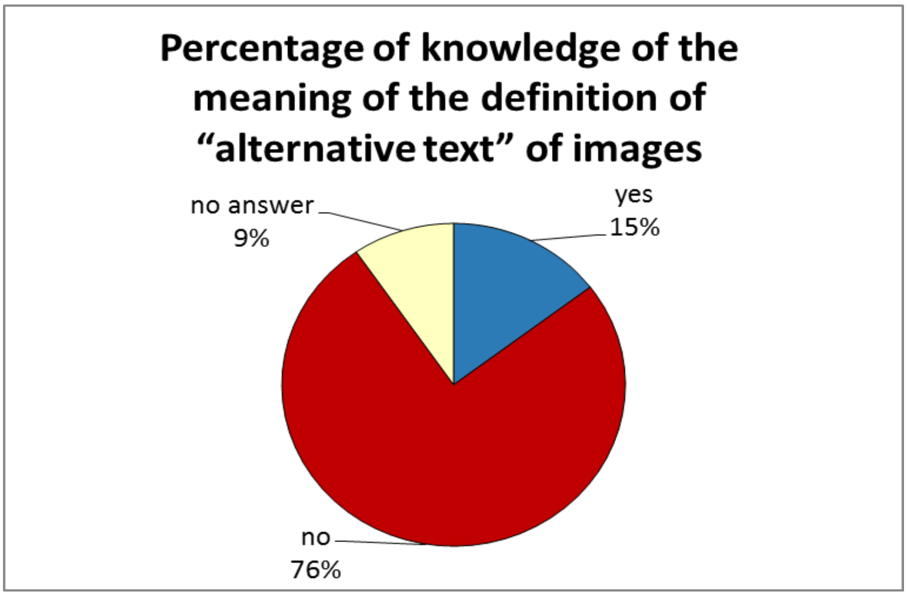


Figure 52 Percentage of participants that declared to know the meaning of “alternative text” for images.

From the point of view of the participants considered as readers of articles with images, they were asked about their experiences in perceiving and understanding figures in academic articles. Unexpectedly, 64 participants (32%) stated to have experienced in some occasion difficulties in seeing the figures published in scientific articles (Figure 53).

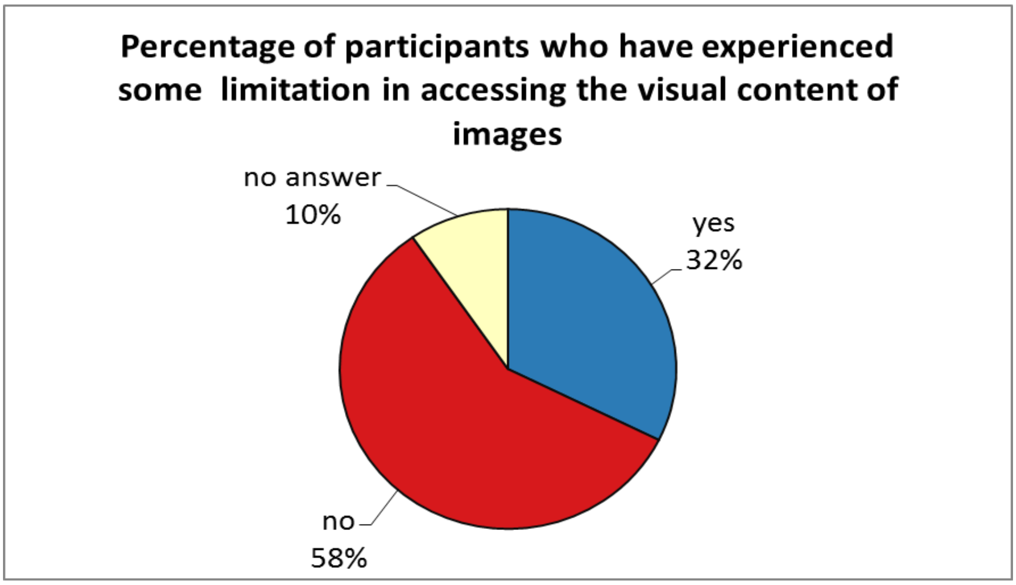


Figure 53 Percentage of participants that declared to have experienced limitations in accessing the visual content of images in scientific articles.

The participants who declared to have experienced limitations in accessing the visual content of scientific articles were asked to specify the type of limitation they experienced (Figure 54).

The most cited limitations were the low quality of the image (11 participants), low resolution (10 participants) and the small font of the text inside the image (9 participants).

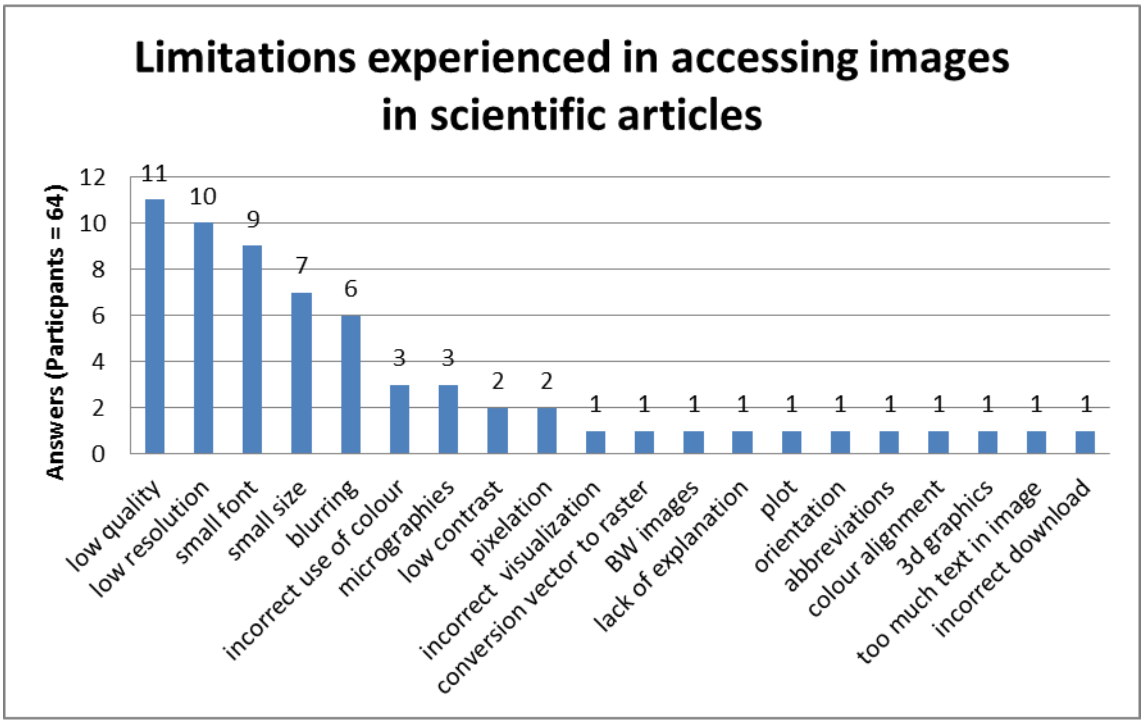


Figure 54 Limitations experienced in accessing the visual content of images in scientific articles according to general categories.

5.2.2 Interviews

As described in the methodology section, the interviews were performed to complete the information collected from the survey and to gain a more qualitative perspective of the opinions and practices of 22 researchers concerning the use and function of the images in academic publications. Besides the answers of the participants, a contextual observation was also performed and notes were taken about the image editing workflow, resources and tools.

5.2.2.1 Results from the contextual observation

The contextual observation was basically useful to find out that all the tasks related to the process of image acquisition/creation and editing were performed by all participants in the computer they used in their workplace. Most participants (19) used the Microsoft Windows operating system. It was observed that the most used image formats were TIFF (10 participants) and JPEG (6 participants), selected according to the publisher's requirements. It was also observed that the most used tools for image editing was Microsoft PowerPoint, selected for its ease of use and for supporting the most common tasks of image editing, such as the creation of annotations (e.g., arrows and letters), cropping, framing and anonymization of personal information in clinical images. A smaller number of participants cited to use Adobe Photoshop.

Concerning the observations of the type of images used, clinical images were often captured by different medical imaging devices: radiography, X-ray Computed Tomography (CT), Magnetic Resonance (MR) and Positron Emission Tomography (PET). In the case of some researchers located in the labs of the Hospital Clinic and the IDIBAPS, the process of accessing and sharing these images stored in a Picture Archiving and Communication System (PACS) was observed. These images usually were accessed in DICOM standard format. However, according to the participants, DICOM metadata were lost when they exported the images to other tools for image editing.

Some participants were observed during their use of repositories for archiving clinical images, for example the "Historia Clínica compartida de Cataluña" (H3C), a system for sharing radiologic images among physicians public hospitals in Catalunya territory and the "Radiological Archive and Image Management" (RAIM), a system for sharing and managing digital images. Biologists working with genomic images cited the use of a public repository called "UCSC genome browser", which allows the generation, sharing and exportation of images in different resolution and dimensions.

During the observations it was detected that the researchers perceived the process of image editing as a set of tasks that required the knowledge of a new language – the visual language – on the use of which they were not trained.

5.2.2.2 Functions of the images in academic publications

The motivations that drive researchers to include images in scientific articles were explored in the interview. According to the clarifications provided by the researchers, the categories initially defined (to illustrate, to synthetize, to support, to emphasize, to show numerical data and other functions) were regrouped and redefined as:

1. Illustrative function (cited by nine participants): The figure illustrates or confirms the article content. For example, P13 said: "*Todas las inferencias que hacemos en nuestro trabajo si no las documentamos con las imágenes no las podemos publicar*" ("We cannot publish all the inferences in our work, unless we do it with images").
2. Summary/synthesis function (five participants): the figure synthesizes easily, clearly or succinctly the information provided textually or in a table. For example P3 cited the renowned aphorism "*una imagen vale más que mil palabras*" ("One picture is worth a thousand words"), adding the comment "c*on texto* […] *se te irían tres párrafos y no acabarías de explicar todos los matices*" ("In a text […], you should use three paragraphs and still not all the details will be explained").
3. Emphasis function (five participants): The figure emphasizes essential details. P6 mentioned the fact that "*especificando el tipo de patología que estamos estudiando, ver cómo se comporta en la imagen es algo que es muy necesario*" ("It is very necessary to specify the type of pathology and see its behavior in the image").

4. Example function (four participants): The figure supports the argument by a visual example. P9 clarifies the concepts explaining that "*Una fotografía va ser muchísimo más ilustrativa que cualquier descripción*" ("A photo is much more illustrative than any other description").
5. Data display function (four participants: The figure shows numerical data). For example, P8 informally said "*Las gráficas son una forma de que los resultados que son un "rollo"* (sic.) *de explicar sean más visuales. Una forma que el resultado entre por la vista*" ("Graphics are a way to show results that are boring (sic.) to be reported. They introduce the results with a visual appeal".

5.2.2.3 Image selection criteria

The interview also highlighted the criteria applied by the researchers for selecting the image to include in their publications. Five selection criteria were identified:

1. Relevance (cited by 14 participants): the "adherence" of the image to the content of the article, the ability to visually reflect the content of the article. P1 explained the concept: "[el criterio para] *seleccionar imágenes* […] *es el que ilumina, que enseña, es más pedagógico respeto al objetivo del trabajo y los hallazgos específicos que has encontrado*" ("[The criterion] for the selection of images […] is to enlighten, to show, it is more pedagogical concerning the objective of the work and its specific findings"). P10 said: "*buscas imágenes para un artículo que den respuesta a las preguntas que planteas incluso en el propio títul*o" ("I look for images that answer the questions outlined in the title of the article itself").
2. Quality (six participants): the ability to present visual information with appropriate resolution, contrast and sharpness. P9 assessed that "*muchas veces la calidad fotográfica condiciona mucho que incluyamos una u otra imagen. De hecho en los artículos de revista te piden una resolución mínima*" ("Often the photographic quality has a lot of influence on the inclusion of one image rather than another image. In fact a minimum resolution is required in journal articles").
3. Impact (four participants): the ability to be interesting and exciting and to attract the attention. P7 clearly stated "*en un congreso donde se escogen las mejores comunicaciones se citan proyectando la imagen que el autor ha escogido como imagen resumen de ese trabajo. Es puro marketing*" ("In a conference where the best presentations are selected, they are cited presenting the image that the author has selected as an abstract of this work. It is pure marketing").
4. Comparison (2 participants): the ability of a multi-panel image to make easier the comparison among the different figures that compose the entire image. For example, P16 emphasized that "*asociamos* [a una imagen] *otro tipo de imagen, que sea los más similar posible o en el corte o en la representación con la otra* […] *para que se vea la semejanza"* ("We associate [to an image] another type of image, which should be very similar to the original one in the cut or in the representation, in order to show the similarity").

5. Density (1 participant): the ability to provide a high density of information. P2 stated that "*A menudo el número de imágenes está limitado, se eligen las figuras que crees que van a aportar más información"* ("Often the number of images is limited, the figures you consider that provide more information are chosen").

5.2.2.4 Common image editing tasks

During the interview, the participants cited a set of tasks they usually performed in the process of editing of images for articles submitted in academic journals:

- Modification of the image resolution (11 participants).
- Editing visual annotations, such as arrows, polygons and circles (seven participants).
- Conversion from one image format to another (six participants).
- Adjustments of brightness and contrast (five participants).
- Editing textual annotations, such as letters and symbols (four participants).
- Renaming the original axis description in the image generated from data analysis software (three participants).
- Changing the font size of text inside the images (three participants).
- Removing personally identifiable information from the figures (anonymization) (two participants).
- Grouping two or more figures into one image (two participants).
- Image cropping (one participant).
- Adding frames and borders to the image (one participant).
- Reorienting the image (one participant).

The main motivation cited for performing one of these tasks is to follow the requirements of the submission guidelines of the publishers.

5.2.2.5 Functions and characteristics of the caption

Concerning the main characteristics of the caption, the viewpoint of the participants can be synthesized by the comment of P1: "*El pie de texto tiene que ser muy simple, muy claro*". According to the explanations provided by the researchers, the main functions of the caption are complementary to the visual information provided by the figure and they are the following:

1. To inform: The caption provides information in order to allow readers to understand the figure. For example, according to P2 "leyendo el pie de figura uno tiene que entender perfectamente qué aparece en la figura" ("Reading the caption, you must perfectly understand what the figure is showing").
2. To provide instructions: the caption provides instructions on how to read the image properly. For example, P21 stated that "*en el pie de figura tienes que explicar que es aquello, que funciones representan y* [desarrollar] *las abreviaturas*" ("In the caption you

have to explain what they are, which functions they represent and give the full form of the abbreviations”).

3. To describe: the caption describes the elements of the figures (“what the image depicts”). For example P15 said “*Como son muchas veces multipanel, haces una descripción de lo que es cada una de ellas.* […] *si es una comparación de dos patologías o sujetos en estudio longitudinal*” (“Since they are often multipanel, you make a description of what each image is […] if it is a comparison between two pathologies of subjects in a longitudinal study”).

In order to specify more in detail the information included in the caption, the 22 interviewees were asked to mention those elements of the figures that they usually include in the caption.

Table XXXVI shows the elements suggested by the participants, ordered by category and field of application. In the case of medical images, there was no agreement about the inclusion of clinical information in the caption.

Table XXXVI Elements of the figures cited by the participants which are usually included in the caption.

Figure element described in the caption	Cited by the participants	Apply to
Identify/explain labels	X	all types of figure
Abbreviations	X	
Color code	X	
Modifications (focus, cropping, etc.)	X	
Source	X	
Description of variables	X	statistical graphics
Unit of measurement of the axis	X	
Statistical analysis details	X	
Areas of interest	X	other figures in Medicine and Biology (photos, etc.)
Magnification rate (scale)	X	
Modality	X	
Clinical information	*	

Figure element described in the caption	Cited by the participants	Apply to
Cut type	x	other figures only in Medicine (photos, etc.)
Reconstruction type	X	

Each element is explained as following:

1. Identify/explain labels: identification and explanation of labels, including symbols, letters, arrows and numbers used to identify in the figure.
2. Abbreviations: explanation of abbreviations and acronyms.
3. Color code: explanation of the meaning of the colors used in the image.
4. Modifications: explanation of the modification, selective digital adjustments and enhancements (cropping, brightness, color, sharpness, etc.) applied to the figure.
5. Source: specification of the external source of the image and appropriate credits.
6. Description of variables: description of the variables used in the graphic.
7. Units of measurement: description of the unit(s) of measurement used in the graphic.
8. Statistical analysis details: description of the results of the statistical analysis showed in the figure, including the type of analysis, information about the sample, standard deviation, statistical significance, etc.
9. Areas of interest: description of the areas of interest marked in the image.
10. Magnification rate: specification of the scale of the image (for example a tissue or cells in a micrograph).
11. Modality: specification of the acquisition modality of the image (e.g. radiography, CT, microscopy, etc.).
12. Clinical information: description of the clinical background of the image.
13. Cut type: specification of the cut type (for example, for a CT the cut can be axial, coronal or sagittal).
14. Reconstruction type: specification of the reconstruction type of the image (for example, for a CT the reconstruction can be a rendering 3D or a mip).

5.2.2.6 Functions and characteristics of the mentions

As a general trend, the answers of the participants indicated that the mentions to the images from the main text are useful elements to provide a “basic understanding” of the image content without looking at the image.

Caption and mentions in the text were usually perceived as related and complementary and specific differences in the content of the two elements were not identified. The authors declared to choose of including information in one text or the other depending on the submission guidelines stated by the publishers. For example, when the guidelines recommend a schematic caption, the author will use the mention in the text for explaining the image in more detail. On the contrary, if the guidelines strictly limit the number of words of the article, the author could take advantage of the space for words offered by the caption and include relevant information related to the figure in it.

This example emphasizes the priority (identified along the majority of the interviews) that the researchers assign to the guidelines established by the publishers when they need to submit an image to an academic journal. 16 participants out of 22 explicitly cited some cases where the submission requirements shaped their process of creation, edition and submission of images and textual descriptions. One participant (P14) even said "*La normativa muy detallada* [se] *complica si tú tienes un artículo magnifico y lo quieres enviar a la mejor revista; si te dicen que no, tienes que volver a hacer formalmente* […] *todo el trabajo porque la otra revista tiene una normativa distinta*" ("When you have a fantastic article and you want to send it to the best journal, very detailed rules make the things complicated; if they said no, you have to do all the formal work again because the other journal has other rules of submission"). The researcher expressed clearly that the "loyal" adherence to the submission guidelines is an essential requirement for avoiding a review with major changes or even the rejection of the manuscript.

According to five participants (P12, P14, P5, P6, P11), the localization of the mentions of the images in the article followed some patterns. Considering the IMRaD model for the structure of the academic articles, the mentions of the images would follow the distribution showed in Table XXXVII.

Table XXXVII Examples of the information provided by the participants and related to images in the main text of the articles according to the IMRaD canonical structure.

Main section of the paper	Related information provided in mentions
Introduction	N/A
Method and Materials	Acquisition modality (P12) Processing techniques (P12) Statistical analysis method (P5) Sample (e.g. number of patients) (P5)
Results	Interpretation of statistical graphics (P6) Data used in statistical graphics, most of times in tabular form (P11)
Discussion	Interpretation of statistical graphics (P6)

Main section of the paper	Related information provided in mentions
	Source data visualized in statistical graphics, most of times by using a table (P11)

5.2.2.7 Functions and characteristics of metadata

Some researchers cited the use of the DICOM format in the original image acquired from the PACS system. However, they confirmed that the DICOM metadata were lost when the images are converted to other formats.

5.2.2.8 Accessibility issues

Participants showed a general lack of awareness of the meaning of “alternative text”. One participant (P13) for example asked “*Qué quieres decir con texto alternativo? Que no sea dentro del texto directamente? Las revistas muy buenas cada vez más piden una figura que sea un "highlight" del artículo”* (“What do you mean by alternative text? It is outside the text, isn’t it? Very good journals are increasingly requiring a figure that is an “highlight" of the article”).

One of the participants (P20) had color blindness (with protanopy, the most common type of color blindness) and affirmed to have experienced in several occasions difficulties in reading graphics with red and green. He suggested the use of blue-yellow combinations as an accessible alternative.

Two participants without special needs (P20 and P22) suggested some practices for avoiding barriers of access, such as to check the true print dimensions of the image, to use color visible to readers with color blindness or to use patterns as an alternative to the color and to apply an appropriate thickness to the lines of the graphics.

5.2.2.9 Difficulties and suggestions in the image submission process

A new subject of interest related to the publishers’ guidelines on images emerged from the analysis of the answers and comments of the participants. They cited some difficulties experienced during the process of editing, submission and publication of images in academic articles:

- Image rejected for low resolution (seven participants).
  Some researchers complained of some image editing tools: “Incluyendo la imagen en el Word pierde de calidad, hay pérdida de resolución” (“When you include a figure in Word, it loses quality, there is a loss of resolution”) (P2) or ”Cuando haces la composición de la imagen en PowerPoint cuesta mucho recuperar las imágenes en condiciones óptimas”

("When you makes a composition of the image in PowerPoint, it is difficult to get back the image in its best conditions")[3] (P14).

Other researchers complained of the slowness of the uploading process of high-resolution images in the submission systems: "Esta imagen se ve bien con una resolución de 300 dpi y las revistas normalmente nos piden 600 dpi, cosa que es imposible porque el aplicativo web no lo coge. Esto es un horror" ("This image can be seen well with a resolution of 300 dpi and usually the journal requires 600 dpi. This is impossible since the application does not accept it. This is horrible") (P18).

- Additional cost for color images (seven participants): According to the type of image, the loss of color was unacceptable for the researchers who choose to pay the additional cost for publishing the images in color or the Open Access option without the limitation of color.
- Very restrictive limitations in the number of figures to include in the article (three participants): "Las figuras dentro de los artículos son muy limitadas […] a veces las fotografías las pones en material suplementario" ("The number of figures inside articles is very limited […] sometimes you opt to include photographs in the supplementary material") (P14).
- Minimum contrast required: (2 participants): They mentioned that it was difficult to comply with this requirement with very complex graphics with a lot of colors and textures.

The participants also made suggestions for improving the publishing process: "*Estaría bien que cuando mandas las imágenes, aunque sean con resultados importantes, no se pidiera el formato definitivo antes de la revisión", "Se podrían minimizar los costes de las figuras, por ejemplo* [publicando] *todas las figuras en la misma página"* ("It could be useful not to require the final format of the submitted image before the revision, inclusive when you the image shows important results,", "Costs for figures could be lowered, for example [publishing] all the figures in the same page") (P18).

5.3 Conclusions on the analysis of authors

5.3.1 Are images in biomedical academic articles a relevant source of information?

The first result is the main role that images have currently in articles in medicine and biology. The importance of the visual information in academic articles, already identified in the analysis of the publishing practices in section 5.1.1, was confirmed by the results of the survey and the interviews. The results have showed the extended use of images in academic articles among researchers and their essential role in providing information that complements the text of the articles. This finding answers affirmatively the question Q4 ("Are images in STEM academic

[3] MS Office allows setting-up the resolution of the images produced.

articles and in particular in biomedical articles a relevant source of information?”) and reinforces the justification of this thesis.

5.3.2 Are current image-related texts appropriate as alternative descriptions of the images in biomedical academic articles?

The second main result is the confirmation of the importance of image-related texts in articles in medicine and biology. In particular, the results of the survey and the interviews confirm that the inclusion of captions and mentions in the main text is a common practice accepted by authors and integrated in the publishing process of images in academic articles.

The functions of image-related texts identified in this study correspond to the uses cited in the style guides and publisher’s guidelines reviewed in section 3.5. Concerning the caption, the functions of “to inform” and “to describe” perceived by the researchers could be aligned with the equivalent functions cited by the publishers (see section 3.5.1).

The information - cited by the participants - that should be included in the caption could be also aligned to the elements recommended by the bibliographic references reviewed in section 3.5.1.

Table XXXVIII shows the elements cited by the researchers aligned with the elements identified in the bibliography. Information related to the content of captions was reviewed in the following references: renowned manuals of style - Chicago Manual of Style (15th ed.), Publication Manual of the American Psychological Association (6th ed.) and American Chemical Society Style guide (3rd ed.) - ; submission guidelines of renowned international publishers publishing journals in biomedicine - Springer, Wiley-Blackwell, New England Journal of Medicine, Nature y Science –; guidelines provided by the ICMJE (International Committee of Medical Journal Editors, 2014).

Table XXXVIII Elements of the figures described in the caption cited by the participants and their correspondence in the bibliography.

Figure element described in the caption	Cited by the participants	According to manuals and submission guidelines	Apply to
Identify/explain labels	X	X	
Abbreviations	X	X	all types of figure
Color code	X		

Figure element described in the caption	Cited by the participants	According to manuals and submission guidelines	Apply to
Modifications (focus, cropping, etc.)	X	X	
Source	X	X	
Description of variables	X	X	statistical graphics
Unit of measurement of the axis	X	X	
Statistical analysis details	X	X	
Areas of interest	X	X	other figures in Medicine and Biology (photos, etc.)
Magnification rate (scale)	X	X	
Modality	X	X	
Clinical information	*	X	
Cut type	X	X	other figures only in Medicine (photos, etc.)
Reconstruction type	X	X	
Method of staining		X	other figures in Biology (photos, etc.)
lanes in electrophoretic gels and blots		X	

In general, the results of the survey and the interviews indicates that the textual information associated to the figures by captions and mentions are used to complement the visual information and, therefore, it does not replace completely the visual information conveyed by the image neither its function when the image is not available. For this reason, currently they can be only partially used as effective text alternative for images. Therefore, the answer to the question Q6 ("Are current image-related texts appropriate as alternative descriptions of the images in STEM and in particular in biomedical academic articles?") is partially negative.

However, due to their widespread use, length, position in the page and explicit reference to their related images (usually by the words *figure* or *fig),* captions and mentions may potentially be used as long text alternatives. In particular, the text included in captions mainly accomplishes the syntactic description of images (what the image looks like) and confirms its usefulness as a text description of the image. The identification of the functions and the content of captions and mentions according to the authors' comments and the reviewed literature provides a solid basis for developing a proposal of possible adjustments in these image-related texts, in order to make them useful as effective text alternatives.

5.3.3 How do authors create and submit images in STEM academic articles?

Concerning the process of creation, editing and submission of images, the results of the survey and the interviews emphasize some issues that can be useful for the definition of a strategy for the inclusion of image accessibility in the publishing process.

The first finding concerns the image formats and tools currently used by the researchers. The results of the survey and the interviews confirmed the widespread use of the JPEG and TIFF formats for the image submission and, in a less extent, the use of other image and document formats that can contain images (PDF, PPT, GIF, PNG, etc.). The SVG format is almost completely ignored. Image formats are chosen according to those recommended in the publishers' guidelines for the submission of articles.

Another finding is that the most used tools for editing images cited by the participants were Microsoft PowerPoint and Adobe Photoshop. Although PowerPoint is not an image editing software, its widespread availability among researchers, its ease of use and the support of the most common image editing tasks justify its extensive use.

Although these tools offer some accessibility features - such as the possibility to associate text alternatives to images, to check the accessibility of the images or to simulate some visual disability - and they also provide documentation about their accessibility support, they present some limitations. For example:

- PowerPoint does not allow adding empty/null alternative text to decorative images. Besides the observations of WebAIM (WebAIM, 2014), this limitation has been also detected directly by the author of the thesis by testing the function for including text alternatives in figures included in PPTX files. The test was performed with PowerPoint 2010.
- Photoshop presents inconsistencies in the manner in which it embeds image descriptions, since they are handled differently depending on the image format (NCAM, 2012). For

example, the description metadata added to PNG files using Photoshop is not available when the image is opened in Adobe Illustrator. This limitation has been checked personally by the author of the thesis.

Another observation is that the participants usually perform a limited set of common image editing tasks (concerning resolution, textual and visual annotations, format conversion, contrast adjustments, etc.). The main reason for performing these tasks is to adjust the image to the submission guidelines requirements.

A third finding is the confirmation of a general lack of awareness of accessibility needs and guideline by the participants. A consequence of this finding is that they do not take into account the requirements for designing accessible images.

These findings corresponds to the three general conditions that should be met for making digital publishing systems and content accessible (Brewer, 2013):

- Availability of technical accessibility standards.
- Availability of tools facilitating the production of accessible content.
- Widespread awareness of the requirements and resources for accessibility.

These conditions, as pointed out before, are currently not fulfilled in the academic publishing of STEM articles in general and in biomedical articles in particular.

Nonetheless, seeing opportunities where others see failures, a comprehensive proposal for the inclusion of accessibility criteria in the authoring workflow of images can be developed by introducing accessibility recommendations integrated into the current practices, thanks to the acquired knowledge of the researcher's common practices and motivations.

The information collected about the author's practices and motivations can be used as the basis for developing specific interventions according the three main components of the Behaviour Change Wheel model (as described in section 1.4):

- The capabilities of the authors, by increasing their knowledge and understanding on how to make images accessible.
- The opportunity of making accessible images, by providing appropriate tools and standards, reviewing the submission guidelines of the publishers or creating a favorable social context.
- The motivation of the authors, by stimulating and persuading authors on reasoning about the ethical and legal consequences of making accessible images.

Considering that researchers highlighted the importance of publishers' guidelines in their decisions when preparing images, publishers can play an essential role in the introduction of the new practices.

The next section of the thesis will be dedicated to describe the proposal according to the results of our analysis and the general scope of application at the three levels just outlined.

Chapter 6

Proposal

6 Proposal

In the previous chapter, the analysis performed with images and articles in STEM academic journals (with a particular attention to biomedical fields) have been described. The analysis of current behaviors and opinions of researchers in medicine and biology commonly submitting images in their academic articles has been also presented. Starting from the results of the analysis and according to the aims of the thesis, the proposal of inclusion of accessibility guidelines for images into the current submission guidelines is presented below. The proposal has the objective to increase the opportunities of an actual adoption of accessibility principles concerning the images within the biomedical academic publishing by minimizing the changes in the current publishing workflow and the extra effort required to the main stakeholders of the process (authors and publishers) in fulfilling these requirements.

The proposal aims to answer the last question of the thesis: How to foster the adoption of accessibility principles concerning images within the STEM publishing and in particular in biomedical publishing? (Q8).

The model and the framework for establishing the proposal is described below.

6.1 Overview of the model for the proposal

As commented in chapter 1, the Behaviour Change Wheel (BCW) is a method for the definition of behavior change interventions supported by evidences in the targeted behavior.

The method was the source of inspiration for building a proposal that starts from the analysis of the target behavior in context and suggests possible interventions to improve the inclusion of accessible images in STEM academic publishing. The analysis described in chapter 5 has identified key behaviors and attitudes of authors during the process of creating, editing and submitting images in STEM articles, in particular in biomedical articles. The results of the analysis are valuable evidences about the image submission process in biomedical academic articles and can be used as a starting point for identifying issues which may need to be addressed.

The BCW model offers a general framework for identifying key behaviors and proposing changes according to the three components of capability, opportunity and motivation (COM-B model of behavior). In order the enable these conditions, different possible activities designed to change behaviors (intervention functions) are selected and specific actions for each intervention are proposed. The proposal of this thesis is structured according to this process and designed according to the three components of intervention of the BCW model:

1. The capabilities of the authors: interventions that should increase their knowledge, understanding and skills on how to make images accessible. The selected intervention functions for this component are:

- Education: increase knowledge or understanding of the needs and barriers of readers with disabilities in accessing images.
- Training**:** facilitate the development of skills on how to make images accessible.

2. The opportunity of making accessible images: interventions that should prompt authors to create accessible images by providing them appropriate tools and standards or creating a favorable social context. The selected intervention functions for this component are:

   - Enablement: provide means or reduce barriers to increase the opportunity of making images accessible.
   - Restriction: introduce measures that prohibit or limit the submission of images which do not conform to accessibility requirements.
   - Environmental restructuring: create a favorable physical and social context by generating corporate awareness of accessibility criteria within the publisher image editing process.

3. The motivation of the authors: interventions that should stimulate and persuade authors on reasoning about the ethical and legal consequences of making accessible images. The selected intervention functions for this component are:

   - Persuasion: inform about social and legal requirements and consequences in order to stimulate the creation of accessible images.
   - Incentivisation: create prospect of reward in the case of submitting accessible images.

The intervention functions and actions proposed in an accessible image workflow are detailed in Table XXXIX, together with the actors responsible of the interventions.

One of the key behaviors identified in the analysis of the authors was the strict follow-up of the publisher guidelines. 73% of the participants of the interview (16 out of 22) explicitly cited some cases where the submission requirements drove their process of editing and submission of images and textual descriptions. Therefore, the proposal assigns the lead role for executing the changes proposed (see Table XXXIX**¡Error! No se encuentra el origen de la referencia.**) to the publishers.

Table XXXIX Intervention functions, actions and actors involved in an accessible image workflow.

Component	**Intervention function**	**Action**	**Actors**
Capability	Education	Inform authors about readers with disabilities. Offer simulation tools in order to perceive images as those readers do.	Publisher
	Training	Provide guidelines, examples and evaluation tools.	Publisher

Component	Intervention function	Action	Actors
Opportunity	Enablement	Improve image editing tools in order to create accessible images.	Image editing software developers
	Restriction	Introduce validations within the submission process.	Publisher
	Environmental restructuring	Include accessibility criteria within the publisher image editing process.	Publisher
Motivation	Persuasion	Inform about disabilities prevalence. Inform about legal requirements. Inform about findability benefits.	Publisher
	Incentivisation	Offer rewards to accessible images (time to be published, discounts in color publishing).	Publisher

The proposed interventions and the actors involved in the behavior change are assigned to specific steps in the "image authoring chain" defined in the chapter 1. Figure 55 shows the steps of "image lifecyle", defined in section 1.4 and reviewed according to the results of the analysis. In this publishing chain, the interventions are focused on the Creation, Editing, Submission, Review and Production steps (which are particularly affected by accessibility issues. The Editing step is also affected by image editing software.

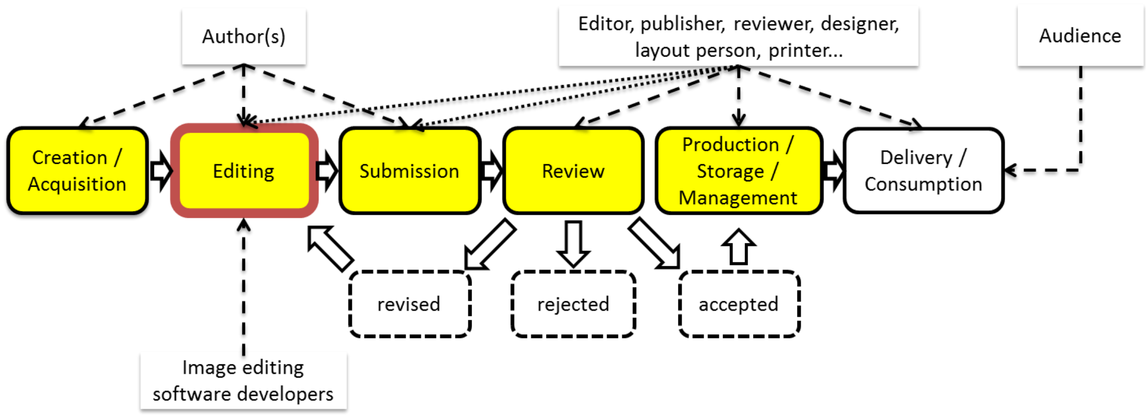


Figure 55 Image authoring workflow including accessibility issues and actors.

Authors are the actors responsible for the Creation/Acquisition, Editing and Submission steps, while publishers are the main actors in the Review, Production/Storage/Management steps. Publishers are also indirectly involved in the Editing and Submission step, since they define the submission guidelines followed by authors and influence their process of image editing.

The Editing step is particularly relevant within the process, since in this step the author prepares the image for its submission in the biomedical article. The actions proposed are intended to help authors in submitting accessible images to the journal, ideally without further interventions in the editorial process. In this approach, the provision of images suitably formed for all kind of users mainly relies on the authors. Furthermore, the authors are considered the main responsible for creating and assigning text alternatives to images (Kelly et al., 2009).

6.2 Perspective of the proposal – image-related texts as text alternatives

As discussed in section 5.1.3, alternative text descriptions are considered essential for conveying the content for science images. In biomedicine, the lack of a meaningful text alternative to images is one of the main barriers that limit the access to the content.

The analysis of captions and mentions in biomedical articles described in section 5.1.3 has emphasized that these texts are widely used and they have characteristics compatible with long text alternatives, suggesting their suitability as potential alternatives to images:

- they have a widespread use: 94,9% of figures analyzed have a caption, 82,1% have a mention;
- they are long enough to be recognized as a detailed description of the figure's content: captions have an average of 802 characters in captions; mentions have an average of 504 characters.
- they are often located on the same page as the figure they refer to: 100% of captions are located next to the images; 58% mentions are located in the same page (70% in the same section).
- 100% of the captions and mentions of the figures analyzed include an explicit reference to their related image by using the words *figure* or *fig,* suggesting the possibility to programmatically associate images to their relative texts.
- The text included in captions mainly accomplishes the syntactic description of images (what the image looks like) and confirms the usefulness of the caption for a text description of the image.

However, the interviews with biomedical researchers showed that existing captions and mentions fall a bit short as text alternatives, since they play just a complementary role of the visual information and they cannot serve as a replacement for them.

In this thesis the use of captions or mentions as a mechanism for providing long text alternatives - a rich, expressive description necessary to explain the details of the graphic – is proposed, under the condition that the authors would improve their content and include a richer description of the visual content. The advantage of the proposal is that these mechanisms are already integrated in the current biomedical image publishing workflow, converting the creation of textual alternatives into a part of the content-writing practices of authors.

In order to help publishers to implement this solution, as a first step, two tools that could be included in the image submission workflow are envisaged: the image text alternatives decision tree, which helps academic authors on deciding where and how to include the alternative text and to select the appropriate image-related, and a checklist for orienting authors in the creation of suitable, effective and meaningful captions as alternative texts for images.

These tools are designed to overcome the current limitations of the solutions offered by accessibility guidelines for creating text alternatives (detailed in section 4.8):

- they lack of specific suggestions on how to textually describe technical images with high information such as those commonly found in medicine and biology fields do not exist. For example, there is no specific guideline on how to describe an X-Ray image.
- They mainly address a generic audience and common publications. They are not related to the current practices of academic authors on the process of creating and selecting images for image submission to academic journals.

The proposed solution adopts a realistic perspective on the implementation of accessibility in image submission and publishing workflow: it is tailored to the specific context (STEM academic publishing) and it is linked to the current practices in image submission and publishing. Therefore, it aims at minimizing the extra work of adapting images to special needs requirements and limiting the changes in the current image publishing workflow required for the implementation of accessibility requirements, increasing the opportunities of an actual application of accessibility principles.

In the next sections (6.3 – 6.5), the general proposal is detailed as following: for each component (capability, opportunity and motivation) specific interventions (education, training, etc.) are detailed. The actors responsible for performing the intervention and the step of the publishing workflow where the intervention has an impact are defined. Each recommendation is based on the results of the analysis of journals, articles and researchers described in chapter 5 and it is supported by the literature review. Related results and bibliography are cited for each recommendation.

6.3 Capability

6.3.1 Education

This intervention has the aim to increase the awareness, knowledge or understanding of the authors about the needs and barriers experienced by readers with disabilities in accessing images.

It is proposed as an answer to the observation that currently most authors are not aware of the need of making image accessible. The results of the survey showed that only 5% of the participants were aware of the accessibility policies of scientific journals and only 15% of the participants declared to know the meaning of “alternative text” of images (section 5.2.1.4).

These results were confirmed by the observations in the interviews, where researchers in biomedicine did not know the meaning of “text alternative” (section 5.2.1.4). Furthermore, four out of 12 biomedical journals analyzed did not offer an accessibility statement and only six journals provided an explicit reference to the use of alternative content for images in their accessibility statements. Image-related accessibility issues were mentioned explicitly in the image submission guidelines for authors only by 25% of the journals analyzed and all of the mentions where isolated references to specific reader barriers and the legibility of the image (see section 5.1.1.1).

Publishers are the main responsible for executing the interventions suggested by the recommendation n.1 and recommendation n.2 described below. These interventions concern the submission guidelines and affect the behaviors of the authors at the editing step.

6.3.1.1 Recommendation n.1

Publishers should inform authors about the accessibility of the resources provided and the barriers that prevent readers with special needs from accessing the content of images and their needs. The information should be referred from the submission guidelines for authors and also from the accessibility statement advertised in the company documentation and on the website, in order to give them more relevance.

Related best practices

An example of accessibility statement is provided in the website of the Nature Publishing Group (2015). A model of accessibility statement (for institutional repositories) has been proposed by Kelly (2011).

6.3.1.2 Recommendation n.2

Publishers should offer simulation tools integrated in the submission system in order to show authors how their images are perceived by readers with visual impairments or when they are printed in black and white.

For example, the submitted images can be presented with color filters simulating color-blindness vision or in black and white.

6.3.2 Training

This intervention is parallel to the intervention concerning education and it aims at improving the skills of authors on how to make accessible images.

This intervention is suggested by the observations in the analysis that publishers do not currently offers common guidelines, examples and evaluation tools on how to make accessible images to the authors. In particular, image-related accessibility issues were mentioned explicitly in the image submission guidelines only by one fourth of the journals analyzed and all of the mentions where isolated references to specific reader barriers and the legibility of the image (see section 5.1.1.1). The analysis of the biomedical articles showed a general incorrect use and even the absence of the *alt* attribute and the absence of the *longdesc* attribute in the HTML versions of the articles and the absence of *alternate text* and the *actual text* attributes in PDF versions (section 5.1.2.3.). Concerning the use of formats supporting accessibility features such as SVG, only 2% of the participants of the survey declared to use SVG (section 5.1.2.4) and only two biomedical journals out of 12 accepted SVG as a submission format for figures (section 5.1.1.2).

The interviews also showed that the main motivation in editing procedures is to follow the requirements of the submission guidelines of the publishers.

Finally, the analysis of the current guidelines of biomedical journals showed a lack of common guidelines for manuscript (and image) submission among publishers: only three journals out of 12 (25%) follow guidelines for manuscript submission provided by international organizations of editors in biomedical field.

Publishers are the main responsible for executing the interventions suggested by the recommendation n.3 and n. 4 described below. These interventions affect the behaviors of the authors at the editing step.

6.3.2.1 Recommendation n.3

Publishers should provide guidelines, examples and evaluation tools on how to make accessible images. This information could be included in the submission guidelines or provided by specific platforms offering interactive courses, video tutorials, webinar and other material on how to make images accessible on their websites.

In particular, publishers should offer training on:

- the creation of effective texts alternatives. In order to facilitate this task and according to the results of the analysis, the use of supporting material is recommended.
- the techniques for including alternative texts in common document formats used for submitting articles, such as PDF and HTML;
- the creation of images using accessible formats (such as SVG);
- the use of common image editing tools currently used by the authors (see next section) for creating accessible images;
- the editing of technical characteristics of the images that have consequences on the perception and the understanding of figures by people with disabilities, in order to guarantee a minimum degree of accessibility in the submitted figures (see the section 6.4 “Opportunity”).

Related best practices

“The Springer author academy” (Springer Publishing, 2015) is a good example of platform providing learning materials to help authors to write, to submit and to publish manuscripts. A course on how to make images in academic manuscripts accessible could be provided according to this model.

6.3.2.2 Recommendation n.4

Publishers should adopt internationally recognized guidelines for manuscript submission, such as those provided by the ICMJE (International Committee of Medical Journal Editors, 2014).

This intervention has two consequences:

- It could foster the standardization of submission guidelines – including guidelines for images - among journals in specific research fields. This standardization would make the process of submitting articles in different journals easier for authors who submit to another journal after having being rejected, or who submit to many journals and have to learn each different requirement. In particular, the tasks of re-editing figures for conforming them to different submission guidelines could be spared when the standardization is applied.
- It could be a starting point for the standardization of requirements for submitting accessible images, under the condition that organizations as ICMJE would recommend the submission of accessible images in their guidelines. Due to this condition, the author of the thesis considers that this recommendation is beyond an adjustment of the current publishing workflow that can be reasonably performed in a short term. It is suggested for a mid-long term intervention.

6.4 Opportunity

6.4.1 Enablement

This intervention aims to increase means (or reducing barriers) of making images accessible by suggesting improvements in the image editing tools support to accessibility features.

It is based on the observation that most participants declared to use Microsoft PowerPoint (98 participants out of 198) and Adobe Photoshop (85 participants out of 198) as image editing tools in the process of editing and publishing biomedical academic images (see section 5.2.2.1). This result was confirmed by the participants of the interviews.

Suggestions are oriented to integrate accessibility features in the same affordable mainstream editing tools used by mainstream users and, at the same time, to allow authors to create descriptions that can survive along the publishing workflow of academic articles.

Image editing software developers are responsible for executing the intervention suggested by the recommendation 5 described below. This intervention affects the behaviors of the authors at the editing step.

6.4.1.1 Recommendation n.5

Common editing tools used by the authors should enable authors to easily export images in high resolution, convert images to SVG format and support the exportation of the images with the *alt* attribute with "*null*" value. This recommendation is addressed to image editing software developers.

As commented in the section 5.3, common editing tools used by the authors, such as Powerpoint and Photoshop, currently support some accessibility features, such as the inclusion of text alternatives to the images they generate. However, they present some limitations that will be detailed below, together with the recommendations on how to fix them.

Microsoft PowerPoint

Although PowerPoint allows the inclusion of a text alternative to a figure, currently it has the limitation of not allowing the inclusion of a "null" alt text, as commented in section 5.3.

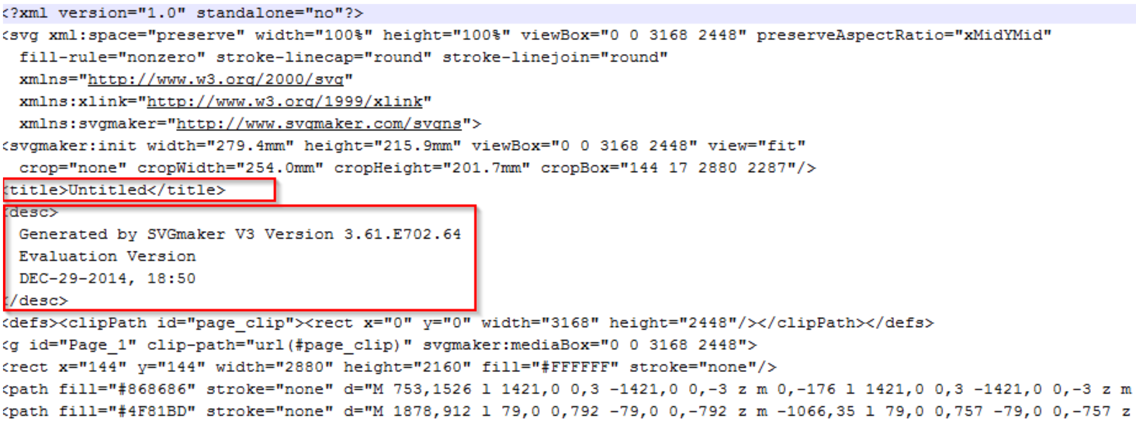
The Lecshare tool (currently discontinued), a tool for checking and improving accessibility in PowerPoint slides with some additional transformation tools, allows assigning a "null" value to the *alt* attribute. The author of the thesis checked the XML code of the PPTX file generated from the tool after adding a "null" *alt* attribute to a figure in a PPTX file. He realized that the software adds a user defined tag for specifying the "null" value in the *alt* attribute text of the <img> tag, as showed in the code below:

```
<p:tagLst
xmlns:a=http://schemas.openxmlformats.org/drawingml/2006/main
xmlns:r="http://schemas.openxmlformats.org/officeDocument/2006/r
elationships"
xmlns:p="http://schemas.openxmlformats.org/presentationml/2006/m
ain">
<p:tag name="ALT_NULL" val="1"/><p:tag name="LONGDESC_NULL"
val="1"/>
</p:tagLst>
```

The example shows that the implementation of the option for adding a "null" alt attribute to the figures edited in PowerPoint is technically feasible. Therefore, Microsoft developers are recommended to implement this option.

Another accessibility-related issue of PowerPoint is the possibility of exporting raster figures from PowerPoint to other image formats (such as JPEG, PNG, BMP or TIFF) in a specified resolution. The default resolution of raster figures exported from PowerPoint is 96 dots per inch (dpi). The export resolution can be changed modifying a registry subkey in Windows. The maximum resolution allowed is 307 dpi (Microsoft, 2015). This process can be applied in PowerPoint 2003/2007/2010/2013 versions. Microsoft developers are recommended to make the option for setting the export resolution more intuitive to non-expert users, and to increase the possible resolutions to 600 dpi at least. Publishers should also offer a section in their guidelines explaining to the authors how to manage this option.

Microsoft PowerPoint does not offer the option to export a PPT(X) file directly to a SVG file. To perform this conversion, several solutions were taken into account by the author, such as online tools, scripts for Powerpoint, Java applications and Java libraries. However, according to the scope and aims of the proposal, it was considered that the most appropriate solution should be easy to install and use, should not require programming knowledge and should work directly with PowerPoint. A possible solution following these conditions is SVGmaker (Software Mechanics Pty Ltd, 2015), a printer driver for Windows systems, which allows generating SVG files by printing from standard Windows programs, including PowerPoint. However, it has the limitations of working only with Windows and of generating SVG files that lose the original title and <desc> tag (Figure 56).



```
<?xml version="1.0" standalone="no"?>
<svg xml:space="preserve" width="100%" height="100%" viewBox="0 0 3168 2448" preserveAspectRatio="xMidYMid"
  fill-rule="nonzero" stroke-linecap="round" stroke-linejoin="round"
  xmlns="http://www.w3.org/2000/svg"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:svgmaker="http://www.svgmaker.com/svgns">
<svgmaker:init width="279.4mm" height="215.9mm" viewBox="0 0 3168 2448" view="fit"
  crop="none" cropWidth="254.0mm" cropHeight="201.7mm" cropBox="144 17 2880 2287"/>
<title>Untitled</title>
<desc>
  Generated by SVGmaker V3 Version 3.61.E702.64
  Evaluation Version
  DEC-29-2014, 18:50
</desc>
<defs><clipPath id="page_clip"><rect x="0" y="0" width="3168" height="2448"/></clipPath></defs>
<g id="Page_1" clip-path="url(#page_clip)" svgmaker:mediaBox="0 0 3168 2448">
<rect x="144" y="144" width="2880" height="2160" fill="#FFFFFF" stroke="none"/>
<path fill="#868686" stroke="none" d="M 753,1526 l 1421,0 0,3 -1421,0 0,-3 z m 0,-176 l 1421,0 0,3 -1421,0 0,-3 z m
<path fill="#4F81BD" stroke="none" d="M 1878,912 l 79,0 0,792 -79,0 0,-792 z m -1066,35 l 79,0 0,757 -79,0 0,-757 z
```

Figure 56 Screenshot of the code of the SVG file created by SVGmaker from a PPT.

Another example for a possible solution could be an extension similar to the one offered by the OpenOffice Impress for exporting files directly to SVG format.

PowerPoint developers are recommended to include a default easy-to-use option for exporting Powerpoint files directly to SVG.

Adobe Photoshop

Adobe Photoshop CC (the latest version of the software) allows the inclusion of descriptive text for the content of figures by metadata. However, as commented in section 5.3, Adobe image tools present some inconsistencies in maintaining the metadata of some specific image formats among different tools. For example, the description metadata added to PNG files using Photoshop is not available when the image is opened in Adobe Illustrator. The recommendation addressed to Adobe is to solve the inconsistencies in handling metadata among different image editing software tools.

6.4.2 Restriction

This intervention aims to restrict the submission of not accessible images by introducing mechanisms of control and validation of the technical features of the images within the submission process.

The restrictions proposed are based on:

- the recommendations concerning the technical aspects of the images on which the perception and understanding of the figures depend;
- the most frequent tasks and barriers in editing images in academic articles observed in the analysis of the interviews.

The technical aspects concerning image accessibility are defined according to technical guidelines offered by WCAG 2.0, which are addressed to overcome barriers for specific user groups. The author of the thesis has related them to specific features of the image, which should be considered in the techniques for applying the recommendation correctly. These guidelines have been detailed in section 4.1.4 and are repeated here in Table XL.

Table XL Image-related accessibility issues according to WCAG 2.0.

WCAG 2.0 Principle and guideline(s)	WCAG 2.0 Success Criterion	Level	Barrier	User group	Related to
1. Perceivable 1.1. Text Alternatives	1.1.1 Text alternatives	A	Rich images lacking equivalent text	blind, cognitive, contextual, search engine	text alternatives
1. Perceivable 1.4. Distinguishable	1.4.1 Use of color	A	Color perception is necessary to understand information	blind, low vision, contextual	use of color
1. Perceivable 1.4. Distinguishable	1.4.3, 1.4.6 Contrast	AA, AAA	Insufficient visual contrast	low vision, contextual	color contrast, resolution, dimensions
1. Perceivable 1.4. Distinguishable	1.4.4 Resize text	AA	Text cannot be resized	low vision, contextual	font size and font family of text in images
3. Understandable 3.1. Readable	3.1.3 Unusual words	AAA	Complex text	cognitive	content of text in images and in caption/mentions
3. Understandable 3.1. Readable	3.1.4 Abbreviations	AAA	Acronyms and abbreviations without expansions	cognitive	content of text in images and in caption/mentions

According to the results of the interviews, common tasks of image editing were:

- Modification of the image resolution (eleven participants).
- Editing visual annotations, such as arrows, polygons and circles (seven participants).
- Conversion from one image format to another (six participants).
- Adjustments of brightness and contrast (five participants).
- Editing textual annotations, such as letters and symbols (four participants).
- Renaming the original axis description in the image generated from adata analysis software (tree participants).
- Changing the font size of text inside the images (three participants).
- Removing personally identifiable information from the figures (two participants).
- Grouping two or more figures into one image (two participants).

The results of the interview also showed that common difficulties during the image submission process were:

- Image rejected for low resolution (seven participants).
- Additional cost for color images (seven participants).
- Very restrictive limitations in the number of figures to include in the article (three participants).
- Minimum contrast required (two participants).

Table XLI shows how the majority of image editing tasks and difficulties in the image submission process concern the same features as the image-related accessibility issues, especially those related to the perception of images.

Table XLI Image-related accessibility issues aligned to the image editing tasks and the difficulties in the submission process cited by the participants of the interview.

Barrier	Related to	Image editing tasks cited by the participants	Difficulties in image submission process
Rich images lacking equivalent text	Text alternatives		
Color perception is necessary to understand information	use of color		Cost for color images
Insufficient visual contrast	visual contrast, resolution, dimensions	Image resolution. Visual annotations. Brightness and contrast.	Low resolution. Minimum contrast.

Barrier	Related to	Image editing tasks cited by the participants	Difficulties in image submission process
Text cannot be resized	font size and font family of text in images	Textual annotations (axis renaming included). Font size of text in images.	
Complex text	content of text in images and in caption/mentions		
Acronyms and abbreviations without expansions	content of text in images and in caption/mentions		
		Format conversion	
		Grouping figures	
		Identifiable information removal	
			Limit of figures

The proposal of restrictions addresses the issues derived from the interview, excluding "identifiable information removal" and "limit of figures" issues, which do not have a clear impact on accessibility.

A general recommendation concerning the restrictions is detailed below, together with recommendations for each specific issue provided by accessibility institutions and researchers, with contributions of the author of the thesis.

The publishers are responsible for executing the intervention suggested by the recommendation 6 described below. This intervention affects the behaviors of the authors at the editing and submission steps.

6.4.2.1 Recommendation n.6

Publishers should specify in the submission guidelines specific restrictions concerning the technical characteristics of the images and introduce a validation step of these restrictions within the submission system, in order to guarantee a minimum accessibility of the figures submitted. In the case the system would detect a lack of conformance of the submitted images to its internal rules, a message could warn the authors that this lack of conformance would provoke a delay in the revision process or even the rejection of the manuscript.

Some suggestions on how to implement them are to perform an automatic analysis of the submitted figure to check the presence of the alternative text, to check the correct format, color contrast, resolution and dimensions of the image, font size and font family of the text in the image, the use of standard units and abbreviations in the text of the image.

6.4.2.2 Specific technical recommendations related to the recommendation n.6

6a. Text alternatives: All informative images should have an equivalent text.

Details on how to create and select effective text alternatives according to current accessibility guidelines have been provided in section 4.8. The solution suggested by the author of the thesis is detailed in section.5.1.3.

Suggestion 6a is based on the success criterion 1.1.1 of the WCAG 2.0: "All non-text content that is presented to the user has a text alternative that serves the equivalent purpose"(W3C, 2014l)

6b. Use of color: Do not to rely on colors as the only means to convey information or functionality or distinguishing a visual element. It is recommended to use a combination of colors and patterns for conveying essential information.

Examples explaining the recommendation have been presented in section 4.1.4.

Suggestion 6b is based on the success criterion 1.4.1 of the WCAG 2.0: "Color is not used as the only visual means of conveying information, indicating an action, prompting a response, or distinguishing a visual element"(W3C, 2014o).

6c. Visual contrast: Provide enough contrast between the different elements of the image and between the informative text in the image and its background: minimum of luminosity contrast ratio of 4.5 to 1 for text less than 18 point if not bold and less than 14 point if bold. Preferred color palettes are green and magenta, turquoise and red, yellow and blue. For greyscale figures, provide at least a 20% greyscale difference between adjacent areas.

Suggestion 6c is based in several recommendations from the literature. According to the WCAG 2.0, the contrast should have the minimum of luminosity contrast ratio of 4.5 to 1 for text less than 18 point if not bold and less than 14 point if bold. If the background is a solid color (or all black or all white), each of the text letters should have 4.5:1 contrast ratio with the background. If the background or the letters vary in relative luminance or are patterned, as a letter marking a ROI on a pattered histological tissue, the background around the letters can be shaded in order to meet contrast requirements.

The great variety of color deficiencies makes the prescription of specific color pairs for ensuring effective contrast to all readers not feasible. The maximum contrast is obtained by the black/white combination. However, WCAG 2.0 suggest the use of colors that are composed predominantly of mid spectral components for light colors and the use of colors in the spectral extremes (blue and red wavelengths) for dark color, in order to enhance the contrast. For creating a sufficient (but not extreme) contrast, a light pastel background rather than a white background behind black text should be used (W3C, 2014b).

According to the principles of the Color Universal Design proposed by Okabe & Ito (2008), it is recommended the use of color palettes that can be easily identified by people with all types of color vision, such green and magenta, turquoise and red, yellow and blue (Wong 2011)(Landini & Perryer 2014). In general, it is recommended to avoid the combination of red and green. Specific recommendations on the use of color in images used in academic articles in biology are (Okabe & Ito, 2008):

- In the fluorescent double-staining micrographs, DNA chips, etc., use the combination of magenta (purple) and green instead of the red/green combination.
- For graphs and line drawings, use label elements of the graph on the graph itself rather than making a separate color-coded key.

For figures in greyscale, high contrast should be used to differentiate features in the image, for example a very pale grey against a very dark grey. Adjacent areas should have at least a 20% greyscale difference, ideally a difference of 25% to 30% (RNIB, 2010a).

6d. Line thickness: The lines used in images should be solid and, at least, 1 point thick.

Suggestion 6d is based in several recommendations from the literature. The lines used in images (such as those used lines, arrows, circles and polygons) should be thick and solid. The thickness of lines should be at least 1 point (NCBI, n. d.). Gray and color lines less than 1 point look broken and jagged due to the small dot pattern used to simulate a shade of gray or color tone (American Mathematical Society, 2008).

6e. Resolution and dimensions: Provide figures in biomedical articles according to standard dimensions in pixels (Table XLII).

Table XLII Universal standard dimensions in pixels for figures included in biomedical articles

	Resolution (dpi)	Minimum width half-page figure (inches)	Minimum width full-page figure (inches)	Minimum width half-page figure (pixels)	Minimum width full-page figure (pixels)
Line art	900	3.40	7	3060	6300
Grayscale art	400	3.40	7	1360	2800
Color art	300	3.40	7	1020	2100

These dimensions are based on recommendations in the literature and the results of the analysis of biomedical articles, which are described below. This suggestion should be considered as a possible solution for figures in biomedical articles and its applicability should be supported by further studies.

Recommendations in the literature and results of the analysis for the suggestion 6e.

WCAG 2.0 do not recommend any specific resolution or dimensions for images and just make a call to high-resolution images without giving any specific number of pixels. The National Council for the Blind of Ireland (NCBI) (NCBI, n. d.) recommends a minimum resolution of 300 dpi for digital or scanned photographs for printing at the size at which they will be used. NIMAS technical specifications recommend a resolution between 300 and 600 dpi for images for print works (National Center on Accessible Instructional Materials, 2006).

In the publishing industry, a resolution of 300 dpi is commonly considered the minimum acceptable standard for ensuring quality for the print reproduction of figures (University of Chicago Press, 2015). According to Lee (Lee, 2014), the most commonly recommended resolution for printing medical images on paper depends on the nature of the images:

- Color picture: 300 dpi
- Black and white pictures: 300 to 600 dpi
- Combination art (photo and text): 600 to 900 dpi
- Line art: 900 to 1,200 dpi.

Lee also states that the standard figure width of most academic journals in medicine is about 86 mm (3,385”) for single column images. According to Laberge (LaBerge & Andriole, 2003b), since most journals are commonly printed on 8.5 by 11 inch paper, the maximum possible image width is about 7 inches (LaBerge & Andriole, 2003b).

According to the analysis performed (chapter 5), the resolution recommended by the guidelines of images biomedical journals are:

- Color images: 266 to 600 dpi for color images.
- Grayscale images (any image with fine shading of gray: 266 to 600 dpi.
- Line art (black and white graphic with no shading, including scanned line drawings and line drawings in bitmap format): 600 to 1200 dpi.

The minimum final size for images recommended by the guidelines of biomedical figures reviewed ranges from 3.50 inches to 4 inches for half-page figures and from 6.33 inches to 8.50 inches for full-page figures (see section 5.1.2.4).

The actual width of figures measured in the biomedical articles in PDF analyzed ranges from 3.16 inches to 3.69 inches for half-page figures and from 6.33 inches to 7.38 inches for full-page figures (section 5.1.2.4).

Table XLIII summarizes the collected information concerning resolution and dimensions. It shows that the image resolution and dimensions derived from the analysis are similar to those specified in the literature.

Table XLIII Dimensions and resolution specified in the bibliography (columns, 2-4), suggested by the guidelines of biomedical journals reviewed and measured in the biomedical articles analyzed.

	Lee	Laberge	NCBI	NIMAS	Biomed journal guidelines	Actual image dimensions in biomed articles
Width (half-page)	3.385”	Not specified	Not specified	Not specified	3.50” – 4“	3.16” – 3.69”
Width (full-page)	Not specified	7”	Not specified	Not specified	6.33” - 8.50”	6.33” – 7.38”
Resolution for Color images	300 dpi	Not specified	300 dpi	300-600 dpi	266 dpi – 400 dpi	N/A
Resolution for Grayscale images	Not specified	Not specified	Not specified	300-600 dpi	266 dpi – 600 dpi	N/A
Resolution for line art	900 – 1200 dpi	Not specified	Not specified	300-600 dpi	300 dpi - 1200 dpi	N/A

In synthesis:

The width of half-page figures recommended by submission guidelines in biomedical articles ranges from 3.50” inches to 4” inches. The width of figures measured in the biomedical articles analyzed is 3.16 - 3.69 inches for half-page figures and 6.33” – 7.38” for full-page figures. These results are in line with the recommendations provided by Lee (3.4 inches approximately as a width for half-page images) and by Laberge (7 inches as a width for full-page images). The submission guidelines of biomedical journals suggest a resolution of 305 dpi for color images, 472 dpi for grayscale images and 983 dpi for line art. These results are in line current recommendations in the publishing industry and in accessibility guidelines.

The universal dimensions in pixels for biomedical figures proposed by the author of the thesis are based on these results.

6f. Text Font size: Taking into account the current font size of text used in biomedical images, the recommendation of accessibility guidelines of a minimum text font size of 12 pt is considered not feasible. All the recommendations of the journals analyzed suggested a text size lower than 12 point (see section 5.1.1.3) and no article among those reviewed used a text inside image with a 12-point font size or higher (see section 5.1.2.4). Minimum text font size more affordable to authors could be 8 pt, on condition that the image has a high resolution and the text is readable under high magnification rates.

Note: Beyond the size, the legibility of the text in raster images depends on its contrast with the background and the font type.

Recommendations in the literature and results of the analysis for the suggestion 6f

The use of text in raster images is generally not recommended (W3C, 2014i). However, a minimum of 12 pt size (ideally 14 pt) for text in images and in the article (and consequently in captions and mentions) is recommended by the literature (W3C, 2014m)(UKAAF, 2012) (NCBI, n. d.) (RNIB, n.d.).

The submission guidelines of biomedical journals reviewed (section 5.1.1.2) recommend a text size between 5 pt to 9 pt. The actual font size of text in images in the biomedical journal analyzed ranges from 6 pt to 8 pt.

6g. Text font family: Text in images and in image-related text should be in sans-serif font type.

Recommendations in the literature for the suggestion 6g

According to the WCAG 2.0, the presentation of text in sans-serif fonts (such as Verdana and Arial) makes the content more accessible (W3C Incubator Group Report, 2014). Sans-serif font family are considered more legible than serif fonts, such as Times New Roman, Georgia or Courier) (Ranaldi & Nisbet, 2010). Sans-serif font types increase the reading performances of readers with learning disabilities (such as dyslexia) over serif fonts (Rello & Baeza-Yates, 2013).

6h. Complex text: Unusual words or phrases should be defined or explained.

Recommendations in the literature for the suggestion 6h

WCAG 2.0 recommends providing the definition of a word or phrase used in an unusual or restricted way (W3C, 2014n).

6i. Acronyms and abbreviations: Abbreviations and acronyms used in textual labels inside the image should be spelled out or defined somewhere in the text (for example in the caption), especially for unusual abbreviations not used elsewhere in the article. The International system of Units (SI) abbreviations are recommended for common units of measurement; SI or standard abbreviations do not need to be spelled out.

Recommendations in the literature for the suggestion 6i

WCAG 2.0 recommend to provide a text expansion or explanation of an abbreviation (W3C, 2014e). ICMJE guidelines (International Committee of Medical Journal Editors, 2014) recommend to use only standard abbreviations and report results in both local and International System of Units (SI). An example of checklist for reviewing SI units in manuscripts is provided by the National Institute of Standards and Technology (National Institute of Standards and Technology, 2004).

6j. Image Formats: Vector images should be used instead of raster images whenever possible. Besides the current most accepted formats by biomedical publishers (PDF, TIFF; JPG, EPS and PPT), publishers should accept SVG format.

Recommendations in the literature for the suggestion 6j

WCAG 2.0 recommend vector images compared to raster images, especially to avoid the loss of quality when the image is enlarged (W3C, 2014i). The advantages of vector formats (and in particular of SVG) to raster images from an accessibility perspective have been described in section 3.2.

6k. Multipanel figures: Minimize the use of multipanel image.

Recommendations in the literature for the suggestion 6k

In section 4.8.1 it has been commented that images in biomedical articles present a high "conceptual complexity" and rely on detailed specialist knowledge and vocabulary (Burford et al., 2003). The complexity of biomedical image can increase dramatically in multipanel images, making the adaptation to blind people very hard to accomplish.

6.4.2.3 Synthesis of the "restriction" recommendations

The technical restrictions to images in the submission system according to the recommendations proposed are summarized in Table XLIV.

Table XLIV Technical restrictions proposed and features of the images to validate in the submission system.

Issue	Literature	Recommendation	Validation in submission system
Text alternatives	WCAG 2.0	All informative images should have an equivalent text.	text alternatives
Use of color	WCAG 2.0	Do not to rely on colors as the only means to convey information or functionality or distinguishing a visual element. It is recommended to use a combination of colors and patterns for conveying essential information.	colors and patterns in images

Issue	Literature	Recommendation	Validation in submission system
Visual contrast	WCAG 2.0; Okabe & Ito; RNIB; NCBI	Provide enough contrast between the informative text in the image and its background: minimum of luminosity contrast ratio of 4.5 to 1 for text less than 18 point if not bold and less than 14 point if bold. Preferred color palettes are green and magenta, turquoise and red, yellow and blue. For greyscale figures, provide at least a 20% greyscale difference between adjacent areas. The lines used in images should solid and with a thickness of at least 1 point.	minimum color contrast color palettes minimum line thickness
Resolution and dimensions	WCAG 2.0; Lee; Laberge	Provide figures in biomedical articles according to standard dimensions in pixels (see Table XLIII).	minimum dimensions
Font size	WCAG 2.0; UKAAF; NCBI, RNIB	Considering the current font size of text used in biomedical images, the recommendation of accessibility guidelines of a minimum text font size of 12 pt is considered not feasible. A more affordable minimum text font size should be a 8 pt.	minimum font size
Font family	WCAG 2.0; Ranaldi; Rello	Text in images and in image-related text should be in sans-serif font type.	sans-serif font type
Unusual words	WCAG 2.0	Unusual words or phrases should be defined or explained.	unusual words

Issue	Literature	Recommendation	Validation in submission system
Acronyms and abbreviations	WCAG 2.0	Abbreviations and acronyms used in textual labels inside the image should be spelled out or defined somewhere in the text (for example in the caption), especially for unusual abbreviations not used elsewhere in the article. The International system of Units (SI) standard abbreviations are recommended for common units of measurement.	abbreviations and standard units
Format	WCAG 2.0	Vector images should be used instead of raster images whenever possible. Besides the current most accepted formats by biomedical publishers (PDF, TIFF; JPG, EPS and PPT), publishers should accept SVG format.	format
Multipanel images	RNIB; Burford	Minimize the use of multipanel image	multipanel images

6.4.3 Environmental restructuring

This intervention aims at introducing a change in the physical and social context of the academic publishing.

It is proposed according to the bibliography reviewed, in particular according to the best practices guidelines for publishers produced by Editeur (Hilderley, 2013) and the principles included in the “Chart for Accessible Publishing” proposed by the Accessible Books Consortium (2014). They have been described in detail in section 1.3.

The publishers are the main responsible for executing the intervention suggested by the recommendation n.7 described below. This intervention affects the image editing workflow at the review, production and delivery steps.

6.4.3.1 Recommendation n.7

The publishers should include accessibility criteria in their publishing workflow based on the following interventions:

- Nominate a senior executive for putting in place a company policy committed with accessible publishing.
- Nominate a person (or team) responsible for the accessibility agenda.
- Raise accessibility awareness among, and provide technical training for, relevant staff:
    - the editorial and design responsible should be aware of issues surrounding accessibility and prepare files suitable for different digital platforms. They should be responsible for the preparation and editing of images with alternative descriptions.
    - The production and IT responsible should be able to provide different formats according to specific requests for accessible material. Text alternatives should be attached to each illustration included in the published material.
- Designate a point of contact to assist persons with print disabilities to access the publications.
- Test digital publications for accessibility and monitor their progress.
- Promote the adoption of accessibility standards throughout the supply chain.
- Support collaboration with organizations representing persons with disabilities.

6.5 Motivation

6.5.1 Persuasion

This action aims to persuade authors of the importance of publishing accessible images.

According to the analysis of the journals, four out of 12 biomedical journals analyzed did not offer an accessibility statement and only two made a reference to compliance with official accessibility legislation in the USA (508 Guidelines) (see section 5.1.1.1).

The motivations concerning the importance of accessibility have been detailed in "justification" section (1.2).

The publishers are the main responsible for executing the intervention suggested by the recommendation n.8 described below. This intervention affects the behaviors of authors at the editing step.

6.5.1.1 Recommendation n.8

Publishers should inform authors about:

- Disabilities prevalence: readers with disabilities are part of the audience of the academic literature and failing at providing fully accessible articles reverts on a negative perception of the service.
- Legal requirements: the universal accessibility of academic articles is regulated by law and the lack of compliance with this regulation could undermine the reputation of publishers and authors.

- Findability benefits: the text alternatives of images are a source of information for search engines and help them in indexing and retrieving figures.

6.5.2 Incentivisation

This action aims at eliciting positive or negative feelings among authors in order to motivate them to submit accessible images.

The publishers are the main responsible for executing the intervention suggested by the recommendation n.9 described below. This intervention affects the behaviors of authors at the editing step.

6.5.2.1 Recommendation n.9

Publishers should offer rewards for submitting accessible images. Some possible incentivisation actions could be:

- The article has priority in the review process.
- The author has a discount to the extra fee for publishing color images in printed version.
- Once the article is published, images with a correct alt text will be more retrievable and they will appear in first place in the page of results.

This last mechanism of incentivisation could be effective under the condition that the journal implement and promote a retrieval system explicitly covering text in text alternatives. Due to this condition, the author of the thesis considers that this recommendation is beyond an adjustment of the current publishing workflow that can be reasonably performed in a short term. It is suggested for a mid-long term intervention.

6.6 Proposal of tools

As commented previously, as a demonstration of possible aids to authors and as a first step to build training material for publishers, two tools were created in order to help authors in selecting and creating text alternatives using common the image-related texts currently adopted in biomedical publishing. They are described in detail below.

6.6.1 The "Image Text Alternatives Decision Tree"

The "Image Text Alternatives Decision Tree" (Figure 57) helps academic authors on deciding where and how to include the alternative text and to select the appropriate image-related text. The tool was inspired by the image sorting decision tree proposed by the DIAGRAM Center (Touch Graphics, Inc., n. d.) and it was contextualized in the academic publishing workflow. Compared to other decision trees, it has the advantage of including the high level categorization of the images' functions resulting from the analysis of the interviews as the initial element of decision for the selection of textual alternatives to images. As described in section 5.2.2.2, five functions for images have been identified:

1. Illustrative function (cited by nine participants): the figure illustrates or confirms the article content.
2. Summary/synthesis function (five participants): the figure synthesizes easily, clearly or succinctly the information provided textually or in a table.
3. Emphasis function (five participants): the figure emphasizes essential details.
4. Example function (four participants): the figure supports the argument by a visual example.
5. Data display function (four participants): the figure shows numerical data.

Regarding the illustrative function, an extra question about the presence of essential information for the article interpretation was added, due to the fact that images with this function may support and clarify the article content (methods, findings, results) and at the same time may contain essential information for the interpretation of the content itself (see section 5.2.2.2).

Once the function is decided, authors are prompted to evaluate the importance of the image content, taking also into account its related text. When an image-related text is considered descriptive enough, the alternative text would be reduced to a brief title identifying and localizing the image. This solution is in line with the techniques proposed by WCAG 2.0 (W3C, 2014d), and the preferences of blind and visually impaired users, as some authors confirm (Tomashek et al., 2013).

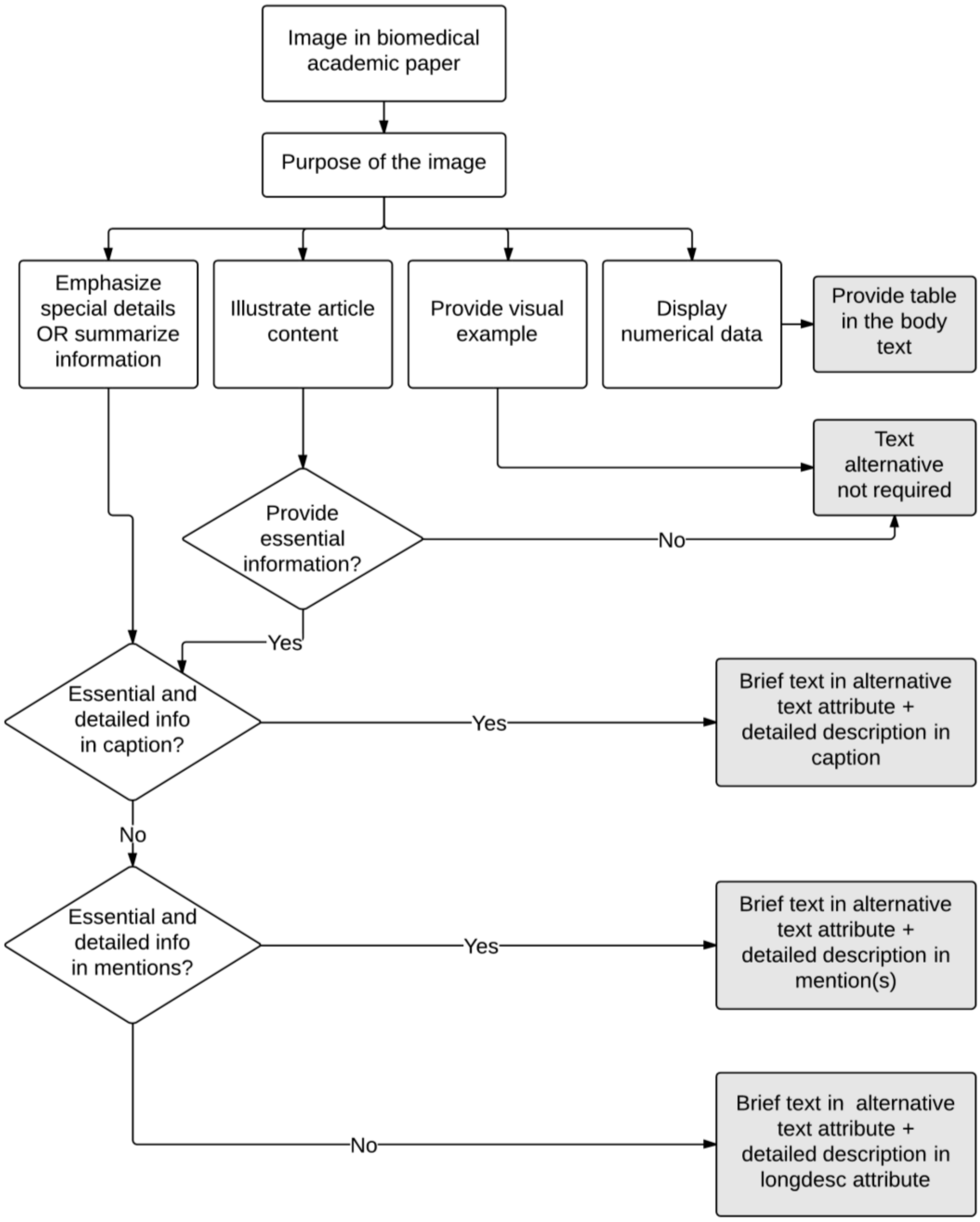


Figure 57 The Image Text Alternatives Decision Tree. Figure created by the author

The example in Figure 58 shows the application of the tool as a two-steps process in which: firstly the author observes different types on biomedical image in its article (a table and decorative figures in the middle of the table (a); a statistical graphic (b); an informative image (c) and a picture of himself (d). Secondly, the author follows the decision tree and decides which is the appropriate text alternative for each image: no text alternatives for figure “a”; a table for figure “b”; a detailed caption and a brief alt text for figure “c” and no text alternative for figure “d”.

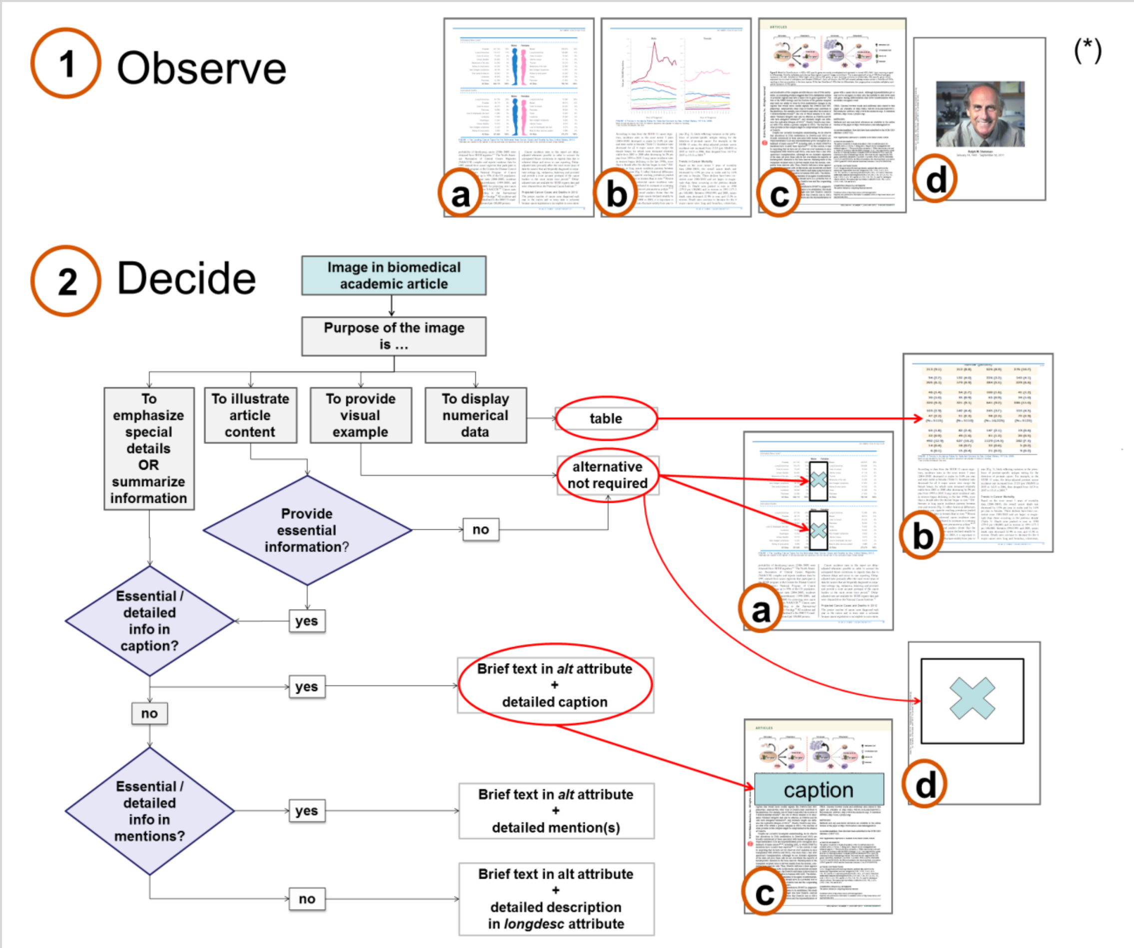


Figure 58 Example of a process for selecting text alternatives using the image text alternative description tree. Image extracted from: (Splendiani & Ribera, 2014b) . *The samples of medical article are extracted from (Siegel, Naishadham, & Jemal, 2012).

6.6.2 Checklist for the content of caption

In section 5.2.2.5, the information that should be included in captions was defined according to the answers of the participants to the interview and the recommendation of the bibliography reviewed (renowned manuals of style - Chicago Manual of Style (15th ed.), Publication Manual of the American Psychological Association (6th ed.) and American Chemical Society Style guide (3rd ed.) -; the submission guidelines of renowned international publishers in biomedicine - Springer, Wiley-Blackwell, New England Journal of Medicine, Nature and Science –; the guidelines provided by the International Committee of Medical Journal Editors) (Table XLV).

Table XLV Elements of the figures described in the caption cited by the participants of the interviews and recommended in the bibliography. In the case of medical images, there was no agreement among the participants about the inclusion of clinical information in the caption.

Figure element described in the caption	**Cited by the participants**	**According to manuals and submission guidelines**	**Apply to**
Identify/explain labels	x	x	all types of figure
Abbreviations	x	x	
Color code	x		
Modifications (focus, cropping, etc.)	x	x	
Source	x	x	
Description of variables	x	x	statistical graphics
Unit of measurement of the axis	x	x	
Statistical analysis details	x	x	
Areas of interest	x	x	other figures in Medicine and Biology (photos, etc.)
Magnification rate (scale)	x	x	
Modality	x	x	
Clinical information	*	x	
Cut type	x	x	other figures only in Medicine (photos, etc.)
Reconstruction type	x	x	
Method of staining		x	other figures in Biology (photos, etc.)
lanes in electrophoretic gels and blots		x	

Starting from these figure elements commonly included in caption, a checklist was developed with the aim of orienting authors in the creation of suitable, effective and meaningful captions as alternative texts for images. The checklist verifies whether the image is identified and labels or abbreviations are correctly applied and ask the author to explain colors, acquisition apparatus and modifications to the original image and to describe areas of interest and clinical background by answering the following questions:

Questions and recommendations applying to all types of figure:

- Is the image clearly identified? Identify it.
- Does the figure use labels? Explain labels, including symbols, letters, arrows and numbers.
- Does the figure use abbreviations? Explain abbreviations and acronyms.
- Does the figure convey information by colors? Explain the meaning conveyed by colors.
- Has any modification/ adjustment/ enhancement been applied to the figure? Explain the modification, selective digital adjustments and enhancements (cropping, brightness, color, sharpness, etc.).
- Does the figure refer to an external source? Specify the external source and appropriate credits.

Questions and recommendations applying to statistical graphics:

- Does the image include variables? Describe the variables.
- Does the figure show Cartesian coordinates (x and y axis)? Describe the units of measurement of the axis used in the graphic.
- Does the image provide a statistical analysis? Specify essential information related to the results of the statistical analysis showed in the figure, including the statistical significance, the type of analysis, information about the sample, the standard deviation, etc.

Questions and recommendations applying to other figures in medicine and biology (photos, etc.):

- Does the figure highlight areas of interest? Explain the areas of interest marked in the image.
- Has a magnification been applied to the figure? Specify the scale of the image.
- Has the figure been acquired in a specific modality? Specify the acquisition modality of the image (e.g. radiography, CT, etc.).
- Is the clinical background included? Describe the clinical background of the image (for example, the clinical profile of the patient showed in the image).

Questions and recommendations applying to other figures only in medicine (photos, etc.):

- Has a specific cut type been applied to the figure? Specify the cut type of the image (for example, for a MRI the cut can be axial, coronal or sagittal).
- Has any reconstruction been applied to the figure? Specify the reconstruction type of the image (for example, for a MRI the reconstruction can be a rendering 3D).

Questions and recommendations applying to other figures only in biology (photos, etc.):

- Has any method of staining been applied to the figure? Specify the staining method used in the image (for example, to enhance contrast in a tissue showed in a micrograph).
- Are electrophoretic gels and blots included in the figure? Detail any relevant information about lanes in electrophoretic gels and blots.

An application of the checklist is showed in the following example, according to the following scenario: a researcher has to submit an image for an article in a biomedical journal, as for example the image in Figure 59.

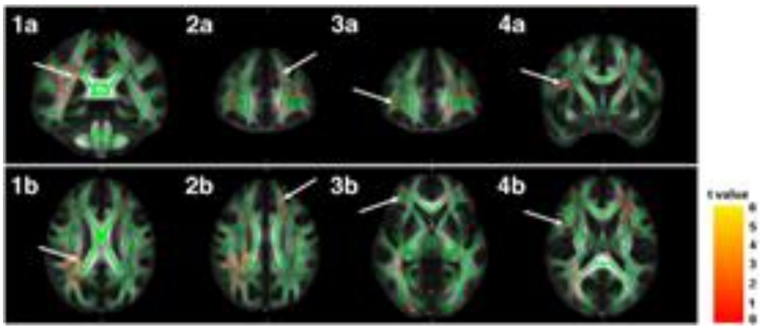


Figure 59 Example of a biomedical visualization for testing the checklist for caption creation. Extracted from: (Soriano-Raya et al., 2014).

He has created a caption (Figure 60) according to the guidelines provided by the publisher of the journal where he wants to submit the article. He wants to check (Table XLVI) if the caption includes all the information required for a meaningful and consistent description of the figure, in order to use it as an effective text alternative.

1 Figure 1. Regional areas of reduction of fractional anisotropy (FA). Clusters showing significant 5
5 reduction of FA in participants with high-grade deep white matter hyperintensities are displayed on 6
7 coronal and axial sections of an FA map. The FA skeleton used for statistical analyses is superposed
4 in green. White arrows at images 1a (y = -37) and 1b (z= 26) show peak-value voxel location of
reduced FA in the right anterior thalamic radiation. White arrows at images 2a (y = 36) and 2b (z =
30) show peak-value voxel of reduced FA in the left inferior fronto-occipital fasciculus (IFOF). White 2
arrows at images 3a (y =36) and 3b (z = - 1) show peak-value voxel of reduced FA in the right IFOF.
White arrows at images 4a (y = 6) and 4b (z = 18) show peak-value voxel of reduced FA in the right
superior longitudinal fasciculus. The color scale indicates the magnitude of t-values with lowest 3
4 appearing in dark red and the highest in bright yellow. Images are displayed in radiologic 8
convention (right side represents left side and left side represents right side of the brain.

Figure 60 Original caption of the Figure 59. The text has been segmented in numbered sections for its identification in the checklist. Extracted from: (Soriano-Raya et al., 2014).

Table XLVI Checklist questions, application and localization of the recommended information in the caption of Figure 59 and specific suggestions for the correct application of the recommended information.

Question	Correct?	Detailed suggestion
Is the figure clearly identified?	Yes (1)	No modification is required.

Question	Correct?	Detailed suggestion
Does the figure use labels?	Yes (2)	No modification is required.
Does the figure use abbreviations?	No (3)	A modification should be applied: substitute “t-values” by “Student's t-test values”.
Does the figure convey information by colors?	Yes (4)	No modification is required.
Has any modification/ adjustment/ enhancement been applied to the figure?	N/A	
Does the figure refer to an external source?	N/A	
Does the figure include variables?	N/A	
Does the figure show Cartesian coordinates (x and y axis)?	N/A	
Does the figure provide a statistical analysis?	No	An addition should be applied: Provide the p value.
Does the figure highlight areas of interest?	Yes (2 & 5)	No modification is required.
Has a magnification been applied to the figure?	N/A	
Has the figure been acquired by a specific modality (e.g. radiography, CT, etc.)?	No (8)	An addition should be applied: include “Diffusion tensor imaging (DTI) of the brain”.
Is the clinical background included?	Yes (6)	No modification is required.
Has a specific cut type been applied to the figure?	No (7)	A modification should be applied: substitute “on coronal and axial sections” by “on coronal (1a-4a images) and axial (1b-4b images) sections”.

Question	Correct?	Detailed suggestion
Has any reconstruction been applied to the figure?	No	An addition should be applied: “For details on preprocessing methods see Materials and Methods section”.
Has any method of staining been applied to the figure?	N/A	
Are electrophoretic gels and blots included in the figure?	N/A	

The final figure caption will be as the one showed in Figure 61.

> Figure 1. Diffusion Tensor Imaging (DTI) of the brain with regional areas of reduction of fractional anisotropy (FA). Clusters showing significant reduction of FA in participants with high-grade deep white matter hyperintensities are displayed on coronal (1a-4a images) and axial (1b – 4b images) sections of an FA map. The FA skeleton used for statistical analyses is superposed in green. White arrows at images 1a (y =37) and 1b (z = 26) show peak-value voxel location of reduced FA in the right anterior thalamic radiation. White arrows at images 2a (y = 36) and 2b (z = 30) show peak-value voxel of reduced FA in the left inferior fronto-occipital fasciculus (IFOF). White arrows at images 3a (y = 36) and 3b (z = 1) show peak-value voxel of reduced FA in the right IFOF. White arrows at images 4a (y = 6) and 4b (z = 18) show peak-value voxel of reduced FA in the right superior longitudinal fasciculus. The color scale indicates the magnitude of Student’s t-test values (P value < 0,05) with lowest appearing in dark red and the highest in bright yellow. Images are displayed in radiologic convention (right side represents left side and left side represents right side of the brain). For details on preprocessing methods see Materials and Methods section.

Figure 61 Revised caption of the Figure 59 with the proposed modifications highlighted.

A brief text alternative should also be provided in the alternative text attribute to locate and identify the image, as recommended in the “Image text alternatives decision tree”. In this case, a suitable alternative text attribute could be: “Diffusion Tensor Imaging (DTI) showing measurement of fractional anisotropy. Details in the text following the figure”.

6.7 Synthesis of the proposal

In this chapter, a proposal for inclusion for image accessibility criteria in the current workflow of biomedical academic publishing has been presented. The proposal is structured according to a set of interventions addressed to the actors of the publishing workflow in order to improve the capability, opportunity and motivation for publishing accessible images. The interventions are expressed as a set of nine recommendations.

Table XLVII synthetizes the recommendations, the actors responsible for the recommended intervention and the step in the publishing process to which the intervention is related.

Table XLVII Recommendations for including accessible criteria in the image publishing workflow, with the related components, intervention functions, number of recommendation (R:ID), actor and step.

Component	Intervention function	R.ID.	Recommendation	Actor	Step
Capability	Education	**1**	Inform authors about the accessibility of the information resources provided and the needs and visual barriers of readers with special needs.	Publisher	editing
		2	Offer tools integrated in the submission system simulating image perception by readers with visual impairments.	Publisher	editing
	Training	**3**	Provide guidelines, examples and evaluation tools on how to make accessible images.	Publisher	editing
		4	Adopt internationally recognized guidelines for manuscript submission.	Publisher	editing
Opportunity	Enablement	**5**	Fix the current shortcomings of common editing tools used by the authors in supporting accessibility.	Image editing software developers	editing
	Restriction	**6**	Introduce a validation step of the technical characteristics of the images related to accessibility within the submission system.	Publisher	editing / submission
	Environmental restructuring	**7**	Include accessibility criteria within the publisher image editing process.	Publisher	review/ production/ delivery
Motivation	Persuasion	**8**	Inform authors about disabilities prevalence, legal requirements and findability benefits of accessibility	Publisher	editing
	Incentivisation	**9**	Publishers should offer rewards for submitting accessible images.	Publisher	editing

Figure 62 illustrates how the recommendations relates to the steps of the publishing workflow and which actors are involved in the change. Although the authors are not showed in the figure, they are the main responsible for the creation and submission of accessible images once the interventions have been performed.

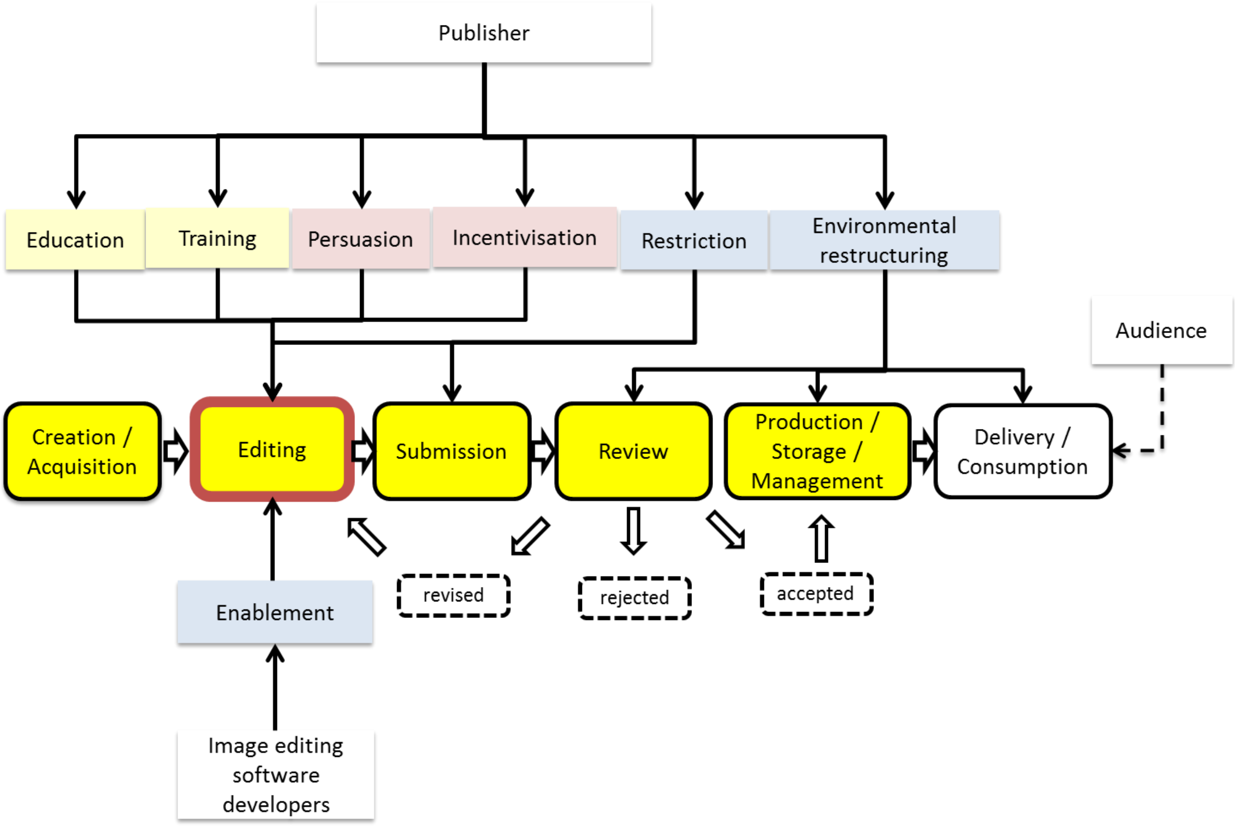


Figure 62 The publishing workflow steps with related recommended interventions.

Chapter 7

Conclusions

7 Conclusions

7.1 Advantages and limitations of the proposal

In the academic publishing of STEM journals, the process of making images accessible is still not accomplished. Publishers do not integrate accessibility requirements in the publishing workflow, they are not aware of the social needs and legal requirements supporting digital accessibility and do not ask authors for accessible images. At the same time, authors do not have guidelines, standard and tools for making images accessible and they are completely unaware of the needs of people with special needs.

The consequence of this situation is that currently the products of academic publishing in general and images in particular, are not accessible to people with visual and cognitive impairments or under contextual limitations. As a side effect of this lack of accessibility, visual information provided by scientific articles can neither be indexed nor easily retrieved.

The thesis proposes the adoption of accessible requirements from the beginning of the "image lifecycle", considering the current policies and practices in each publication step of the actors involved in the publishing workflow. Specific interventions for the inclusion of accessibility criteria are introduced inside the process of producing academic articles and connected to the current practices of authors and publishers. The adoption of interventions that aim to modify the behaviors of authors of scientific images according to the Behaviour Change Wheel model contributes to solve the issue of making scientific images accessible in a "sustainable" way, avoiding severe changes of the current practices on the creation, editing and submission of images in academic articles.

In this thesis, the current state of images in biomedical publications and other STEM publications has been reviewed from the accessibility viewpoint and the study has been performed in different stages. Firstly, the public who benefit of accessible publishing and the current mechanism for making images accessible has been defined (chapter 4). Secondly, the lack of accessible policies and practices for images has been stated by reviewing a group of highly cited journals and articles (chapter 5). Thirdly, the publishers' requisites for the image submission and for the characteristics of the images published in the articles have been assessed. In particular, the characteristics of the text commonly used to describe images and its potential use as a text alternative has been stated (chapter 5). Fourthly, a qualitative study of practices and opinions of researchers in medicine and biology has been performed, with the aim to define the characteristics, uses and functions of image and image-related texts in academic articles.

The sample of journals and articles in biomedicine, computer science and mathematics selected of the analysis is small compared to the whole number of journals and articles in these fields. However, by choosing the journals with the highest impact factor, the most valued publications are described and those which lead publishing directions in each research field.

The sample of 198 researchers participating in the survey and the sample of 22 researcher interviewed is quite small compared to the entire number of researchers in medicine and biology in Spain. However, taking into account the lack of specific studies on the topic in the literature at a national and international level, the study brings new and interesting data on author practices.

The main contribution of the thesis is the proposal of the integration of accessibility requirements into the image publishing workflow, according to the practices, policies and opinions of the main actors in the publishing chain and in particular of the authors, the main responsible for the creation, editing and submission of images in scientific articles.

The contributions, key issues, advantages and limitations of the thesis are detailed below.

7.2 Contributions of the thesis

Most relevant contributions of the thesis are detailed below according to the objectives and main questions of the thesis outlined in chapter 1.

7.2.1 Objective 1. To define the profile of the target reader of accessible images.

The main target readers who benefit of producing accessible image in academic articles are users with blindness, low vision and cognitive impairments. Also readers with a situational limitation using a "disabling technology" (e.g., a mobile device with a reduced screen) or under specific environmental conditions (e.g., accessing digital text in poor or excessive lighting conditions). Finally, search engines, although not human, work in a similar way to users who are blind and also benefit from accessibility improvements.

7.2.2 Objective 2. To review the current solutions offered by the accessibility discipline for accessible images.

1. WCAG 2.0 offers universally design principles for accessible images according to the needs of users and the barriers they experience when accessing images:
   - Add an equivalent textual description to the image.
   - Any information presented in color should also be presented in a way that does not depend on color perception, for example using patterns
   - Provide sufficient visual contrast between all visual information (including text inside images) and its background.
2. The definition of these elements allows images of STEM articles to be designed with universal accessibility.
3. Among these recommendations, the use of text alternatives is considered the most effective. The most common mechanisms for the inclusion of text alternatives in the widespread document formats used in academic publishing (HTML and PDF) are: the *alt*

attribute and the *longdesc* attribute of the *img* element in HTML documents and the *alternate text* and *actual text* attributes in PDF documents. They currently present some limitations on meeting the requirements of good text alternatives such as availability, discoverability (visible for sighted people and programmatically determined) and the possibility to include structured text.

4. Image-related texts used in STEM articles, such as captions, mentions in the main text, text inside the image and metadata, may be a suitable solution for conveying text alternatives. They are technically supported by the mechanisms currently available for associating text alternatives to images, including figures in academic articles in PDF and HTML formats.
5. Other possible solutions for providing text alternatives are offered by the Library and Information Science discipline and consist in the use of metadata within the same file of the image or the use of external metadata, located in a separate file. In the case of internal metadata the text description embedded in the image file may be easily preserved along the editing and publishing image workflow. However, the lack of interoperability among formats and tools for authoring/editing images and for metadata handling currently prevents this method from being totally effective.
6. Semantic Web technologies may offer an opportunity to enrich alternative descriptions of images by providing structured and machine-understandable information readable by assistive technology and to automate the process of assigning text alternatives.
7. Natural-language summarization techniques are a suitable solution to enable the automatic text description of graphics. Until now, they are mainly focused on highly-structured standard graphics (such as statistical graphics) and they may be extended to other kind of figures, taking into account that the text automatically extracted from figures without standardized visual elements (such as biomedical figures) is difficult to understand when taken out of context.
8. Currently, organizations and institutions in the accessibility area offer many guidelines on how to write efficient text alternatives to them. In particular they offer: recommendations about which content should be included in a text alternative in order to make it effective; decision trees for choosing the appropriate text alternative according to the function of the image and its context; templates with basic information and some gaps to fill with the particular information of the graphic; examples of specific type of images with their corresponding alternative textual descriptions. However they have the following shortcomings:

   - They do not provide detailed guidelines on how to textually describe complex or technical images with high information density, as those commonly used for example in medicine and biology fields.
   - They are mainly addressed to a generic audience and intended for common publications and do not address all the categories of target public defined in this research.

- They are not related to the current practices and purposes of academic authors in the process of creating and selecting images for image submission to academic journals and they do not take into account the different functions of the images defined according to the authors' opinions.
- They are mostly based on sets of best practices that different institutions have developed in the accessibility field according to their experience, rather than on common theoretical principles on visual information creation and description.

7.2.3 Objective 3. To assess whether the figures in STEM academic articles are currently accessible, with a particular attention to the biomedical discipline.

Currently, images are not accessible in highly-cited academic articles in the analyzed STEM fields (biomedicine, computer science and mathematics). In particular, images do not have text alternatives associated to them. These conclusions confirm the two initial assumptions of the thesis: images are currently not accessible in academic articles in STEM fields and they do not have text alternatives associated to them.

When a text alternative is applied (especially in the *alt* attribute of *img* element in HTML version), its content is incorrect. Text alternatives are absent or incorrect even in articles belonging to journals that have an accessibility policy on the website and offer references to the use of text alternative in the author's submission guidelines.

The design requirements concerning the visual features of the figures (use of color, minimum color contrast, legible font type, etc.) are partially conformed in the articles of the three fields. Articles from the three research fields do not usually have images with a minimum color contrast and the text inside the images uses very small font sizes (between 7 or 8 points). Mathematical articles have specific characteristics as they include a reduced number of color images; text of figures and captions do not commonly use abbreviations and are not in sans-serif font type.

The lack of accessibility of scientific images confirms the need for improving their accessibility and reinforces the justification of the thesis.

7.2.4 Objective 4. To assess whether the figures in STEM academic articles are relevant, with a particular attention to the biomedical discipline.

Figures are extensively used in STEM articles and, quantitatively speaking, they are currently an important element of information. However, there are differences in the total number of figures among publications in different disciplines. The average number of figures per page in biomedical articles is one per 3.33 pages, in computer science one per 2.4 pages and in mathematics one per 4.5 pages.

Researchers in medicine and biology extensively use images in their academic articles. According to them, images have an essential role in providing information that complements the text of the articles.

7.2.5 Objective 5. To assess whether the publishers of STEM journals currently offer effective instructions for authors to make their images accessible.

The guidelines specifying how to submit images the journals in biomedicine, computer science and mathematics do not offer effective instructions on how to prepare and submit accessible images.

Only very few guidelines of the journals in the three research fields occasionally mention issues related to image legibility and their impact on readers with special needs.

Currently, there is no agreement among publishers of journals in the same research field and among different research fields concerning image submission requirements.

Only few journals explicitly declare to follow international guidelines for submission instructions, suggesting that currently there is no standard for the image submission requirements.

7.2.6 Objective 6. To explore to potential use of current image-related texts currently used in STEM and in particular in biomedical academic articles as alternative descriptions of the images.

The use of captions and mentions of images in the main text of the article and text labels in figures is widespread in biomedicine, computer science and mathematics.

Captions and mentions have promising features for serving as image alternatives: they are semantically dense and often they can be programmatically associated with the images. In particular, captions and mentions in biomedical articles have characteristics compatible with long text alternatives: they are used in almost all the articles, they are long enough for them to be recognized as a detailed description of the figure's content, and they are often located on the same page as the figure they refer to.

Although currently the caption is used by authors as a complement for the visual information and, therefore, it does not replace completely the visual information conveyed by the image neither its function when the image is not available, the text included in captions mainly provides a syntactic description of images (what the image looks like) and could be used as a text description of the image applying some specific adjustments according to the type of image they describe.

Concerning the variability of the function and content of captions and mentions in different research areas, some "implicit patterns" seem to depend on the type of image, its function and its intended message. Biomedical figures present illustrations of complex biological processes that in general require detailed textual descriptions. On the other hand, mathematical images are mainly schematic line graphs representing math expressions and abstract concepts which seem to require a reduced number of words for their description.

Text labels and other text inside figures can be automatically recognized and extracted easily only from vector images, and also for raster images with the help of optical character recognition. However, this type of text cannot be used as an effective mechanism for providing text alternatives to images as it is meaningless when separated from its visual context.

Metadata are sparingly used in figures in biomedical and computer science articles and almost not at all in mathematics. They mostly concern the technical features of the image, while the descriptive metadata are limited to the *author* and *title* fields, and even so, the actual use of the title is uninformative. Currently, metadata cannot be considered a feasible mechanism to provide a text alternative to images, as their use is hindered by imaging software and authors do not use them.

7.2.7 Objective 7. To assess how and why authors currently create and submit images in biomedical academic articles.

The image file formats most used in academic articles by researchers in medicine and biology are JPEG and TIFF and, in a less extent, other image and document formats that can contain images (PDF, PPT, GIF, PNG, etc.). Image formats are chosen according to the publishers' guidelines for the submission of articles. Publishers do not take into account a format accessibility support for their selection. For example, the SVG format is almost completely ignored.

The most used tools for editing images are Microsoft PowerPoint and Adobe Photoshop. Although PowerPoint is not an image editing software, its widespread availability among researchers, its ease of use and the support of the most common image editing tasks justify its extensive use. Even though these tools offer some accessibility features and also provide documentation about their accessibility support, they currently present some limitations that hinder their effective use in the creation of accessible figures.

Authors of academic articles in medicine and biology generally are not aware of accessibility needs and guidelines and they do not apply accessible design guidelines in the creation, editing and submission of figures in academic articles.

The authors strictly follow the submission requirements of the publisher, which drove their process of editing and submission of images and textual descriptions.

The authors include images in scientific articles according to five different functions:

1. to illustrate or confirm the article content (illustrative function)
2. to synthesize easily, clearly or succinctly the information provided textually or in a table (summary/synthesis function);
3. to emphasize essential details (emphasis function);
4. to support the argument by a visual example (example function);
5. to show numerical data (display function).

The authors select figures to include in their publications according to five criteria:

1. the "adherence" of the image to the content of the article (relevance criterion);
2. the ability to present visual information with appropriate resolution, contrast and sharpness (quality criterion);

3. the ability to be interesting and exciting and to attract the attention (Impact criterion);
4. the ability of a multi-panel image to make easier the comparison among the different figures that compose the entire image (Comparison criterion);
5. the ability to provide a high density of information (density criterion).

The findings summarized in this section have the limitation to be based on the practices and attitudes of researchers mostly belonging to one university (University of Barcelona). Further studies are required in order to compare and validate the results according to the practices of researchers from other institutions in Spain and other countries.

7.2.8 Objective 8. To propose recommendations on how to foster the adoption of accessibility principles concerning the images within the biomedical academic publishing, and for extension in all STEM fields.

A set of interventions addressed to the actors of the publishing workflow of scientific articles are defined in order to improve the capability, opportunity and motivation for publishing accessible images. The interventions are expressed as a set of nine recommendations, addressed to specific actors of the publishing workflow and specifying the step in the publishing process:

1. Publishers should inform authors about the accessibility of the articles provided and the needs and visual barriers of readers with special needs (editing step).
2. Publishers should offer tools integrated in the submission system simulating image perception by readers with visual impairments (editing step).
3. Publishers should provide guidelines, examples and evaluation tools on how to make accessible images (editing step).
4. Publishers should adopt internationally recognized guidelines for manuscript submission (editing step), and in particular for image submission.
5. Image editing software developers should fix the current shortcoming of common editing tools used by the authors in supporting accessibility (editing step).
6. Publishers should introduce a validation step of the technical characteristics of the images related to accessibility within the submission system (editing/submission steps).
7. Publishers should include accessibility criteria within the publisher image editing process (review/production/delivery steps).
8. Publishers should inform authors about disabilities prevalence, legal requirements and findability benefits of accessibility (editing step).
9. Publishers should offer rewards for submitting accessible images.

Two tools are created in order to help authors in selecting and creating text alternatives using common image-related texts currently adopted in biomedical publishing: the “image text alternatives decision tree”, which helps academic authors on deciding where and how to include the alternative text and to select the appropriate image-related, and a checklist for orienting authors in the creation of effective captions as alternative texts for images. These tools have the advantage of connecting abstract accessibility guidelines to the current publishing practices. They convert the creation of accessible images into a part of the content-writing practices of authors.

The proposal shows that the integration of accessibility requirements in the image creation and submission process in most cases does require an extra effort but not excessive compared to the current process of creation and submission of images. In particular, the proposal includes recommendations on how to make accessible images based on current practices in image editing, in order to minimize authors extra work for adapting images to special needs and proposes the use of suitable, effective and meaningful captions and mentions as alternative text for image, in order to spare authors (and publishers) the task of creating alt texts.

The main advantage of the proposal is that, while current accessibility guidelines are often general and disconnected from the policies and practices of the specific context, the proposal is tailored to the specific context of biomedical academic publishing and is linked to the current practices in image submission and publishing. Therefore, it aims at minimizing the extra work of adapting images to special needs requirements and limiting the changes in the current image publishing workflow required for the implementation of accessibility requirements, increasing the opportunities of an actual application of accessibility principles.

Furthermore, this approach emphasizes the issue that a general answer to image accessibility – especially concerning the textual description of visual content - is not feasible.

The main limitation of the proposal is to be based on the practices of a small number of researchers in the specific fields of medicine and biology. The model of textual image description should be tested for checking its validity and reliability with experts in each specific academic domain. Further studies should also be performed evaluating the extent to which the solutions proposed can be implemented into the publishing workflow according to factors as costs, conflicts with other policies, etc. However, taking into account the lack of studies that relate the current practices in the publishing of visual information in academic article to accessibility requirements, this study can be reasonably assumed as valuable for the novelty of its approach and for the scope of its vision, since it aims at covering the mismatch between accessibility guidelines and the ability of organizations to implement these changes from an "holistic" perspective. This perspective does not only take into account the technical requirements for achieving an effective implementation of accessibility into the academic publishing workflow, but also considers the practices, objectives and motivations of the researchers involved in the process of authoring and submitting visual scientific figures.

7.3 Future research

Despite the contributions of this thesis, there remain some outstanding issues that require further explorations:

- The proposed model of interventions should be validated. In particular, the model for describing image textually should be tested to check its validity and reliability with experts in each specific academic domain.
- Further studies should be performed evaluating the extent to which the solutions proposed can be implemented into the publishing workflow according to factors as costs and conflicts with other policies.
- The proposal is focused on providing accessible images in HTML and PDF documents, the current main formats for academic publishing dissemination. EPUB3, the most promising format for providing accessible documents in academic publishing, should be taken into account in further studies.
- Once validated, the model could be developed further and extended to images in other research fields.
- The issue of how to bind textual description to images and maintain them along the publishing process is still unsolved. The thesis suggests the use of embedded metadata stored in images as a way for solving the issue. In particular, a possible research line may concern the exploration of XMP metadata for conveying and preserving text alternatives.
  Finally, the thesis has identified Semantic Web technologies as potential mechanisms for providing semantic text descriptions of images in academic publications. Further research may explore how to share common criteria and join mechanisms between Accessibility and Semantic Web research fields for improving a "universal access" to images by humans and machines.

Chapter 8

Contributions

8 Contributions

This section lists all the publications derived from the studies developed in the thesis and disseminated by articles in published academic journals and by presentations and poster sessions in international conferences and workshops.

The interdisciplinary perspective of the thesis about the alternative representation of images described in chapter 4 was presented in the 13th International Conference on Computers Helping People with Special Needs (ICCHP'13) (Splendiani et al., 2012).

The review of the current guidelines for writing text alternatives of images described in chapter 4 were presented in a poster session in the III International Seminar on LIS Education and Research (LIS-ER) (Splendiani & Ribera, 2015b). The potentialities of the Semantic Web technologies for improving image accessibility described in chapter 4 was presented in the W3C Workshop on Annotations (Splendiani & Ribera, 2014b).

The analysis and the results of the analysis described in chapter 5 and part of the proposal described in chapter 6 were published in the Journal of the Association for Information Science and Technology (Impact Factor: 2.23) (Splendiani & Ribera, 2015a), in the Procedia Computer Science (Splendiani & Ribera, 2014a) and in the Journal of Digital Imaging (Splendiani et al., 2014) (Impact Factor: 1.20). They were also presented in the 28th Annual International Technology and Persons with Disabilities Conference (CSUN 2013) (Splendiani, Ribera, & Centelles, 2014).

The proposal for the inclusion of accessibility criteria in the publishing workflow of images described in Chapter 6 was presented in the 6th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Infoexclusion (DSAI 2015) (Splendiani & Ribera, 2015c) and in the XV International Conference on Human Computer Interaction (Interacción 2014) (Splendiani & Ribera, 2014b).

Chapter 9

Bibliography

9 Bibliography

Abu Doush, I., Pontelli, E., Son, T. C., Simon, D., & Ma, O. (2010). Multimodal Presentation of Two-Dimensional Charts: An Investigation Using Open Office XML and Microsoft Excel. *ACM Transactions on Accessible Computing*, *3*(2), 8:1–8:50. http://doi.org/10.1145/1857920.1857925

ACCESS-ed. (2011). Equivalent Text Descriptions. Available at http://access-ed.r2d2.uwm.edu/EqTDs/ [Accessed 9 April 2015].

Accessible Books Consortium. (2014). Accessible Book Consortium. Available at http://accessiblebooksconsortium.org/portal/en/index.html [Accessed 9 April 2015].

Accessible Publishing Consortium. (2014). Charter for Accessible Publishing. Available at http://accessiblebooksconsortium.org/portal/en/charter.html [Accessed 9 April 2015].

Adobe Systems Incorporated. (2006). PDF Reference Sixth edition. Available at http://www.adobe.com/content/dam/Adobe/en/devnet/acrobat/pdfs/pdf_reference_1-7.pdf [Accessed 9 April 2015].

Agarwal, S., & Yu, H. (2009). FigSum: automatically generating structured text summaries for figures in biomedical literature. In *Proceedings of the 2009 Annual Symposium of the American Medical Information Association (AMIA 2009)*, 6–10. San Francisco, CA. American Medical Information Association.

AIM Commission. (2013). The Advisory Commission on Accessible Instructional Materials in Postsecondary Education for Students with Disabilities Commission Report. Available at http://www2.ed.gov/about/bdscomm/list/aim/index.html [Accessed 9 April 2015].

Altmanninger, K., & Wöß, W. (2008). Accessible Graphics in Web Applications: Dynamic Generation, Analysis and Verification. En K. Miesenberger, J. Klaus, W. Zagler, & A. Karshmer (Eds.), *Computers Helping People with Special Needs* (pp. 378-385). Springer Berlin Heidelberg.

American Association for the Advancement of Science. (2015). Preparing Your Manuscript: Art & Figures. Available at http://www.sciencemag.org/site/feature/contribinfo/prep/prep_revfigs.xhtml [Accessed 9 April 2015].

American Chemical Society, (2006). *The ACS Style Guide: Effective Communication of Scientific Information*, 3rd ed. Anne M. Coghill and Lorrin R. Garson. Oxford University Press.

American Mathematical Society. (2008). Author Resource Center. Available at http://www.ams.org/publications/authors/authors [Accessed 9 April 2015].

American Psychological Association. (2010). *Publication Manual of the American Psychological Association, Sixth Edition*. Washington, DC: American Psychological Assoc.

Anighoro, A., & Rastelli, G. (2013). Enrichment factor analyses on G-protein coupled receptors with known crystal structure. *Journal of Chemical Information and Modeling*, *53*(4), 739-743. http://doi.org/10.1021/ci4000745

Annual Reviews. (2014). Instructions for the preparation of the manuscripts. Available at http://www.annualreviews.org/userimages/ContentEditor/1343842722976/AuthorHandbook-BLUE.pdf [Accessed 9 April 2015].

Antani, S., Demner-Fushman, D., Li, J., Srinivasan, B. V., & Thoma, G. R. (2008). Exploring use of images in clinical articles for decision support in evidence-based medicine. In Proceedings of the SPIE-IS&T Electronic Imaging Symposium, 6815. San Jose, CA. SPIE--The International Society for Optical Engineering http://doi.org/10.1117/12.766778

Apostolova, E., You, D., Xue, Z., Antani, S., Demner-Fushman, D., & Thoma, G. R. (2013). Image retrieval from scientific publications: Text and image content processing to separate multipanel figures. Journal of the American Society for Information Science and Technology, 64(5), 893–908. doi:10.1002/asi.22810

Association of American Publishers. (2013). EPUB 3 Implementation Project White Paper. Available at http://publishers.org/press/117/ [Accessed 9 May 2015].

Attwood, T. K., Kell, D. B., McDermott, P., Marsh, J., Pettifer, S. R., & Thorne, D. (2009). Calling International Rescue: knowledge lost in literature and data landslide! *The Biochemical Journal*, *424*(3), 317-333. http://doi.org/10.1042/BJ20091474

Audhkhasi, K., & Narayanan, S. (2013). A globally-variant locally-constant model for fusion of labels from multiple diverse experts without using reference labels. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(4), 769-783. http://doi.org/10.1109/TPAMI.2012.139

Ault, H. K., Deloge, J. W., Lapp, R. W., Morgan, M. J., & Barnett, J. R. (2002). Evaluation of Long Descriptions of Statistical Graphics for Blind and Low Vision Web Users. En K. Miesenberger, J. Klaus, & W. Zagler (Eds.), *Computers Helping People with Special*

Needs (pp. 517-526). Springer Berlin Heidelberg. Available at http://link.springer.com/chapter/10.1007/3-540-45491-8_99

Aydıngöz, Ü. (2005). Figures, tables, and references: integral but sometimes neglected components of scientific. *Diagn Interv Radiol*, *11*, 67–68.

Ballesteros, E., Ribera, M., Pascual, A., & Granollers, T. (2014). Reflections and proposals to improve the efficiency of accessibility efforts. *Universal Access in the Information Society*, 1-4. http://doi.org/10.1007/s10209-014-0356-1

Barnes, C. (2002). 'Emancipatory disability research': project or process? *Journal of Research in Special Educational Needs*, *2*(1), 2: no. doi: 10.1111/j.1471-3802.2002.00157.x

Beaudoin, J. (2009). An investigation of image users across disciplines: A model of image needs, retrieval and use. *Proceedings of the American Society for Information Science and Technology*, *45*(1), 1 - 5. http://doi.org/10.1002/meet.2008.14504503118

Benetech. (2012). Poet image Description - Bookshare Image Description Guidelines. Available at https://diagram.herokuapp.com/edit_book/description_guidance [Accessed 9 April 2015].

Benetech. (2013). Accessibility Metadata Project Best Practices Guide Version 1.0.1. Available at http://www.a11ymetadata.org/resources/ [Accessed 9 April 2015].

Benetech. (2014). Born Accessible. Available at http://benetech.org/our-programs/literacy/born-accessible/ [Accessed 9 April 2015].

Berkenbilt, J. (2014). What is QPDF? Available at http://qpdf.sourceforge.net/ [Accessed 9 April 2015].

Bhatia, S., Lahiri, S., & Mitra, P. (2009). Generating synopses for document-element search. In *Proceedings of the 18th ACM conference on Information and knowledge management* (CIKM 2009) (pp. 2003–2006). Hong Kong, China: ACM. doi:10.1145/1645953.1646287

Bias, R. G., & Mayhew, D. J. (2005). *Cost-justifying Usability: An Update for an Internet Age*. Morgan Kaufmann.

Bigham, J. P., Jayant, C., Ji, H., Little, G., Miller, A., Miller, R. C., Miller, R., Tatarowicz , A., White, Y., White, S. & Yeh, T. (2010). VizWiz: Nearly Real-time Answers to Visual Questions. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology* (pp. 333–342). New York, NY, USA: ACM. http://doi.org/10.1145/1866029.1866080

Bigham, J. P., Kaminsky, R. S., Ladner, R. E., Danielsson, O. M., & Hempton, G. L. (2006). WebInSight:making web images accessible. In *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility* (Assets '06) (pp. 181–188). New York, NY, USA: ACM. http://doi.org/10.1145/1168987.1169018

Birbeck, M. (2009). Introduction to RDFa II. Available at http://alistapart.com/article/introduction-to-rdfa-ii [Accessed 9 May 2015].

Bizer, C., Eckert, K., Meusel, R., Mühleisen, H., Schuhmacher, M., & Völker, J. (2013). Deployment of RDFa, Microdata, and Microformats on the Web – A Quantitative Analysis. In *The Semantic Web–ISWC 2013* (pp. 17–32). Springer. Available at http://link.springer.com/chapter/10.1007/978-3-642-41338-4_2

Blackwell, A., & Engelhardt, Y. (2002). A Meta-Taxonomy for Diagram Research. En M. A. MSc, B. M. D. rer nat, & P. O. M. MSc (Eds.), *Diagrammatic Representation and Reasoning* (pp. 47-64). Springer London. Available at http://link.springer.com/chapter/10.1007/978-1-4471-0109-3_3

Book Industry Study Group. (2013). EPUBTest: Testsuite. Available at http://epubtest.org/testsuite/ [Accessed 9 May 2015].

Borrego, Á., & Frías, J. A. (2004). Más allá de la cantidad: la incorporación de los métodos cualitativos a la investigación en Información y Documentación. In Frías Montoya, J.A. y Ríos Hilario, A. B. (editores) *Metodología de investigación en Información y Documentación*. Salamanca: Ediciones Universidad de Salamanca, 2004, 191-211.

Braille Authority of North America. (2011). Guidelines and Standards for Tactile Graphics. Available at http://www.brailleauthority.org/tg/web-manual/index.html [Accessed 9 May 2015].

Brajnik, G. (2009). Barrier walkthrough. Available at http://users.dimi.uniud.it/~giorgio.brajnik/projects/bw/bw.html [Accessed 9 May 2015].

Brajnik, G. (2011). The troubled path of accessibility engineering: an overview of traps to avoid and hurdles to overcome. *SIGACCESS Access. Comput.*, (100), 1–11. http://doi.org/10.1145/1982572.1982573

Brewer, J. (2013). Accessibility Issues in Digital Publishing: DRM, Accessibility Content, and User Interface Control. Available at http://www.w3.org/2012/08/electronic-books/submissions/webooks2013_submission_41.html [Accessed 9 May 2015].

Brown, A., Jay, C., & Harper, S. (2010). Using qualitative eye-tracking data to inform audio presentation of dynamic Web content. *New Rev. Hypermedia Multimedia*, *16*(3), 281–301. http://doi.org/10.1080/13614568.2010.542253

Burford, B., Briggs, P., & Eakins, J. P. (2003). A Taxonomy of the Image: On the Classification of Content for Image Retrieval. *Visual Communication*, *2*(2), 123-161. http://doi.org/10.1177/1470357203002002001

Campbell, R., Pentz, E., & Borthwick, I. (2012). *Academic and Professional Publishing* (1 edition). Oxford: Chandos Publishing.

Capiel, G. (2014). Annotation as a Tool for Accessibility for Blind and Vision Impaired Students. Available at http://www.w3.org/2014/04/annotation/submissions/benetech_annotation_position.html [Accessed 9 May 2015].

Carbonell, A., Mazo, A., Serras, F., & Corominas, M. (2013). Ash2 acts as an ecdysone receptor coactivator by stabilizing the histone methyltransferase Trr. *Molecular Biology of the Cell*, *24*(3), 361-372. http://doi.org/10.1091/mbc.E12-04-0267

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in information visualization: using vision to think*. Morgan Kaufmann.

Cavender, A., Trewin, S., & Hanson, V. (2014). | Accessible Writing Guide. Available at http://www.sigaccess.org/welcome-to-sigaccess/resources/accessible-writing-guide/ [Accessed 9 May 2015].

CEN. (2008). *DPA (Document Processing for Accessibility) worKshop (2008). cen/ isss cwa 15778:2008*. Available at ftp://ftp.cenorm.be/PUBLIC/CWAs/DPA/accessible-CWA15778-2008-Feb.pdf [Accessed 9 May 2015].

Cerf, V. G. (2012). Why is Accessibility So Hard? *Commun. ACM*, *55*(11), 7–7. http://doi.org/10.1145/2366316.2366341

Channin, D. S., Mongkolwat, P., Kleper, V., Sepukar, K., & Rubin, D. L. (2010). The caBIG annotation and image Markup project. *Journal of Digital Imaging*, *23*(2), 217-225. http://doi.org/10.1007/s10278-009-9193-9

Chi, E. H., & Riedl, J. T. (2002). An operator interaction framework for visualization systems. In *Proceedings of the 1998 IEEE Symposium on Information Visualization* (INFOVIS '98) (pp. 63-70) IEEE Computer Society Washington, DC, USA.

Choudhury, S. R., Mitra, P., Kirk, A., Szep, S., Pellegrino, D., Jones, S., & Giles, C. L. (2013). Figure Metadata Extraction from Digital Documents. In *12th International Conference on Document Analysis and Recognition (ICDAR)* (pp. 135-139). Washington, DC, IEEE. http://doi.org/10.1109/ICDAR.2013.34

Christiansen, A., Lee, D. J., & Chang, Y. (2007). Finding relevant PDF medical journal articles by the content of their figures. In *Proceedings of SPIE, Medical Imaging: PACS and Imaging Informatics* (Vol. 6516, pp. 65160K–65160K). San Diego, CA. SPIE. http://doi.org/10.1117/12.709911

Ciancarini, P., Iorio, A. D., Nuzzolese, A. G., Peroni, S., & Vitali, F. (2013). Semantic Annotation of Scholarly Documents and Citations. In M. Baldoni, C. Baroglio, G. Boella, & R. Micalizio (Eds.), *AI*IA 2013: Advances in Artificial Intelligence* (pp. 336-347). Springer International Publishing. Available at http://link.springer.com/chapter/10.1007/978-3-319-03524-6_29

Clark, J. (2003). *Building accessible websites*. New Riders Pub. Available at http://joeclark.org/book/ [Accessed 9 May 2015].

Clark, L. (2002). *Liverpool Central Primary Care Trust Accessible Health Information: Project Report*. Liverpool Central Primary Care Trust, Liverpool. Available at http://disability-studies.leeds.ac.uk/files/library/Clark-Laurence-liverpool-NHS.pdf [Accessed 9 May 2015].

Codogno, P., Mehrpour, M., & Proikas-Cezanne, T. (2012). Canonical and non-canonical autophagy: variations on a common theme of self-eating? *Nature Reviews Molecular Cell Biology*, *13*(1), 7-12. http://doi.org/10.1038/nrm3249

Cohen, W. W., Wang, R., & Murphy, R. F. (2003). Understanding captions in biomedical publications. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 499–504). ACM. Available at http://dl.acm.org/citation.cfm?id=956809

Committee on Publication Ethics. (2015). Guidelines | Committee on Publication Ethics: COPE. Available at http://publicationethics.org/resources/guidelines [Accessed 9 May 2015].

Coonin, B. (2002). Establishing accessibility for e-journals: a suggested approach. *Library Hi Tech*, *20*(2), 207-220. http://doi.org/10.1108/07378830210432570

Costa, J. (1998). *La Esquematica: Visualizar La Información.* Paidos Iberica, S.A.

Creaser, C. (2013). Accessibility of top 1,000 titles of 2012 - Report. Available at http://www.rnib.org.uk/sites/default/files/Top_1000_books_2012.pdf [Accessed 9 May 2015].

Cryer. (2013). Teaching STEM subjects to blind and partially sighted learners - RNIB. Available at https://www.rnib.org.uk/knowledge-and-research-hub/research-reports/education-research/stem [Accessed 9 May 2015].

Cunningham, K. (2012). *Accessibility Handbook* (1 edition). Beijing: O'Reilly Media.

DAISY Consortium. (2008). DAISY 3 Structure Guidelines - Index of DTBook elements. Available at http://www.daisy.org/z3986/structure/SG-DAISY3/index-of-elements.html [Accessed 9 May 2015].

DAISY Consortium. (2014). Tobi - Image Description Manual. Available at http://www.daisy.org/tobi/image-description-manual#what_are_image_descriptions [Accessed 9 May 2015].

DAISY Consortium. (n. d.). Available at http://www.daisy.org/ [Accessed 9 May 2015].

Darrah, M. A. (2013). Computer Haptics: A New Way of Increasing Access and Understanding of Math and Science for Students Who are Blind and Visually Impaired. *Journal of Blindness Innovation and Research*, *3*(2). Available at https://nfb.org/images/nfb/publications/jbir/jbir13/jbir030202.html

Dasiopoulou, S., Giannakidou, E., Litos, G., Malasioti, P., & Kompatsiaris, Y. (2011). A Survey of Semantic Image and Video Annotation Tools. In G. Paliouras, C. D. Spyropoulos, & G. Tsatsaronis (Eds.), *Knowledge-Driven Multimedia Information Extraction and Ontology Evolution* (pp. 196-239). Springer Berlin Heidelberg. Available at http://link.springer.com/chapter/10.1007/978-3-642-20795-2_8

Del Valle. (2001). El Análisis documental de la fotografía. Available at http://www.ucm.es/info/multidoc/prof/fvalle/artfot.htm [Accessed 9 May 2015].

Demir, S., Oliver, D., Schwartz, E., Elzer, S., Carberry, S., Mccoy, K. F., & Chester, D. (2010). Interactive SIGHT: textual access to simple bar charts. *New Rev. Hypermedia Multimedia*, *16*(3), 245–279. http://doi.org/10.1080/13614568.2010.534186

Denzin, N. K., & Lincoln, Y. S. (2005). *The SAGE handbook of qualitative research*. Thousand Oaks: Sage Publications.

DIAGRAM Center. (2014a). Diagrammar: A Framework for Making Images and Graphics Accessible. Available at http://diagramcenter.org/standards-and-practices/content-model.html [Accessed 9 May 2015].

DIAGRAM Center. (2014b). Image Guidelines for EPUB 3. Available at http://diagramcenter.org/59-image-guidelines-for-epub-3.html [Accessed 9 May 2015].

DIAGRAM Content Working Group. (2014). Accessible Image Sample Book. Available at http://diagramcenter.org/standards-and-practices/accessible-image-sample-book.html [Accessed 9 May 2015].

Divoli, A., Wooldridge, M. A., & Hearst, M. A. (2010). Full Text and Figure Display Improves Bioscience Literature Search. *PLoS ONE*, *5*(4), e9619. http://doi.org/10.1371/journal.pone.0009619

DO-IT knowledge database. (2013). How long can an «alt» attribute be? Available at http://www.washington.edu/doit/articles?1257 [Accessed 9 May 2015].

Dublin Core Metadata Initiative. (2014). DCMI Metadata Terms. Available at http://dublincore.org/documents/dcmi-terms/ [Accessed 9 May 2015].

Duke, D. J. (2004). Drawing attention to meaning. *Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, *7*(6), 673-682. http://doi.org/10.1089/cpb.2004.7.673

Dunn, C., Rabren, K. S., Taylor, S. L., & Dotson, C. K. (2012). Assisting Students With High-Incidence Disabilities to Pursue Careers in Science, Technology, Engineering, and Mathematics. *Intervention in School and Clinic*, 48 (1), 47-54. http://doi.org/10.1177/1053451212443151

Eakins, J., & Graham, M. (1999). *Content-based Image Retrieval: A report to the JISC Technology Applications Programme*. Available at http://www.jisc.ac.uk/uploaded_documents/jtap-039.doc

Eakins, J. P. (2001). Retrieval of Still Images by Content. En M. Agosti, F. Crestani, & G. Pasi (Eds.), *Lectures on Information Retrieval*, 1980, 111-138. Berlin, Heidelberg: Springer

Berlin Heidelberg. Available at http://www.springerlink.com/content/5331yhf7hnbaacx7/

Elsevier Publishing. (2012). Author Artwork Instructions. Available at cdn.elsevier.com/assets/pdf_file/0010/109963/Artwork.pdf [Accessed 9 May 2015].

Elzer, S., Schwartz, E., Carberry, S., Chester, D., Demir, S., & Wu, P. (2008). Accessible bar charts for visually impaired users. In Proceedings of the IASTED International Conference on Telehealth/Assistive Technologies (pp. 55–60). Anaheim, CA, USA: ACTA Press. Available at http://dl.acm.org/citation.cfm?id=1722763.1722774

Engelhardt, Y. (2002). *The language of graphics: a framework for the analysis of syntax and meaning in maps, charts and diagrams*. (PhD thesis). Faculteit der Natuurwetenschappen, Wiskunde en Informatica, FNWI: Institute for Logic, Language and Computation (ILLC). Available at http://dare.uva.nl/record/105970 [Accessed 9 May 2015].

Enser, P. (2008). The evolution of visual information retrieval. *Journal of Information Science*, *34*, 531–546. http://doi.org/10.1177/0165551508091013

ETSI. (2014). New European Standard on accessibility requirements for public procurement of ICT products and services. Available at http://www.etsi.org/news-events/news/754-new-european-standard-on-accessibility-requirements-for-public-procurement-of-ict-products-and-services [Accessed 9 May 2015].

EUAIN (EuropEan accEssiBlE information nEtworK). (2008). The freedom to explore: outcomes of the euain project. Available at pdf.aminer.org/000/288/522/a_model_for_accessible_information_networks_findings_of_the_euain.pdf [Accessed 9 May 2015].

Facebook. (2014). Open Graph protocol. Available at http://ogp.me/ [Accessed 9 May 2015].

Faro, S., & Lecroq, T. (2013). The exact online string matching problem: A review of the most recent results. *ACM Computing Surveys*, *45*(2). Available at http://dl.acm.org.sire.ub.edu/citation.cfm?id=2431212

Fathulla, K., & Basden, A. (2007). Frameworks for understanding diagrams: enriching the debate. *Applied Semiotics Journal.* 8(19) 1204-6140. Available at http://nectar.northampton.ac.uk/388/

Faulkner, S. (2012). HTML5 Accessibility Chops: title attribute use and abuse. Available at http://www.paciellogroup.com/blog/2012/01/html5-accessibility-chops-title-attribute-use-and-abuse/ [Accessed 9 May 2015].

FECYT. (2013). Indicadores bibliométricos de la actividad científica española 2010. Available at http://icono.fecyt.es/informesypublicaciones/Documents/indicadores%20bibliometricos_web.pdf [Accessed 9 May 2015].

Fox, S. (2007). E-patients with a disability or chronic disease. Washington, DC: Pew Internet and American Life Project. Available at http://www.pewinternet.org/~/media/Files/Reports/2007/EPatients_Chronic_Conditions_2007.pdf.pdf [Accessed 9 May 2015].

Fox, S. (2008). The engaged e-patient population. *Washington, DC: Pew Internet & American Life Project*. Available at

http://nicolaziady.com/wp-content/uploads/2011/07/PIP_Health_Aug08.pdf.pdf [Accessed 9 May 2015].

Fredj, Z. B., & Duce, D. A. (2007). GraSSML: Accessible Smart Schematic Diagrams for All. *Univers. Access Inf. Soc.*, *6*(3), 233–247. http://doi.org/10.1007/s10209-007-0085-9

García, R. (2010). A Semantic Web Approach to Digital Rights Management. (PhD thesis). Department of Technologies, Universitat Pompeu Fabra,Barcelona, Spain. Available at http://rhizomik.net/html/~roberto/thesis/ AC [Accessed 9 May 2015].

Gardner, J. (2014). Can Mainstream Graphics be Accessible? *Information Technology and Disabilities Journal*, *XIV*(1). Available at http://itd.athenpro.org/volume14/number1/gardner.html

Gardner, J., & Bulatov, V. (2006). Scientific Diagrams Made Easy with IVEOTM. En K. Miesenberger, J. Klaus, W. L. Zagler, & A. I. Karshmer (Eds.), *Computers Helping People with Special Needs* (pp. 1243-1250). Springer Berlin Heidelberg. Available at http://link.springer.com/chapter/10.1007/11788713_179

Gardner, J., Bulatov, V., & Kelly, R. (2009). Making journals accessible to the visually impaired: the future is near. *Learned Publishing*, *22*(4), 314-319. http://doi.org/10.1087/20090408

Garrish, M. (2014). EPUB 3 Accessibility Guidelines. Available at http://www.idpf.org/accessibility/guidelines/ [Accessed 9 May 2015].

Glyph & Cog. (2014). Xpdf: A PDF Viewer for X. Available at http://www.foolabs.com/xpdf/home.html [Accessed 9 May 2015].

Goede, P. A., Lauman, J. R., Cochella, C., Katzman, G. L., Morton, D. A., & Albertine, K. H. (2004). A Methodology and Implementation for Annotating Digital Images for Context-appropriate Use in an Academic Health Care Environment. *Journal of the American Medical Informatics Association : JAMIA*, *11*(1), 29-41. http://doi.org/10.1197/jamia.M1247

Goncu, C., Marriott, K., & Hurst, J. (2010). Usability of Accessible Bar Charts. In A. K. Goel, M. Jamnik, & N. H. Narayanan (Eds.), *Diagrammatic Representation and Inference* (pp. 167-181). Springer Berlin Heidelberg. Available at http://link.springer.com/chapter/10.1007/978-3-642-14600-8_17

Google, Inc., Yahoo, Inc., Microsoft Corporation, & Yandex. (2011). schema.org. Available at http://schema.org/ [Accessed 9 May 2015].

Graham, R. N. J., Perriss, R. W., & Scarsbrook, A. F. (2005). DICOM demystified: a review of digital file formats and their use in radiological practice. *Clinical radiology*, *60*(11), 1133-1140. http://doi.org/10.1016/j.crad.2005.07.003

Grammel, L., Tory, M., & Storey, M. A. (2010). How Information Visualization Novices Construct Visualizations. *IEEE Trans. Visualization and Computer Graphics*, *16*(6). http://doi.org/10.1109/TVCG.2010.164

Groves, K. (2013). Playing with Longdesc. Available at http://karlgroves-sandbox.com/longdesc.php [Accessed 9 April 2015].

Groves, K. (2014). Longdesc - Where are the alternatives? Available at http://www.karlgroves.com/2014/08/26/longdesc-where-are-the-alternatives/ [Accessed 9 May 2015].

Gruber, T. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*, 199-220. http://doi.org/10.1006/knac.1993.1008

Hagen, H. (2010). Tagged PDF in ConTEXt. TUGboat, 31(3), 197-202. Available at http://www.tug.org/TUGboat/tb31-3/tb99hagen.pdf [Accessed 9 May 2015].

Hames. (2007). *Peer Review and Manuscript Management in Scientific Journals: Guidelines for Good Practice*. Malden, MA ; Oxford: Wiley.

Hartley, J., Cabanac, G., Kozak, M., & Hubert, G. (2014). Research on tables and graphs in academic articles: Pitfalls and promises. *Journal of the Association for Information Science and Technology*, 66 (2), 428–431. http://doi.org/10.1002/asi.23208

Hearst, M. (2003). Information visualization: Principles, promise, and pragmatics. Handouts of the tutorial at CHI 2003 Conference on Human Factors in Computing Systems.

Henry, S. L. (2007). *Just Ask: Integrating Accessibility Throughout Design*. Available at http://uiaccess.com/accessucd/ [Accessed 9 May 2015].

Herman, I. (2009). W3C Semantic Web FAQ. Available at http://www.w3.org/2001/sw/SW-FAQ#swgoals [Accessed 9 May 2015].

Hersh, W., Müller, H., & Kalpathy-Cramer, J. (2009). The ImageCLEFmed medical image retrieval task test collection. *Journal of Digital Imaging: The Official Journal of the Society for Computer Applications in Radiology*, *22*(6), 648-655. http://doi.org/10.1007/s10278-008-9154-8

Hilderley, S. (2013). Accessible Publishing – Best Practice Guidelines for Publishers. Available at http://www.accessiblebooksconsortium.org/inclusive_publishing/en/accessible_best_practice_guidelines_for_publishers.html [Accessed 9 May 2015].

Hojas, Á. O. (2008). *Instrucciones y criterios para la producción de libros DAISY*. Organización Nacional de Ciegos Españoles, Madrid, Spain.

Hollink, L., Schreiber, G., Wielinga, B., & Worring, M. (2004). Classification of user image descriptions. *International Journal of Human Computer Studies*, *61*, 601-626. http://doi.org/10.1016/j.ijhcs.2004.03.002

Horton, S., & Sloan, D. (2014). Accessibility in Practice: A process-driven approach to accessibility. In *Inclusive Designing: Joining Usability, Accessibility, and Inclusion* (Langdon, P.M., Lazar, J., Heylighen, A., Dong, H. (Eds.), pp. 105 - 118). Springer. Available at http://blog.paciellogroup.com/2014/03/accessibility-practice-process-driven-approach-accessibility/

Hyland, K., & Tse, P. (2007). Is There an «Academic Vocabulary»? *TESOL quarterly* 41(2), 235-53. http://doi.org/10.1002/j.1545-7249.2007.tb00058.x.

Iezzoni, L. I., & O'Day, B. L. (2005). *More than Ramps: A Guide to Improving Health Care Quality and Access for People with Disabilities*. Oxford University Press.

International Committee of Medical Journal Editors. (2014). Preparing for Submission. Available at http://www.icmje.org/recommendations/browse/manuscript-preparation/preparing-for-submission.html [Accessed 9 May 2015].

International Digital Publishing Forum. (2014a). EPUB 3 Accessibility Guidelines - Figures. Available at http://www.idpf.org/accessibility/guidelines/content/xhtml/figures.php [Accessed 9 May 2015].

International Digital Publishing Forum. (2014b). EPUB3 Overview Recommended Specification 26 June 2014. Available at http://www.idpf.org/epub/301/spec/epub-overview.html [Accessed 9 May 2015].

International Publishers Association. (2013). International Publishers Association endorses EPUB 3 as a global standard. Available at http://idpf.org/ipa_endorses_epub_3 [Accessed 9 May 2015].

IRRI photos. (2007). *Rice Genetics III_p789a* [Photo]. Available at https://www.flickr.com/photos/ricephotos/6943707382/ [Accessed 9 May 2015].

ISO. (2008). ISO 9241-171:2008 - Ergonomics of human-system interaction -- Part 171: Guidance on software accessibility. Available at http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=39080 [Accessed 9 May 2015].

ISO. (2011) ISO 9999:2011 - Assistive products for persons with disability -- Classification and terminology. Available at http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=50982 [Accessed 9 May 2015].

ISO. (2012). PDF/UA-1 (ISO 14289-1:2012) - Document management applications -- Electronic document file format enhancement for accessibility -- Part 1: Use of ISO 32000-1 (PDF/UA-1). Available at http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?ics1=35&ics2=240&ics3=30&csnumber=54564 [Accessed 9 May 2015].

ISO. (2013) ISO 32000-1:2008 Document management - Portable document format - Part 1: PDF 1.7. Available at http://www.iso.org/iso/catalogue_detail.htm?csnumber=51502 [Accessed 9 May 2015].

ISO/IEC. (2008). ISO 9241-171:2008 - ISO/IEC TS 20071-11:2012 - Information technology -- User interface component accessibility -- Part 11: Guidance for alternative text for images. Available at http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=59423 [Accessed 9 May 2015].

Jackson, G. W., Davidson, H. C., Wiggins, R. H., 3rd, & Harnsberger, H. R. (2001). Electronic submission of academic works: a survey of current editorial practices of radiologic journals. *Journal of digital imaging: the official journal of the Society for Computer Applications in Radiology*, *14*(2). http://doi.org/10.1007/s10278-001-0008-x

Jaimes, A., & Chang, S. (1999). A conceptual framework for indexing visual information at multiple levels. In *Proceedings of the SPIE- IS&T Internet Imaging,* 3964, 2–15. San Jose, CA: SPIE. http://doi.org/10.1117/12.373443

JISC Digital Media. (2014a). Choosing a File Format for Digital Still Images. Available at http://www.jiscdigitalmedia.ac.uk/guide/choosing-a-file-format-for-digital-still-images [Accessed 13 January 2015].

JISC Digital Media. (2014b). Guide - File Formats and Compression. Available at http://www.jiscdigitalmedia.ac.uk/guide/file-formats-and-compression/ [Accessed 13 January 2015].

JISC Digital Media. (2014c). Introduction to metadata. Available at http://www.jiscdigitalmedia.ac.uk/guide/an-introduction-to-metadata [Accessed 13 January 2015].

JISC Digital Media. (2014d). Approaches to describiing images. Available at http://www.jiscdigitalmedia.ac.uk/guide/approaches-to-describing-images [Accessed 13 January 2015].

JISC Digital Media. (s. f.). Accessible images and how to create them. Available at http://vle.jisctechdis.ac.uk/xerte/play_395 [Accessed 13 January 2015].

Jones, M. G., & Broadwell, B. (2008). Visualization Without Vision: Students with Visual. In E. J. K. Gilbert, J. K. Gilbert, M. Reiner, & M. Nakhleh (Eds.), *Visualization: Theory and Practice in Science Education* (Vol. 3, pp. 283-294). Springer Netherlands. Available at http://dx.doi.org/10.1007/978-1-4020-5267-5_12

Julesmcn. (2006). *Hailey disease*. Available at http://commons.wikimedia.org/wiki/File:HHD_R_Axilla.jpg [Accessed 9 May 2015].

Kahn, C. E., Jr, Kalpathy-Cramer, J., Lam, C. A., & Eldredge, C. E. (2012). Accurate determination of imaging modality using an ensemble of text- and image-based classifiers. *Journal of digital imaging: the official journal of the Society for Computer Applications in Radiology*, *25*(1), 37-42. http://doi.org/10.1007/s10278-011-9399-5

Karshmer, A., Gupta, G., & Pontelli, E. (2007). Mathematics and accessibility: a survey. In *Proc. 9th International Conference on Computers Helping People with Special Needs* (3118, 664–669). Available at http://pdf.aminer.org/000/288/497/remathex_reader_and_editor_of_the_mathematical_expressions_for_blind.pdf

Kawanaka, S., Kobayashi, M., Takagi, H., & Asakawa, C. (2009). Accessibility commons: a metadata repository for web accessibility. *ACM SIGWEB Newsletter*, 1:1–1:7. http://doi.org/10.1145/1514984.1514985

Keane, K., & Laverentz, C. (2014). Interactive Scientific Graphics - Recommended Practices for Verbal Description. Available at diagramcenter.org/wp-content/uploads/2014/06/Interactive_Scientific_Graphics_Submission23june2014.docx [Accessed 9 May 2015].

Kelly, B. (2011). Web Accessibility, Institutional Repositories and BS 8878. Available at http://ukwebfocus.wordpress.com/2011/01/24/web-accessibility-institutional-repositories-and-bs-8878/ [Accessed 9 May 2015].

Kelly, B., Nevile, L., Sloan, D., Fanou, S., Ellison, R., & Herrod, L. (2009). From Web accessibility to Web adaptability. *Disability and Rehabilitation: Assistive Technology*, 4(4), 212-226. http://doi.org/10.1080/17483100902903408

Kemp, J. (2007). Obsteric ultrasonograph of a healthy female at 16 weeks. In *Wikipedia, the free encyclopedia*. Available at http://en.wikipedia.org/wiki/File:Obsteric_ultrasonograph.jpg [Accessed 9 May 2015].

Keysers, D., Renn, M., & Breuel, T. M. (2007). Improving accessibility of html documents bygenerating image-tags in a proxy. In Assets '07 (Assets '07), 249–250. New York, NY, USA: ACM. http://doi.org/10.1145/1296843.1296896

Kim. (2010). Caption-based topical descriptors for microscopic images as published in academic papers, *Health Information & Libraries Journal,* 27(3): 235–243. http://doi.org/10.1111/j.1471-1842.2010.00897.x

Kopeček, I., & Ošlejšek, R. (2008). Dialogue-Based Processing of Graphics and Graphical Ontologies. In P. Sojka, A. Horák, I. Kopeček, & K. Pala (Eds.), *Text, Speech and Dialogue*, 5246, 601-608. Berlin, Heidelberg: Springer Berlin Heidelberg. Available at http://www.springerlink.com/content/63882j8k73up4241/

Korpela, J. (2012). Guidelines on using ALT texts in IMG elements. Available at http://www.cs.tut.fi/~jkorpela/html/alt.html [Accessed 9 May 2015].

LaBerge, J. M., & Andriole, K. P. (2003a). Digital Image Processing: A Primer for JVIR Authors and Readers Part 1: The Fundamentals. *Journal of Vascular and Interventional Radiology*, 14(10),1223-3. Available at http://www.ncbi.nlm.nih.gov/pubmed/14551267

LaBerge, J. M., & Andriole, K. P. (2003b). Digital image processing: a primer for JVIR authors and readers: part 2: digital image acquisition. *Journal of Vascular and Interventional Radiology*, *14*(11), 1365-1372. Available at http://doi.org/http://www.ncbi.nlm.nih.gov/pubmed/14605101

Lacey, A., & Luff, D. (2001). *Qualitative data analysis*. Trent Focus Sheffield. Available at http://www.rds-yh.nihr.ac.uk/wp-content/uploads/2013/05/9_Qualitative_Data_Analysis_Revision_2009.pdf [Accessed 9 May 2015].

Lacy, L. W. (2005). *OWL: representing information using the web ontology language*. Trafford Publishing.

Landau, S. (2013). An Interactive web-based tool for sorting textbook images prior to adaptation to accessible format: Year 1 Final Report. Available at diagramcenter.org/images/decisiontree_finalreport.doc [Accessed 9 May 2015].

Lee, J. H. (2014). Handling digital images for publication. *Science Editing*, *1*(2), 58–61. http://dx.doi.org/10.6087/kcse.2014.1.58

Lévesque, V. (2005). Blindness, technology and haptics. Center for Intelligent Machines. Technical Report.

Levine, D. (2010). How to Obtain Images from Picture Archiving and Communication Systems and Ready Them for Publication. *Radiology*, *257*(3), 603-608. http://doi.org/10.1148/radiol.10091423

Lewis, P., Noble, S., & Soiffer, N. (2010). Using accessible math textbooks with students who have learning disabilities. In Proceedings *of the 12th international ACM SIGACCESS conference on Computers and accessibility*, 139–146. New York, NY, USA: ACM. http://doi.org/10.1145/1878803.1878829

Liang, H., Xue, Y., & Chase, S. K. (2011). Online health information seeking by people with physical disabilities due to neurological conditions. *International Journal of Medical Informatics*, *80*(11), 745-753. http://doi.org/10.1016/j.ijmedinf.2011.08.003

Logicmighty Ltd. (2014). Logicmighty free software. Available at http://www.logicmighty.co.uk/Home/FreeDownloads.aspx [Accessed 9 May 2015].

Lohse, G. L., Biolsi, K., Walker, N., & Rueter, H. H. (1994). A classification of visual representations. *Communications of the ACM*, *37*(12), 36-49. http://doi.org/10.1145/198366.198376

Lord, P., Cockell, S., & Stevens, R. (2012). Three Steps to Heaven: Semantic Publishing in a Real World Workflow. *Future Internet*, *4*(4), 1004-1015. http://doi.org/10.3390/fi4041004

Lorenz, M., & Horstmann, M. (2004). Semantic access to graphical web resources for blind user. In *3rd International Semantic Web Conference (ISWC2004)*. Available at http://www.ei.sanken.osaka-u.ac.jp/iswc2004/posters/PID-YIHSOXCY-1090249595.pdf

Macnamara, J. R. (2005). Media content analysis: its uses, benefits and best practice methodology. *Asia-Pacific Public Relations Journal*, 6 (1), 1-34. Available at http://search.informit.com.au/documentSummary;dn=200705762;res=IELAPA

Marshall, C. C. (2009). *Reading and Writing the Electronic Book*. San Rafael, Calif.: Morgan and Claypool Publishers.

Massachusetts Medical Society. (2015). Author Center - Technical Guidelines. Available at http://www.nejm.org/page/author-center/technical-guidelines [Accessed 9 May 2015].

Maus, J. (2012). Maximum Intensity Projection (MIP) of a wholebody positron emission tomography (PET) acquisition. En *Wikipedia, the free encyclopedia*. Available at http://en.wikipedia.org/w/index.php?title=File:PET-MIPS-anim.gif&oldid=500509457 [Accessed 9 May 2015].

Metadata Working Group. (2010). Guidelines For Handling Image Metadata (v2.0). Available at www.metadataworkinggroup.org/pdf/mwg_guidance.pdf [Accessed 9 May 2015].

Michie, S., Stralen, M. M. van, & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, *6*(1), 42. http://doi.org/10.1186/1748-5908-6-42

Microsoft. (2003). The Market for Accessible Technology. Available at http://www.microsoft.com/enable/research/phase1.aspx [Accessed 9 May 2015].

Microsoft. (2015). How to change the export resolution of a PowerPoint slide. Available at http://support.microsoft.com/kb/827745/en-us [Accessed 9 May 2015].

Milowski, A. (2012). Using RDFa to Annotate Images. Available at http://www.milowski.com/journal/entry/2012-05-24T09:17:12.327-07:00/ [Accessed 9 May 2015].

Mirri, S., Salomoni, P., & Prandi, C. (2011). Augment Browsing and Standard Profiling for Enhancing Web Accessibility. In Proceedings *of the International Cross-Disciplinary Conference on Web Accessibility*, 5:1–5:10. New York, NY, USA: ACM. http://doi.org/10.1145/1969289.1969297

Moon, N.W., R. L. T. N. W. (2011). Accommodating Students with Disabilities in Science, Technology, Engineering, and Mathematics (STEM): Findings from Research and Practice for Middle Grades through University Education. Atlanta: Center for Assistive Technology and Environmental Access, College of Architecture, Georgia Institute of Technology. Available at http://www.catea.gatech.edu/scitrain/accommodating.pdf [Accessed 9 May 2015].

Moore, R. (2009). Ongoing Efforts to Generate «Tagged PDF» using pdfTeX. *Towards a Digital Mathematics Library.* Grand Bend, Ontario, Canada, 125-131. Available at http://dml.cz/bitstream/handle/10338.dmlcz/702553/DML_002-2009-1_14.pdf

Morris, S., Barnas, E., LaFrenier, D., & Reich, M. (2013). *The Handbook of Journal Publishing*. Cambridge University Press.

Müller, H., Kalpathy-Cramer, J., Demner-Fushman, D., Antani, S., & García Seco, A. (2012). Medical Image Classification and Retrieval 2012. Available at http://www.imageclef.org/2012/medical [Accessed 9 May 2015].

Murphy, S. (2005). Accessibility of graphics in technical documentation for the cognitive and visually impaired. In Proceedings of the 23rd annual international conference on Design of communication: documenting & designing for pervasive information (SIGDOC '05), 12–17. New York, NY, USA: ACM. http://doi.org/10.1145/1085313.1085320

Myers, C., & Capiel, G. (2014). Practical Properties Guide - Accessibility Metadata Project. Available at https://wiki.benetech.org/display/a11ymetadata/Practical+Properties+Guide#PracticalPropertiesGuide-HTMLDocumentExample [Accessed 9 May 2015].

National Center on Accessible Instructional Materials. (2006). NIMAS Technical Specification v1.1. Available at http://aim.cast.org/experience/technologies/spec-v1_1 [Accessed 9 May 2015].

NISO. (2005) ANSI/NISO Z39.86-2005 - Specifications for the Digital Talking Book. Available at http://www.niso.org/workrooms/daisy/Z39-86-2005.html [Accessed 9 April 2015].

NISO. (2012). ANSI/NISO Z39.98-2012 - Authoring and Interchange Framework for Adaptive XML Publishing Specification. Available at http://daisy.niso.org/standards/z39-98-2012.html [Accessed 9 May 2015].

NISO. (2015). ANSI/NISO Z39.14-1997 (R2009) Guidelines for Abstracts. Available at http://www.niso.org/apps/group_public/project/details.php?project_id=32 [Accessed 9 May 2015].

National Institute of Standards and Technology. (2004). SI Unit rules and style conventions checklist. Available at http://physics.nist.gov/cuu/Units/checklist.html [Accessed 9 May 2015].

Nature Publishing Group. (2012). How to submit : Nature Genetics. Available at http://www.nature.com/ng/authors/submit/index.html [Accessed 9 May 2015].

Nature Publishing Group. (2014). Manuscript Formatting Guide For authors. Available at http://www.nature.com/nature/authors/gta/#a5.8 [Accessed 9 May 2015].

Nature Publishing Group. (2015). Nature.com - Accessibility Statement. Available at http://www.nature.com/info/accessibility_statement.html

NCAM. (2009a). Effective Practices for Description of Science Content within Digital Talking Books. Available at http://www.nature.com/info/accessibility_statement.html [Accessed 9 May 2015].

NCAM. (2009b). NCAM Publishes Guidelines for Describing Science, Technology, Engineering and Math (STEM) Images. Available at http://ncam.wgbh.org/about/news/ncam-publishes-guidelines-for- [Accessed 9 May 2015].

NCAM. (2011). DIAGRAM Center Working Paper – April 2011. Potential Use of Image Description Metadata for Accessibility. Available at http://diagramcenter.org/images/documents/ncam%20diagram%20image%20metadata%20paperapril2011.pdf [Accessed 9 May 2015].

NCAM. (2012). DIAGRAM CENTER WORKING PAPER – Image Metadata Addendum, April, 2012. Available at http://diagramcenter.org/images/documents/image_metadata_paper-addendum.pdf [Accessed 9 May 2015].

NCBI. (n. d.). Make It Clear Guidelines. Available at http://www.ncbi.ie/services/services-for-organisations/making-print-and-multimedia-accessible-mcs/make-it-clear [Accessed 9 May 2015].

NEMA. (2014). DICOM PS3.10 2014b - Media Storage and File Format for Media Interchange - DICOM Standards Committee. Available at http://medical.nema.org/medical/dicom/current/output/html/part10.html [Accessed 9 May 2015].

Niedderer, K., Mackrill, J., Clune, S., Lockton, D., Ludden, G., Morris, A., … Hekkert, P. (2014). *Creating Sustainable Innovation through Design for Behaviour Change: Full Project Report*. University of Wolverhampton, CADRE. Available at http://wlv.openrepository.com/wlv/handle/2436/336632 [Accessed 9 May 2015].

Okabe, M., & Ito, K. (2008). Color Universal Design (CUD) - How to make figures and presentations that are friendly to Colorblind people -. Available at http://jfly.iam.u-tokyo.ac.jp/color/ [Accessed 9 May 2015].

Open University. (2013). Description Guidelines (CC). Available at http://kn.open.ac.uk/public/workspace.cfm?wpid=3881 [Accessed 9 May 2015].

Palacios, A., & Romañach, J. (2007). El modelo de la diversidad. La Bioética y los Derechos Humanos como herramientas para alcanzar la plena dignidad en la diversidad funcional. *Derechos y libertades: Revista del Instituto Bartolomé de las Casas*, *11*(17), 231-239.

Panofsky, E. (1993). *Meaning in the visual art*. London: Penguin.

Partin, B. (n. d.). Public Health Image Library (PHIL). Available at http://phil.cdc.gov/phil/details.asp

PDF Association. (2012). NVDA goes PDF/UA! The PDF Association Steps Up to Fund Development of the World's First PDF/UA Conforming Screen Reader. Available at http://www.pdfa.org/2012/06/nvda-goes-pdfua-the-pdf-association-steps-up-to-fund-development-of-the-world%E2%80%99s-first-pdfua-conforming-screen-reader/

PDF Association. (2013). Assistive technology for in-depth testing. Available at http://www.pdfa.org/2013/06/assistive-technology-for-in-depth-testing/ [Accessed 9 May 2015].

Pemberton, S. (2003). The Kiss of the Spiderbot. *interactions*, *10*(1), 44–. http://doi.org/10.1145/604575.604594

Peňáz, P., & Ondra, S.. (2012). E-Simplicity versus e-Complexity: Some Remarks on the Perspectives of the Universal Design for Learning. In *Proceedings of The International Workshop on Digitization and E-Inclusion in Mathematics and Science* (DEIMS12), 1-18, Tokyo: Nihon University. Available at http://www.muni.cz/research/publications/1102625

Perera, C. (2012). The Evolution of E-Health – Mobile Technology and mHealth. *Journal of Mobile Technology in Medicine*, *1*(1), 1-2. http://doi.org/10.7309/jmtm.1

Peres, S. C., Best, V., Brock, D., Shinn-Cunningham, B., Frauenberger, C., Hermann, T., … Stockman, T. (2008). Chapter 5 - Auditory Interfaces. In *HCI Beyond the GUI*, 147-195. San Francisco: Morgan Kaufmann. Available at http://www.sciencedirect.com/science/article/pii/B9780123740175000055

Petrie, H., Harrison, C., & Dev, S. (2005). Describing images on the web: a survey of current practice and prospects for the future. *Proceedings of Human Computer Interaction International (HCII)*, Las Vegas, Nevada. Available at http://www.users.cs.york.ac.uk/~petrie/HCII05_alt_text_Paper.pdf

Petrie, H., & Kheir, O. (2007). The relationship between accessibility and usability of websites (pp. 397–406). New York, NY, USA: ACM. http://doi.org/10.1145/1240624.1240688

Pinto Molina, M. (2002). *Indización y resumen de documentos digitales y multimedia : técnicas y procedimientos*. Gijón: Trea.

Pollock, J. T. (2009). *Semantic Web For Dummies* (1 edition). Hoboken, N.J: Wiley.

Power, C., & Jürgensen, H. (2010). Accessible presentation of information for people with visual disabilities. *Universal Access in the Information Society*, *9*(2), 97–119. http://doi.org/10.1007/s10209-009-0164-1

Public Health Image Library. (n. d.). Computed tomography scan of lungs showing classic snowstorm appearance of acute histoplasmosis. Available at http://phil.cdc.gov/phil/details.asp?pid=472

Rafkind, B., Lee, M., Chang, S.-F., & Yu, H. (2006). Exploring text and image features to classify images in bioscience literature. In Proceedings *of the Workshop on Linking Natural Language Processing and Biology: Towards Deeper Biological Literature Analysis,* 73–80. Stroudsburg, PA, USA: Association for Computational Linguistics. Available at http://dl.acm.org/citation.cfm?id=1567619.1567632

Ranaldi, F., & Nisbet, P. (2010). Guide for Accessible textual resources - Accessible Text: Guidelines for Good Practice. Available at: http://www.callscotland.org.uk/Resources/Books/Accessible-Text/ [Accessed 9 May 2015].

Real Decreto Legislativo. (2013) Real Decreto-ley 1/2013, de 29 de noviembre, por el que se aprueba el Texto Refundido de la Ley General de derechos de las personas con discapacidad y de su inclusión social. *Boletín Oficial del Estado*, 3 de diciembre de 2013, núm. 289, pp. 95635-95673.

Reitz, J. M. (2004). *Dictionary for library and information science*. Libraries Unlimited. Available at http://www.abc-clio.com/ODLIS/odlis_A.aspx [Accessed 9 May 2015].

Rello, L., & Baeza-Yates, R. (2013). Good fonts for dyslexia. In Proceedings *of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, 14. Available at http://dl.acm.org/citation.cfm?id=2513447

Renz, H., Brandtzaeg, P., & Hornef, M. (2012). The impact of perinatal immune development on mucosal homeostasis and chronic inflammation. *Nature Reviews. Immunology*, *12*(1), 9-23. http://doi.org/10.1038/nri3112

Ribera, M. (2008). Are PDF Documents Accessible? *Information Technology and Libraries*, *27*(3), 25-43. http://doi.org/10.6017/ital.v27i3.3246

Ribera, M. (2010a). *Publicaciones científicas accesibles para personas ciegas y deficientes visuales*. Madrid: Organización Nacional de Ciegos Españoles, Asesoría de Servicios Sociales para Afiliados.

Ribera, M. (2010b). La accesibilidad en los repositorios. *4as Jornadas OS Repositorios* (Barcelona, 3-5 marzo 2010). Available at http://openaccess.uoc.edu/webapps/o2/handle/10609/1726?mode=full [Accessed 9 May 2015].

Ribera, M., & Splendiani, B. (2014). Could semantic web and accessibility be BFF (best friends for ever) in image annotation? In *Footnotes, comments, bookmarks, and marginalia on the Web. A W3C Workshop on Annotations*, San Francisco, California. Available at www.w3.org/2014/04/annotation/submissions/W3C-Workshop_on_Annotations-position_paper_UB.pdf [Accessed 9 May 2015].

RNIB. (2010a). ImageShare Accessible image acceptance criteria. Available at https://load2learn.org.uk/training/imagetrain/images_guidelines/guidelines_docs/Image_acceptance_guidelines_v7.doc [Accessed 9 May 2015].

RNIB. (2010b). ImageShare complexity guidelines. Available at https://load2learn.org.uk/training/imagetrain/images_guidelines/image_complexity/complexity_intro.php [Accessed 5 December 2014].

RNIB. (2013). Creating accessible eBooks. Available at https://www.rnib.org.uk/professionals/publishing/publishertechnical/ebooksaccessible/Pages/accessible_ebook_creation.aspx [Accessed 5 December 2014]

RNIB, & Dyslexia Action. (2014). Browse resources - Images. Available at https://load2learn.org.uk/hierarchy/browse/2/139/Images [Accessed 9 May 2015].

RNIB. (n. d.). Clear Print Guidelines from the Royal National Institute for Blind People (RNIB) Available at http://www.psncorp.com/Downloads/RNIB_Clear_Print_Guidelines.pdf [Accessed 9 May 2015].

Rolandi, M., Cheng, K., & Pérez-Kriz, S. (2011). A Brief Guide to Designing Effective Figures for the Scientific Paper. *Advanced Materials*, *23*(38), 4343–4346. http://doi.org/10.1002/adma.201102518

Ronallo, J. (2012). HTML5 Microdata and Schema.org. *The Code4Lib Journal*, (16). Available at http://journal.code4lib.org/articles/6400 [Accessed 9 May 2015].

Roth, P., Richoz, D., Petrucci, L., & Pun, T. (2001). An audio-haptic tool for non-visual image representation (Vol. 1, pp. 64-67 vol.1). In *Signal Processing and its Applications, Sixth International Symposium*, IEEE. http://doi.org/10.1109/ISSPA.2001.949776

Rubin, J., & Chisnell, D. (2011). *Handbook of Usability Testing: Howto Plan, Design, and Conduct Effective Tests*. John Wiley & Sons.

Sacco, J., Corona, G., & Johannesen, L. (2004). Out of the Blue: Usability Testing at IBM. *AFB AccessWorld Magazine, 5* (5). Available at https://www.afb.org/afbpress/Pub.asp?DocID=aw050507 [Accessed 9 May 2015].

Schalitz, B. (2007). Das accessibility-Paket. Available at http://www.babs.gmxhome.de/download/da_pdftex/dok_pdf.pdf [Accessed 9 May 2015].

Schepers, D. A. (2013). Current State of Authoring Accessible SVG « Reinventing Fire. Available at http://schepers.cc/authoring-accessible-svg [Accessed 9 May 2015].

Science. (2012). Preparing Efficient Figures for Initial Submission. Available at http://www.sciencemag.org/site/feature/contribinfo/prep/prep_subfigs.xhtml [Accessed 9 May 2015].

Sedghi, S., Sanderson, M., & Clough, P. (2008). A study on the relevance criteria for medical images. *Pattern Recogn. Lett.*, *29*(15), 2046–2057. http://doi.org/10.1016/j.patrec.2008.07.003

Sedghi, S., Sanderson, M., & Clough, P. (2011). Medical image resources used by health care professionals. *Aslib Proceedings*, *63*(6), 570-585. http://doi.org/10.1108/00012531111187225

Sedghi, S., Sanderson, M., & Clough, P. (2012). How do health care professionals select medical images they need? *Aslib Proceedings*, *64*(4), 437-456. http://doi.org/10.1108/00012531211244815

Segura, B., Ibarretxe-Bilbao, N., Sala-Llonch, R., Baggio, H. C., Martí, M. J., Valldeoriola, F., … Junqué, C. (2013). Progressive changes in a recognition memory network in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *84*(4), 370-378. http://doi.org/10.1136/jnnp-2012-302822

Shatford Layne, S. (1986). Analyzing the Subject of a Picture: A Theoretical Approach. *Cataloging & Classification Quarterly*, *6*(3), 39-62. http://doi.org/10.1300/J104v06n03_04

Shawn Lawton, H., & Duffy, M. (2005). Essential Components of Web Accessibility. Available at http://www.w3.org/WAI/intro/components.php [Accessed 9 May 2015].

Siegel, R., Naishadham, D., & Jemal, A. (2012). Cancer statistics, 2012. *CA: A Cancer Journal for Clinicians*, *62*(1), 10-29. http://doi.org/10.3322/caac.20138

Siu, Y.T. (2014). 3D Printing for Accessible Materials in Schools – Final Report. Available at http://diagramcenter.org/wp-content/uploads/2014/06/3D_FinalReport_SIU_3.docx [Accessed 9 May 2015].

Sloan, D., & Kelly, B. (2011). Web accessibility metrics for a post digital world. En *RDWG Symposium on Website Accessibility Metrics*. University of Bath. Available at http://www.w3.org/WAI/RD/2011/metrics/ [Accessed 9 May 2015].

Smeureanu, I., & Isaila, N. (2011). Aspects of Mathematics Learning Objects Creation for Persons with Visual Disabilities. *Economy Informatics*, *1*(11), 34–44. Available at https://www.researchgate.net/publication/268015858_Aspects_of_Mathematics_Learning_Objects_Creation_for_Persons_with_Visual_Disabilities [Accessed 9 May 2015].

Smith, R. (2006). The Trouble with Medical Journals. *JRSM*, *99*(3), 115-119. http://doi.org/10.1258/jrsm.99.3.115

Social Security Administration. (2010). Social Security Administration Guide: Alternate text for images. Available at www.cio.gov/documents/alttextguide_1.1.pdf [Accessed 9 May 2015].

Software Mechanics Pty Ltd. (2015). SVGmaker. Available at http://www.svgmaker.com/ [Accessed 9 May 2015].

Sojka, P., Ruzicka, M., Kucbel, M., & Jarmar, M. (2013). Accessibility Issues in Digital Mathematical Libraries. In *Proceedings of the Conference Universal Learning Design*. 89 -97, 9. Brno: Masaryk University. Available at http://www.fi.muni.cz/usr/sojka/presentations/sojka-brno-uld-pres2011.pdf [Accessed 9 May 2015].

Soriano-Raya, J. J., Miralbell, J., López-Cancio, E., Bargalló, N., Arenillas, J. F., Barrios, M., … Mataró, M. (2014). Tract-specific fractional anisotropy predicts cognitive outcome in a community sample of middle-aged participants with white matter lesions. *Journal of Cerebral Blood Flow & Metabolism*. http://doi.org/10.1038/jcbfm.2014.26

Southwell, K. L., & Slater, J. (2012). Accessibility of digital special collections using screen readers. *Library Hi Tech*, *30*(3), 457-471. http://doi.org/10.1108/07378831211266609

Splendiani, B., & Ribera, M. (2014a). Accessible Images in Computer Science Journals. *Procedia Computer Science*, *27*, 9-18. http://doi.org/10.1016/j.procs.2014.02.003

Splendiani, B., & Ribera, M. (2014b). How to Textually Describe Images in Medical Academic Publications. In Proceedings *of the XV International Conference on Human Computer Interaction*, 67:1–67:3. New York, NY, USA: ACM. http://doi.org/10.1145/2662253.2662320

Splendiani, B., & Ribera, M. (2015a). Accessibility of graphics in STEM research articles: Analysis and proposals for improvement. *Journal of the Association for Information Science and Technology*. http://doi.org/10.1002/asi.23464

Splendiani, B., & Ribera, M. (2015b). Alternative texts to images: a review of guidelines. In *III International Seminar on LIS Education and Research* (LIS-ER), Barcelona (Spain).

Splendiani, B., & Ribera, M. (2015c). A proposal for the inclusion of accessibility criteria in the publishing workflow of images in biomedical academic articles. In *6th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Infoexclusion* (DSAI 2015).

Splendiani, B., Ribera, M., & Centelles Velilla, M. (2014). Are figures accessible in mathematics academic journals? Available at http://diposit.ub.edu/dspace/handle/2445/54814 [Accessed 9 May 2015].

Splendiani, B., Ribera, M., García, R., & Salse, M. (2012). An interdisciplinary approach to alternative representations for images. In *13th International Conference on Computers Helping People with Special Needs* (ICCHP'13), Linz (Austria). Available at http://www.uld-conference.org/conference/linz-2012/paper-details?paper=174 [Accessed 9 May 2015].

Splendiani, B., Ribera, M., Garcia, R., & Termens, M. (2014). Do Physicians Make Their Articles Readable for Their Blind or Low-Vision Patients? An Analysis of Current Image Processing Practices in Biomedical Journals from the Point of View of Accessibility. *Journal of Digital Imaging*, 1-24. http://doi.org/10.1007/s10278-014-9674-3

Springer Publishing. (2015). Springer Author Academy. Available at http://academy.springer.com/ [Accessed 9 May 2015].

Springer Publishing. (n. d.). Springer - Manuscript Guidelines Journals. Available at http://www.springer.com/authors/manuscript+guidelines?SGWID=0-40162-12-331200-0 [Accessed 9 May 2015].

Srinivasarao, V., Pingali, P., & Varma, V. (2011). Effective Term Weighting in ALT Text Prediction for Web Image Retrieval. In X. Du, W. Fan, J. Wang, Z. Peng, & M. A. Sharaf (Eds.), *Web Technologies and Applications* (6612, 237-244). Berlin, Heidelberg: Springer Berlin Heidelberg. Available at http://www.springerlink.com/index/10.1007/978-3-642-20291-9_24

Stangl, A., Kim, J., & Yeh, T. (2014). 3D Printed Tactile Picture Books for Children with Visual Impairments: A Design Probe. In Proceedings *of the 2014 Conference on Interaction Design and Children*, 321–324. New York, NY, USA: ACM. http://doi.org/10.1145/2593968.2610482

Stefik, A. M., Hundhausen, C., & Smith, D. (2011). On the design of an educational infrastructure for the blind and visually impaired in computer science. In Proceedings *of the 42nd ACM technical symposium on Computer science education*, 571–576. New York, NY, USA: ACM. http://doi.org/10.1145/1953163.1953323

Stephanidis, C. (2009). *The Universal Access Handbook* (1st ed.). Boca Raton, FL, USA: CRC Press, Inc.

Stephanidis, C. (2011). Universal Access in Human-Computer Interaction. Users Diversity. In *Proceedings of 6th International Conference, UAHCI*, Held as Part of HCI International 2011, Orlando, FL, USA. Part II. Lecture Notes in Computer Science 6766, Springer.

Stephanidis, C. (2013). Design 4 All. In *The Encyclopedia of Human-Computer Interaction* (2.[a] ed.). Aarhus, Denmark: The Interaction Design Foundation. Available at http://www.interaction-design.org/encyclopedia/design_4_all.html [Accessed 9 May 2015].

Strübing, A. (2012). LATEX und barrierefreies PDF. Available at http://www.dante.de/events/Archiv/dante2012/Programm/Vortraege/vortrag-struebing.pdf [Accessed 9 May 2015].

Sutton, J. (2002). A Guide to Making Documents Accessible to People Who Are Blind or Visually Impaired. Available at http://www.sabeusa.org/wp-content/uploads/2014/02/A-Guide-to-Making-Documents-Accessible-to-People-Who-are-Blind-or-Visually-Impaired.pdf [Accessed 9 May 2015].

Swenson, A. M. (2013). A Second Look at Large Print Materials Revised. Available at http://ctserc.org/brailleliteracy/12%20A%20Second%20Look%20at%20Large%20Print%20Materials%20Revised.pdf?2fa6f942252db2ec6c621fe255459617=f5ef442384ab1e06d6a98b92783e6be6 [Accessed 9 May 2015].

Takagi, H., & Ishihara, T. (2007). Technology Advances and Standardization Toward Accessible Business Graphics. In C. Stephanidis (Ed.), *Universal Access in Human-Computer Interaction. Applications and Services* (pp. 426-435). Springer Berlin Heidelberg. Available at http://link.springer.com/chapter/10.1007/978-3-540-73283-9_48

Tang, L. (2012). *Producing Informative Text Alternatives for Images* (PhD thesis). Department of Computer Science, University of Saskatchewan. Available at: http://ecommons.usask.ca/handle/10388/ETD-2012-09-657 [Accessed 9 May 2015].

TAPoR Team. (2012). Comparator - Plain Text (TAPoRware). University of Alberta. Available at http://www.tapor.ca/?id=61 [Accessed 9 May 2015].

Public Law 100-407 (1988). Technology-Related Assistance for Individuals with Disabilities Act of 1988. Available at http://ataporg.org/legislative.html [Accessed 9 May 2015].

The New England Journal of Medicine. (2012). NEJM — Author Center - Images in Clinical Medicine. Available at http://www.nejm.org/page/author-center/images-in-clinical-medicine [Accessed 9 May 2015].

Thompson, J. B. (2013). Books in the Digital Age: The Transformation of Academic and Higher Education Publishing in Britain and the United States. 1 edition. Cambridge, UK ; Malden, MA: Polity, 2005.

Thyssen, A. (2012). ImageMagick v6 Examples - Annotating Images. Available at http://www.imagemagick.org/Usage/annotating/ [Accessed 9 May 2015].

Tomashek, D. B., Edyburn, K. D., Baumann, R., & Smith, R. O. (2013). The Case For Next Generation Text Description Solutions For Visual Information Accessibility. *RESNA Annual Conference 2013*. Available at http://www.resna.org/sites/default/files/legacy/conference/proceedings/2013/CAC/Tomashek.html [Accessed 9 May 2015].

Touch Graphics, Inc. (n. d.). Decision Tree - Image Sorting Tool. Available at http://diagramcenter.org/decision-tree.html [Accessed 9 May 2015].

UKAAF. (2012a). UKAAF Guidance - Describing images. Available at http://www.ukaaf.org/formats-and-guidance#accessible [Accessed 9 May 2015].

UKAAF. (2012b). UKAAF minimum standards: Clear and large print. Available at http://www.ukaaf.org/guidancedocuments/pdf/ms03/finish [Accessed 9 May 2015].

United Nations. (2006). Convention on the Rights of Persons with Disabilities. Available at http://www.un.org/disabilities/convention/conventionfull.shtml [Accessed 9 May 2015].

United States Department of Education. (2012). Letter from the Office of Office of Special Education and Rehabilitative Services of the United States DEpartment of Education. Available at http://www.google.es/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCIQFjAA&url=http%3A%2F%2Fwww2.ed.gov%2Fpolicy%2Fspeced%2Fguid%2Fidea%2Fmemosdcltrs%2Fosers-nimas.doc&ei=FHHXU-uRL4ib0QWsnIGACg&usg=AFQjCNG30T1P4evT9a22FoINXan3MVITWA&sig2=9qKDtSkQH-Ep7KrK-LibuQ&bvm=bv.71778758,d.d2k&cad=rja [Accessed 9 May 2015].

University of Chicago Press. (2006) The Chicago Manual of Style. The essential guide for writers, editors, and publishers.15th edition. Berkeley, California: University of Chicago Press.

University of Chicago Press. (2015). Manuscript Preparation - Artwork. Available at http://www.press.uchicago.edu/infoServices/prep-art.html [Accessed 9 May 2015].

University of Delaware. (2013). UDel SIGHT: Summarizing Information GrapHics Textually: SIGHT. Available at https://www.eecis.udel.edu/wiki/SIGHT/index.php/SIGHT/CorporaAndExperiments [Accessed 9 May 2015].

University of Hong Kong (2001). Vocabulary Profiler tool. Available at http://www4.caes.hku.hk/vocabulary/profile.htm [Accessed 9 May 2015].

University of Minnesota Duluth. (2014). Information technology and sytems & services - longdesc Tools. Available at http://www.d.umn.edu/itss/training/online/moodle_downloads/accessibility_104/longdesc_tools.html [Accessed 9 May 2015].

USERLab. (2012). Text Alternatives Tool for Images (TATI). University of Saskatchewan. Available at: http://userlab.usask.ca/TATI/Instructions.php. [Accessed 9 May 2015].

US Section 508. (2002a). Section 508 of the Rehabilitation Act (29 U.S.C. 794d), as amended by the Workforce Investment Act of 1998 (P.L. 105-220), August 7, 1998. Available at https://www.section508.gov/ [Accessed 9 May 2015].

US Section 508. (2002b). 1194.22 Web-based intranet and internet information and applications. Available at http://section508.gov/section-508-standards-guide#Web [Accessed 9 May 2015].

US Section 508. (2011). Creating Alternative Text for Images. Available at http://section508.gov/creating-alternative-text-images [Accessed 9 May 2015].

Uthman, E. (2006). Seminoma (microscopic, H&E) [Photo]. Available at https://www.flickr.com/photos/euthman/268018285/ [Accessed 9 May 2015].

Vanderheiden, G. (2000). Universal Design of Consumer Products: Current Industry Practice and Perceptions. Available at http://trace.wisc.edu/docs/ud_consumer_products_hfes2000/index.htm [Accessed 9 May 2015].

Van Harmelen, F., & Fensel, D. (1999). Practical Knowledge Representation for the Web. Available at http://www.cs.vu.nl/~frankh/postscript/IJCAI99-III.html. [Accessed 9 May 2015].

Varma, D. R. (2012). Managing DICOM images: Tips and tricks for the radiologist. *The Indian Journal of Radiology & Imaging*, *22*(1), 4-13. http://doi.org/10.4103/0971-3026.95396

Vision Australia. (2014). Colour Contrast Analyser 2.2 for Web Pages. Available at http://www.visionaustralia.org/digital-access-cca [Accessed 9 May 2015].

W3C. (1999). Web Content Accessibility Guidelines 1.0. Available at http://www.w3.org/TR/WCAG10/ [Accessed 9 May 2015].

W3C. (2000a). Accessibility Features of SVG. Available at http://www.w3.org/TR/SVG-access/ [Accessed 9 May 2015].

W3C. (2000b). Techniques For Accessibility Evaluation And Repair Tools. Available at http://www.w3.org/TR/AERT [Accessed 9 May 2015].

W3C. (2007). The Semantic Web Layer Cake. Available at http://www.w3.org/2007/03/layerCake.svg [Accessed 9 May 2015].

W3C. (2008). Web Content Accessibility Guidelines (WCAG) 2.0. Available at http://www.w3.org/TR/WCAG/ [Accessed 9 May 2015].

W3C. (2010). XML Essentials. Available at http://www.w3.org/standards/xml/core [Accessed 9 May 2015].

W3C. (2012a). RDFa Lite 1.1 W3C Recommendation 07 June 2012. Available at http://www.w3.org/TR/2012/REC-rdfa-lite-20120607/ [Accessed 18 april 2013].

W3C. (2012b). Understanding WCAG 2.0. A guide to understanding and implementing Web Content Accessibility Guidelines 2.0. Available at http://www.w3.org/TR/UNDERSTANDING-WCAG20/Overview.html [Accessed 9 May 2015].

W3C. (2013). Shared Web Experiences: Barriers Common to Mobile Device Users and People with Disabilities. Available at http://www.w3.org/WAI/mobile/experiences.html [Accessed 9 May 2015].

W3C. (2014a). ARIA10 Example 1. Available at http://www.w3.org/WAI/WCAG20/Techniques/working-examples/ARIA10/ARIA10-example1.html [Accessed 9 May 2015].

W3C. (2014b). Contrast (Minimum): Understanding SC 1.4.3. Available at http://www.w3.org/TR/UNDERSTANDING-WCAG20/visual-audio-contrast-contrast.html [Accessed 9 May 2015].

W3C. (2014c). G73: Providing a long description in another location with a link to it that is immediately adjacent to the non-text content | Techniques for WCAG 2.0. Available at http://www.w3.org/TR/2014/NOTE-WCAG20-TECHS-20140408/G73 [Accessed 9 May 2015].

W3C. (2014d). G74: Providing a long description in text near the non-text content, with a reference to the location of the long description in the short description | Techniques for WCAG 2.0. Available at http://www.w3.org/TR/WCAG20-TECHS/G74 [Accessed 9 May 2015].

W3C. (2014e). G102: Providing the expansion or explanation of an abbreviation. Techniques for WCAG 2.0. Available at http://www.w3.org/TR/2014/NOTE-WCAG20-TECHS-20140916/G102 [Accessed 9 May 2015].

W3C. (2014f). H53: Using the body of the object element. Available at http://www.w3.org/TR/2014/NOTE-WCAG20-TECHS-20140916/H53 [Accessed 9 May 2015].

W3C. (2014g). HTML5: Techniques for providing useful text alternatives. Available at http://www.w3.org/TR/html-alt-techniques/ [Accessed 9 May 2015].

W3C. (2014h). Implementation Report W3C HTML5 Image Description Extension (longdesc), 4 September 2014. Available at http://w3c.github.io/test-results/html-longdesc/cr-report.html [Accessed 9 May 2015].

W3C. (2014i). Resize text: Understanding SC 1.4.4. Available at http://www.w3.org/TR/UNDERSTANDING-WCAG20/visual-audio-contrast-scale.html [Accessed 9 May 2015].

W3C. (2014j). Scalable Vector Graphics (SVG) 2. Available at http://www.w3.org/TR/SVG2/ [Accessed 9 May 2015].

W3C. (2014k). Techniques for WCAG 2.0 - G94: Providing short text alternative for non-text content that serves the same purpose and presents the same information as the non-text content. Available at http://www.w3.org/TR/WCAG20-TECHS/G94.html [Accessed 9 May 2015].

W3C. (2014l). Understanding Success Criterion 1.1.1 | Understanding WCAG 2.0. Available at http://www.w3.org/TR/UNDERSTANDING-WCAG20/text-equiv-all.html [Accessed 9 May 2015].

W3C. (2014m). Understanding WCAG 2.0 - Distinguishable: Understanding Guideline 1.4. Available at http://www.w3.org/TR/UNDERSTANDING-WCAG20/visual-audio-contrast.html [Accessed 9 May 2015].

W3C. (2014n). Unusual Words. Understanding Success Criterion 3.1.3 - Understanding WCAG 2.0. Available at http://www.w3.org/TR/2015/NOTE-UNDERSTANDING-WCAG20-20150226/meaning-idioms.html [Accessed 9 May 2015].

W3C. (2014o). Use of Color: Understanding Success Criterion 1.4.1 - Understanding WCAG 2.0. Available at http://www.w3.org/TR/UNDERSTANDING-WCAG20/visual-audio-contrast-without-color.html [Accessed 9 May 2015].

W3C. (2015). HTML5 Image Description Extension (longdesc). W3C Recommendation 26 February 2015. Available at http://www.w3.org/TR/html-longdesc/ [Accessed 9 May 2015].

W3C Incubator Group Report. (2007). Image Annotation on the Semantic Web. Available at http://www.w3.org/2005/Incubator/mmsem/XGR-image-annotation-20070814 [Accessed 9 May 2015].

W3C Incubator Group Report. (2014). Visual Presentation: Understanding Success Criterion 1.4.8 - Understanding WCAG 2.0. Available at http://www.w3.org/TR/UNDERSTANDING-WCAG20/visual-audio-contrast-visual-presentation.html [Accessed 9 May 2015].

W3C Research and Development Working Group. (2015). Tips for Accessibility-Aware Research. Available at http://www.w3.org/WAI/RD/wiki/Tips_for_Accessibility-Aware_Research#Reporting_research_in_an_inclusive_way [Accessed 9 May 2015].

W3C-WAI Act Project. (2014). Images Concepts | Images | WAI Web Accessibility Tutorials. Available at http://www.w3.org/WAI/tutorials/images/ [Accessed 9 May 2015].

Wang, X., Erdelez, S., Alen, C., Anderson, B., Cao, H., & Shyu, C.-R. (2011). Medical image describing behavior: a comparison between an expert and novice. In *Proceedings of the 2011 iConference* (pp. 792–793). New York, NY, USA: ACM. http://doi.org/10.1145/1940761.1940907

Ware, M. (2013). *Peer Review. An Introduction and Guide*. Publishing Research Consortium (PRC). Available at http://www.tandf.co.uk/journals/pdf/PRC-PeerReview-Guide-2013.pdf [Accessed 9 May 2015].

Ware, M., & Mabe, M. (2012). *The STM Report: An overview of scientific and scholarly journal publishing*. Mark Ware Consulting. Available at http://www.stm-assoc.org/2012_12_11_STM_Report_2012.pdf [Accessed 9 May 2015].

Watson, L. (2014). Solving the longdesc problem. Available at http://tink.co.uk/2013/03/solving-the-longdesc-problem/ [Accessed 4 MArch 2014]

World Blind Union. (n. d.). WBU Priorities and Goals. Available at http://www.worldblindunion.org/English/our-work/our-priorities/Pages/default.aspx [Accessed 9 May 2015].

Web Accessibility Iniciative (WAI). (2005). Introduction to Web Accessibility. Available at http://www.w3.org/WAI/intro/accessibility.php [Accessed 9 May 2015].

WebAIM. (2012). Appropriate use of alternative text. Available at http://webaim.org/techniques/alttext/ [Accessed 9 May 2015].

WebAIM. (2014). PowerPoint Accessibility. Available at http://webaim.org/techniques/powerpoint/ [Accessed 9 May 2015].

Whitington, J. (2011). *PDF Explained* (1 edition). Sebastopol, Calif: O'Reilly Media.

Wikipedia. (2013). Anatomical MRI gif of the brain, concurrently showing progressive slices through transverse, sagittal, and coronal planes. Available at http://en.wikipedia.org/wiki/File:BrainMRI3planes.gif [Accessed 9 May 2015].

Wikipedia. (2015). Electrocardiogram 12 derivations male 23yo Japanese. Available at http://commons.wikimedia.org/wiki/File:Electrocardiogram_12derivations_male_23yo_Japanese.png [Accessed 9 May 2015].

Wiley - Blackwell Publishing. (2015). Wiley-Blackwell - Electronic Artwork - An introduction. Available at http://authorservices.wiley.com/bauthor/illustration.asp [Accessed 9 May 2015].

World Health Organization (WHO). (2011). World report on disability. Available at http://www.who.int/disabilities/world_report/2011/en/index.html [Accessed 9 May 2015].

World Health Organization (WHO). (2014). Visual impairment and blindness. Fact Sheet N°282. Available at http://www.who.int/mediacentre/factsheets/fs282/en/ [Accessed 9 May 2015].

World Health Organization (WHO). (n. d.). Change the Definition of Blindness. Available at http://www.who.int/blindness/Change%20the%20Definition%20of%20Blindness.pdf Accessed 9 May 2015].

You, D., Antani, S., Demner-Fushman, D., Rahman, M. M., Govindaraju, V., & Thoma, G. R. (2011). Automatic identification of ROI in figure images toward improving hybrid (text and image) biomedical document retrieval In *SPIE Proceedings, Document Recognition and Retrieval* XVIII (Vol. 78740). San Francisco Airport, CA: SPIE doi:10.1117/12.873434

Yu, H., Agarwal, S., Johnston, M., & Cohen, A. (2009). Are figure legends sufficient? Evaluating the contribution of associated text to biomedical figure comprehension. *Journal of Biomedical Discovery and Collaboration*, *4*(1), 1. http://doi.org/10.1186/1747-5333-4-1

Yu, H., & Lee, M. (2006). Accessing bioscience images from abstract sentences. *Bioinformatics*, *22*(14), e547–e556. http://doi.org/10.1093/bioinformatics/btl261

Zhang, Y., & Wildemuth, B. M. (2009). Qualitative analysis of content. In B. Wildemuth (Ed.), *Applications of Social Research Methods to Questions in Information and Library Science* (pp.308-319). Westport, CT: Libraries Unlimited.

Chapter 10

Appendix

10 Appendix

10.1 Appendix A

Encuesta on-line sobre uso y descripción de imágenes en publicaciones científicas

soy Bruno Splendiani, becario de doctorado del departamento de Biblioteconomía y Documentación de la Universitat de Barcelona.

En mi tesis estoy estudiando el proceso de creación, recuperación y uso de imágenes por parte de investigadores en los ámbitos de medicina y biología, con el objetivo de mejorar el proceso de descripción del contenido visual en las publicaciones científicas. La investigación aportará una serie de beneficios a los autores: una mejora en la recuperación de la información visual del artículo (aumentando la posibilidad de ser citado) y la ampliación del acceso al artículo a lectores con ciertas discapacidades visuales o en condiciones de acceso con dispositivos móviles. También aportará beneficios para los lectores de

los artículos y para los editores.

Bienvenido!

Hay 28 preguntas en esta encuesta

1¿Cuál es su género?

Por favor seleccione **sólo una** de las siguientes opciones:

Femenino

Masculino

2 ¿En qué área trabaja actualmente?

Por favor, marque las opciones que correspondan:

Enseñanza

Investigación

Medicina clínica

Documentación y gestión de la información

Otro:

3 ¿Cuántos años de experiencia tiene en su trabajo?

Por favor, marque las opciones que correspondan:

0 - 1 años

1 - 3 años

3 - 5 años

5 - 10 años

10 - 20 años

más de 20 años

4 ¿Cuál es su posición académica actual?

Por favor seleccione **sólo una** de las siguientes opciones:

Catedrático de Universidad

Titular de Universidad

Titular de Escuela Universitaria

Contratado docente

Personal Externo

Becario

Estudiante (sin beca)

No tengo vinculación con la universidad

5 ¿De qué facultad(es) depende el área en la cual trabaja actualmente?

Por favor seleccione **sólo una** de las siguientes opciones:

Medicina

Biología

Otro

6 ¿Cuál es su área de interés específica?

Por favor, marque las opciones que correspondan:

Anatomía Patológica, Farmacología y Microbiología

Biología Celular, Inmunología y Neurociencias

Ciencias Clínicas

Ciencias Fisiológicas

Cirugía y Especialidades quirúrgicas

Medicina

Obstetricia i Ginecología, Pediatría y Radiología y Medicina Física

Odontostomatología

Patología y Terapéutica Experimental

Psiquiatría y Psicobiología Clínica

Salud Pública

Otro:

7 ¿Cuál es su área de interés específica?

Por favor, marque las opciones que correspondan:

Biología Animal

Biología Celular

Biología Vegetal

Bioquímica y Biología Molecular

Cristalografía, Mineralogía y Depósitos Minerales

Ecología

Estadística

Estratigrafía, Paleontología y Geociencias Marinas

Fisiología (biología)

Genética

Microbiología

Nutrición y Bromatología

8 ¿Suele utilizar imágenes para ilustrar sus publicaciones?

Por favor seleccione **sólo una** de las siguientes opciones:

Sí

No

9 ¿En qué tipo de publicación suele utilizar imágenes?

Por favor, marque las opciones que correspondan:

Abstract

Artículo

Artículo en la prensa

Actos de congresos

Capítulo de libro

Informe de investigación

Informe de trabajo

Libro

Material docente

Publicaciones multimedia

Revisión

Revisión en prensa

Otro:

10 ¿Por qué utiliza imágenes en su publicaciones? *

Por favor, marque las opciones que correspondan:

La información no puede ser presentada de manera fácil y clara de otra forma, como tabla o texto

Para ilustrar el contenido del artículo

Para enfatizar detalles esenciales

Para enseñar datos numéricos

Para soportar el argumento con un ejemplo visual

Otro:

11 ¿Qué tipo de imagen suele utilizar en sus publicaciones científicas?

Por favor, marque las opciones que correspondan:

Cromatografía, gel

Estructuras químicas

Fórmulas matemáticas

Fotografías (imágenes ópticas, incluyendo imágenes microscópicas y endoscópicas)
Gráficos (incluyendo diagramas e ilustraciones)
Medicina nuclear (excluyendo PET)
PET (tomografía por emisión de positrones)
Radiografía (incluyendo imágenes fluoroscópicas y angiográficas)
Resonancia magnética
Secuencias genéticas
Señales y ondas (electroencefalograma, electrocardiograma, electromiografía)
Tomografía axial computarizada (TAC)
Ultrasonido (incluyendo color Doppler e imágenes de ultrasonido 3D)
Otro:

12 ¿Qué tipo de imagen suele utilizar en sus publicaciones científicas?

Por favor, marque las opciones que correspondan:
Estructuras químicas
Imágenes de morfología celular
Imágenes de proteínas motoras
Otras micrografías
Fórmulas matemáticas
Fotografías (excluyendo imágenes microscópicas)
Gráficos (incluyendo diagramas e ilustraciones)
Secuencias de imágenes en timelapse
Secuencias genéticas
Otro:

13 ¿Qué tipo de imagen suele utilizar en sus publicaciones científicas?

Por favor, marque las opciones que correspondan:
Estructuras químicas
Fórmulas matemáticas
Fotografías
Gráficos (incluyendo diagramas e ilustraciones)
Otro:

14

¿Qué formato utiliza normalmente cuando trabaja con imágenes que ilustran sus publicaciones? Por favor, para dos o más formatos especifique un motivo del porqué los utiliza.

Por favor, seleccione todas las opciones que correspondan y escriba un comentario:
BMP
DICOM
DOC

(E)PS
GIF
JPEG
JPEG2000
PDF
PNG
PPT
PSD
RAW (formato nativo de cámara
digital)
SVG
TIFF
Otro:

15 ¿Cada cuánto tiempo crea o busca imágenes para ser publicadas en sus artículos científicos?

Por favor seleccione **sólo una** de las siguientes opciones:

Diariamente
Semanalmente
Mensualmente
Nunca

16 ¿Hay alguna anotación asociada a la imagen que considere especialmente relevante a la hora de buscar a una imagen? *

Por favor, marque las opciones que correspondan:

Anotación textual, como símbolos, letras u otras etiquetas
Anotación visual, como punteros (línea, flecha, etc.) o áreas destacadas (polígonos, etc.)
Ninguna
Otro:
Ejemplo de imagen con anotaciones

17 ¿Qué fuentes suele consultar para encontrar una imagen con la intención de utilizarla en una publicación científica?

Por favor, seleccione todas las opciones que correspondan y escriba un comentario:

Amigos y compañeros de trabajo
Artículos científicos
CDs y DVDs
Colecciones del departamento
Colecciones personales
Libros
Recurso en la Web (por favor,

especifique un recurso favorito, por
ejemplo Google Images)

Sistemas específicos de almacenamiento, transmisión y descarga de imágenes digitales (por ejemplo los servidores PACS para imágenes radiológicas)

Otro:

18 Durante la lectura de un artículo científico, normalmente obtiene información relevante sobre las imágenes contenidas en el artículo mirando a…

Por favor, marque las opciones que correspondan:

el título de la imagen

la información técnica asociada a la imagen en general (por ejemplo, los metadatos DICOM, que incluyen información sobre quien, adonde y como se ha tomado una imagen)

la leyenda de la imagen

las anotaciones de la imagen (flechas, símbolos, círculos, etc.)

las referencias en el artículo

Ninguna de las anteriores

Otro:

Preguntas relacionadas con la descripción textual de imágenes

19 ¿Suele añadir alguna anotación a las imágenes que utiliza en sus publicaciones?

Por favor, marque las opciones que correspondan:

Anotación textual, como símbolos, letras u otras etiquetas

Anotación visual, como punteros (línea, flecha, etc.) o áreas destacadas (polígonos, etc.)

No añado ninguna anotación

Ejemplo de imagen con anotaciones

20 ¿Suele asociar a la imagen que usa para un artículo científico una descripción textual?

Por favor seleccione **sólo una** de las siguientes opciones:

Sí

No

21 ¿Cuándo suele crear la descripción textual de la imagen? *

Por favor, marque las opciones que correspondan:

Cuando creo la imagen

Cuando envío la imagen al editor para su publicación en un artículo científico

Cuando edito/modifico la imagen

Otro:

22 ¿Por qué no suele asociar a la imagen una descripción textual?

Por favor, marque las opciones que correspondan:

La imagen es suficientemente autoexplicativa

La imagen está suficientemente descrita en el artículo

La imagen está suficientemente descrita en la leyenda

No encuentro un motivo por hacerlo

Otro:

23 ¿Qué herramienta utiliza normalmente para editar/anotar las imágenes que utiliza en sus publicaciones?

Por favor, marque las opciones que correspondan:

Adobe Illustrator

Adobe Photoshop

Gimp

Irfanview

Microsoft PowerPoint

No utilizo ninguna herramienta de edición de imágenes

Otro:

24 Por favor, envíe una imagen representativa entre aquellas que utiliza en sus publicaciones, incluyendo una descripción de la misma.

25 ¿Conoce el significado de la definición de "texto alternativo" de una imagen?

Por favor seleccione **sólo una** de las siguientes opciones:

Sí

No

26 ¿Conoce las políticas de accesibilidad en revistas científicas para personas con discapacidad visual?

Por favor seleccione **sólo una** de las siguientes opciones:

Sí

No

27 ¿Ha experimentado alguna vez alguna dificultad para ver bien una imagen publicada en un artículo científico?

Por favor seleccione **sólo una** de las siguientes opciones:

Sí

No

28

Por favor, indique que tipo de dificultad ha experimentado en ver bien la imagen

Por favor, escriba su respuesta aquí:

La encuesta ha terminado. Muchas gracias por su colaboración.

10.2 Appendix B

Primer email de contacto para la encuesta

Estimado/a {FIRSTNAME} {LASTNAME},

soy Bruno Splendiani, becario de doctorado del departamento de Biblioteconomía y Documentación de la Universitat de Barcelona.

En mi tesis estoy estudiando el proceso de creación, recuperación y uso de imágenes por parte de investigadores en los ámbitos de medicina y biología, con el objetivo de mejorar el proceso de descripción del contenido visual en las publicaciones científicas. La investigación aportará una serie de beneficios a los autores: una mejora en la recuperación de la información visual del artículo (aumentando la posibilidad de ser citado) y la ampliación del acceso al artículo a lectores con ciertas discapacidades visuales o en condiciones de acceso con dispositivos móviles. También aportará beneficios para los lectores de los artículos y para los editores.

Siendo usted miembro del grupo de investigación de la UB en "{ATTRIBUTE_1}", le agradecería que rellenara la siguiente encuesta online:

{SURVEYURL}

Toda la información que se proporcione en la encuesta será recogida de forma anónima, tratada de forma confidencial y utilizada únicamente con fines de investigación. El tiempo estimado de realización de la encuesta es de 10 minutos.

Agradeciéndole de antemano por su ayuda, reciba un cordial saludo.

Bruno Splendiani

Dep. Bibioteconomia i Documentació

C/ Melcior de Palau 140, 08014 Barcelona

Universitat de Barcelona

Tel. 93.403.40.17

Email recordatorio para la encuesta

Estimado/a {FIRSTNAME} {LASTNAME}:

Recientemente se le invitó a participar en la "Encuesta online sobre uso y descripción de imágenes en publicaciones científicas”. Quería recordarle que se encuentra todavía disponible si desea participar. Para hacerlo, por favor pulse en el siguiente enlace: {SURVEYURL}

Como miembro del grupo de investigación de la UB en "{ATTRIBUTE_1}", su colaboración ayudará a mejorar el estudio de las necesidades en el proceso de publicación de imágenes en revistas científicas.

Nuevamente le agradezco su colaboración.

Atentamente,

{ADMINNAME} ({ADMINEMAIL})

Email de confirmación de encuesta completada

Estimado/a {FIRSTNAME} {LASTNAME}:

Este correo es para confirmarle que ha completado la encuesta titulada "{SURVEYNAME}" y sus respuestas han quedado correctamente guardadas.

Le agradezco por su participación y le recuerdo de qué si tiene alguna duda o consulta adicional, puede contactarme al email splendiani@ub.edu.

Reciba un cordial saludo,

{ADMINNAME}

Email de contacto para entrevista

Estimado {FIRSTNAME} {LASTNAME},

mi nombre es Bruno Splendiani, becario de doctorado del departamento de Biblioteconomía y Documentación de la Universitat de Barcelona.

Quería agradecerle de nuevo por haber completado la encuesta online sobre el uso y la descripción de imágenes en publicaciones científicas. Gracias a su colaboración (y a la de muchos otros investigadores) he podido definir las tendencias actuales en el proceso de creación, recuperación y uso de imágenes por parte de investigadores en ámbito médico.

Para poder completar los datos recogidos en la encuesta, estoy realizando unas entrevistas con personal investigador y una observación del contexto en el que trabajan a diario con imágenes médicas. Por eso le quería preguntar si tuviera disponibilidad en realizar una entrevista de aproximadamente media hora sobre su trabajo diario con imágenes, su creación, manipulación y uso. Su opinión sería muy relevante para mi tesis, debido a su perfil de médico y académico, que además publica artículos científicos incluyendo imágenes.

Agradeciéndole de antemano por su disponibilidad, le saludo cordialmente,

Bruno Splendiani

Dep. Bibioteconomia i Documentació

C/ Melcior de Palau 140, 08014 Barcelona

Universitat de Barcelona

Tel. 93.403.40.17

10.3 Appendix C

Guion de las entrevistas para el estudio sobre opiniones y prácticas de investigadores en Medicina y Biología en la publicación de imágenes científicas

Primero quería hacerte unas preguntas generales sobre tu trabajo y perfil. Después te preguntaré sobre porqué necesitas imágenes en tu trabajo y como las utilizas.

1. Podrías decirme cuál es tu formación y experiencia en tu trabajo?

2. En qué consiste tu trabajo y cuáles son tus intereses en la investigación?

3. Para qué necesitas imágenes en tu trabajo y como las incorporas en tu trabajo?

4. Por qué utilizas habitualmente imágenes en tus publicaciones? Qué función tienen?

5. Qué criterios utilizas para seleccionar las imágenes que incluyes en tus publicaciones?

6. Cómo y porqué sueles modificar/editar una imagen desde su creación hasta su publicación in una revista científica?

7. Qué tecnologías y formatos utilizas normalmente cuando trabajas con imágenes?

Podrías enseñarme un artículo científico en el qué has utilizado imágenes? (Alternativa: piensa en la última vez que has utilizado una imagen para una publicación)

8. Por qué creas descripciones textuales de la imagen (figura)?

9. Qué función y peso tienen las descripciones textuales dentro del artículo?

10. Podrías describirme el típico proceso de creación de la pié de texto de las imágenes en un artículo científico? Por ejemplo, qué elementos esenciales de la imagen sueles describir textualmente? Utilizas algunas pautas en la creación de alternativas textuales de las imágenes?

11. Podrías describirme el típico proceso de creación de la referencias a las imágenes en un artículo científico?

Podrías enseñarme un artículo relevante en tu investigación que contenga imágenes? (alternativa: piensa en la última vez que leíste un artículo con imágenes relevante para tu investigación)

12. Consideras que las imágenes han sido esenciales para entender el artículo? Qué te han aportado? Si no estuvieran que perderías?

13. Consideras que los textos descriptivos de las imágenes han sido esenciales para entender las imágenes del artículo?

14. Como lees normalmente artículos científicos con imágenes?

15. Has tenido alguna vez algún problema en leer las imágenes dentro de un artículo? Qué tipo de problemas ha experimentado?

Podrías enseñarme como buscaste el artículo con imágenes relevante en tu investigación? (alternativa: piensa en la última vez que buscaste un artículo con imagen)

16. Qué tipo de criterio has utilizado para buscar la imagen?

17. Qué información textual asociada a la imagen te ha sido útil para encontrar una imagen?

10.4 Appendix D

Documento 1

Consentimiento informado

Para participar al estudio:

"Estudio sobre uso y descripción de imágenes en publicaciones científicas"

Yo *(nombre y apellido del participante)*.. ,

Con DNI núm. ....

He leído el *Documento de Información* que me ha sido entregado sobre el estudio.

He podido hacer preguntas sobre el estudio.

He recibido información suficiente sobre el estudio.

He hablado con el investigador Bruno Splendiani

Con DNI núm X6915638K.

Entiendo que mi participación es voluntaria.

Doy libremente mi conformidad para participar al estudio .

Data, ...

Signaturas:

El participante El investigador

Para más información o cualquier tema relacionado con el proyecto os podéis dirigir a:

Bruno Splendiani

Departamento de Biblioteconomía i Documentación

Universitat de Barcelona

Melcior de Palau, 140. 08014 Barcelona

Teléfono: 93 403 40 17 Correo electrónico: splendiani@ub.edu

Documento 2

Documento de autorización para ser grabado

Dentro del estudio:

"Estudio sobre uso y descripción de imágenes en publicaciones científicas"

Yo *(nombre y apellido del participante)*.. ,

con DNI núm...

Autorizo la grabación de mi voz durante la entrevista que se realiza dentro del estudio "Uso y descripción de imágenes en publicaciones científicas". Esta grabación solo podrá ser utilizada con finalidades científicas para el análisis de los datos recogidos en el proyecto o para divulgar los resultados. Bajo ningún concepto se podrá hacer un uso da los datos que pueda vulnerar mi imagen o mi dignidad personal. Tampoco se podrá hacer un uso comercial de los mismos.

Data, ...

Signatura:

El participante

Para más información o cualquier tema relacionado con el proyecto os podéis dirigir a:

Bruno Splendiani

Departamento de Biblioteconomía i Documentación

Universitat de Barcelona

Melcior de Palau, 140. 08014 Barcelona

Teléfono: 93 403 40 17 Correo electrónico: splendiani@ub.edu

10.5 Appendix E

Ejemplo de codificación de las respuestas a la entrevista relacionadas con el tema de pie de figura en artículos científicos generado a través del software Nvivo (v.10).

<Internals\\P1 transcription> - § 3 references coded [16,02% Coverage]

Reference 1 - 7,61% Coverage

Es obligado mencionarlo, decir mínimamente lo que quieres decir con aquella figura o tabla, que normalmente se envía en hojas aparte, casi todas las revistas te lo piden de enviarla en hojas aparte, ellos la incluyen. Todas tienen que tener pie de texto. Muchas te lo hacen poner en una hoja con pie de texto. Pie de texto tiene que ser muy simple, muy claro, muy informativo de tal forma que solo mirando la figura y el pie de texto tú tienes que ver el mensaje.

Reference 2 - 1,14% Coverage

Pero en los resultados tienes que ponerlo y con pie de texto obligado.

Reference 3 - 7,27% Coverage

Lo ideal es que sobrara el pie de texto. Tienen que pretender que aunque el pie de texto sea obligado, enfatice o deje más claro, que casi sobre, aunque sea obligado. No debería ser esencial para entender la figura. Si me pasaras figuras sin pie de texto, en la mayor parte aunque fuera sin pie de texto yo creo que te podría decir lo que hay en aquella figura. La mejores revistas de índex medicus, medline, pubmed en el campo de biomedicina.

<Internals\\P10_transcription> - § 1 reference coded [11,41% Coverage]

Reference 1 - 11,41% Coverage

hay revistas que te dicen "tienes que explicar lo que es la imagen en el pie de figura" Y lo que sería el cuerpo explicas porque es interesante, el caso..... En otra revista el pie de figura tiene que ser una cosa esquemática entonces luego explicas todas las referencias..."en la figura tal se ve esto" y el cuerpo el caso clínico.

<Internals\\P11_transcription> - § 3 references coded [15,85% Coverage]

Reference 1 - 9,43% Coverage

En algunos artículos el pie de figura sirve solo para describir la imagen, describir la situación que puede acompañar a la existencia de la lesión representada en la imagen. Porque a lo mejor se trata de una lengua no con una verruga, sino una lengua seca, resquebrajada, agrietada, en la cual lo importante no son las grietas, sino aquello que las gritas representan, que significa sequedad de boca, deshidratación. Entonces esa lengua no nos dice "lengua agrietada" sino "lengua de paciente deshidratado, cuya manifestación clínica en la boca es una lengua agrietada".

Reference 2 - 4,70% Coverage

Es muy importante destacarlo en el pie de figura. Sobre todo los lectores no muy conocedores del tema que se está tratando, pueden ignorarlo. Entonces cuanto más claro sea el pie de figura y más claramente este descrito en el texto tanto mejor para la comprensión del trabajo entero.

Reference 3 - 1,72% Coverage

Nunca he encontrado recomendaciones de los editores sobre como describir la imagen en el pie de figura.

<Internals\\P12_transcription> - § 3 references coded [11,55% Coverage]

Reference 1 - 7,32% Coverage

Es obligado poner pie de figura y cada vez más tiene que ser muy explicativo. Incluso con algún parámetro estadístico, aunque no sea muy frecuente. Explico elementos concretos de la figura. En el pie de figura se indican los elementos que necesitas para entender de manera prácticamente autónoma esa figura.

Reference 2 - 1,88% Coverage

mientras en el pie de figura dices "aumenta estadísticamente" y añades la "p".

Reference 3 - 2,35% Coverage

La imagen podría ser interpretada sin el pie de figura, pero tienes que saber que has hecho antes.

<Internals\\P13_transcription> - § 2 references coded [7,21% Coverage]

Reference 1 - 4,51% Coverage

En el pie de página incluyo los signos que utilizo que significa, para que la gente entienda fácilmente.

Viene a ser una ampliación de esto (?) pero que lo has explicado aquí en la discusión.

Reference 2 - 2,70% Coverage

Aquí (En el pie de página?) puedes ir a ver el valor en cambio aquí (?) lo explico el valor, digo qué valor tiene.

10.6 Appendix F

Results of the survey

Participants Recruitment

1005 emails were sent to the participants selected with an introductory presentation of the study (see the appendix) and a link to the survey. 797 emails were delivered (Figure 63).

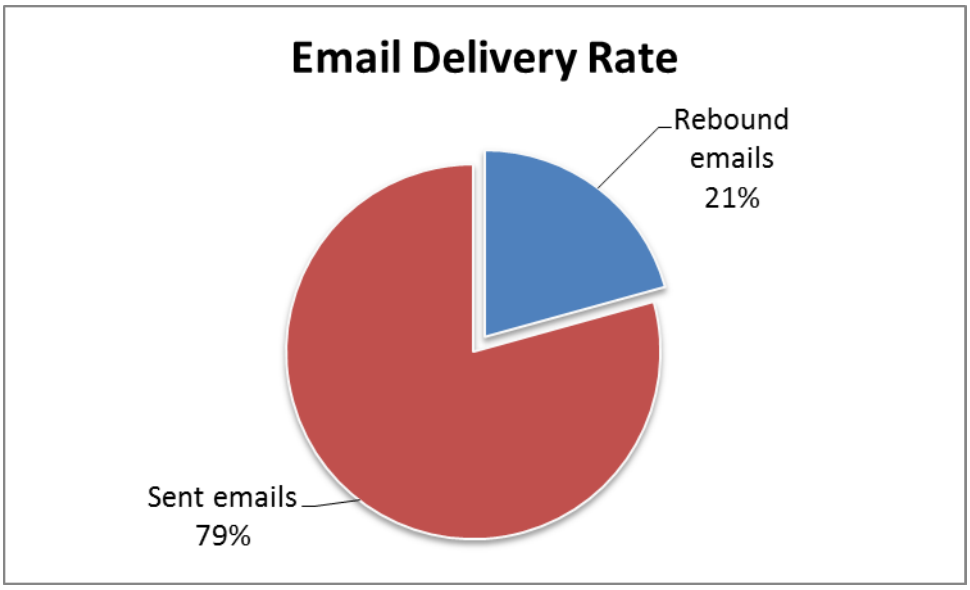


Figure 63 Percentage of the email delivery rate of the survey

198 participants responded to the questionnarie: 179 participants completed the questionnaire, 19 completed the questionnaire partially, while 551 participants did not answered the questionnaire (Figure 64).

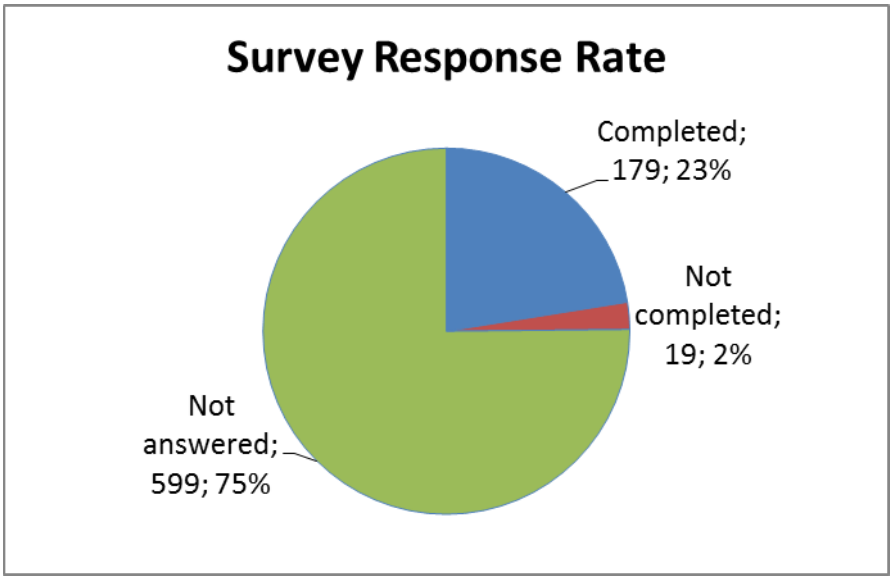


Figure 64 Survey response rate of the survey

Demographics

54% of the participants that responded the survey (partially or totally) were male, 46% female (Figure 65).

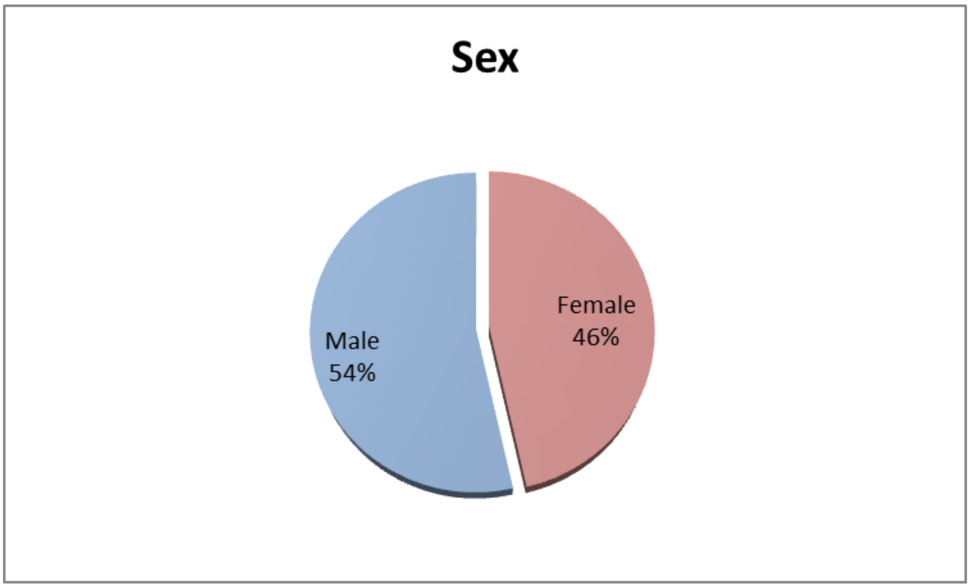


Figure 65 Distribution of the participants of the survey per sex

Figure 66 shows the job profile of the participants.

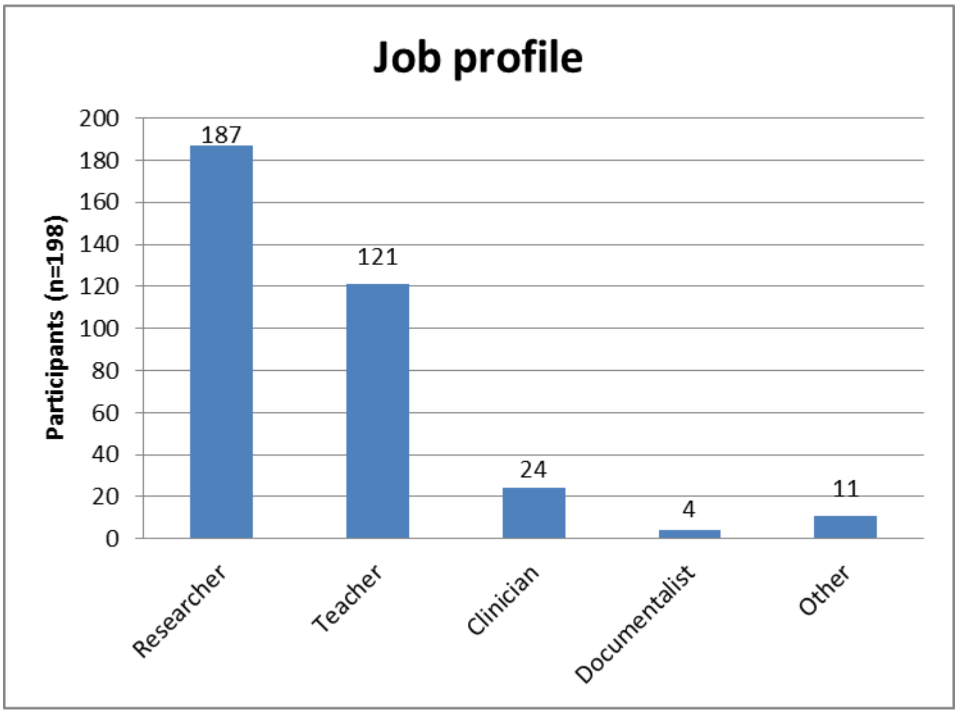


Figure 66 Distribution the participants per job profile

The great majority of answers specified the job profile as researcher (187 answers) and academic teacher (121 answers). 24 participants defined themselves as “Clinician” and four as “Documentalist”. Other profiles (11 participants) included technicians, administrative assistants and scientific divulgators. One category does not exclude others.

More than half of the participants (53%) had 20 or more years of experience in their work, 13% between 10 and 20 years, 22 % between 5 and 10 years and 12% less than 5 years (Figure 67).

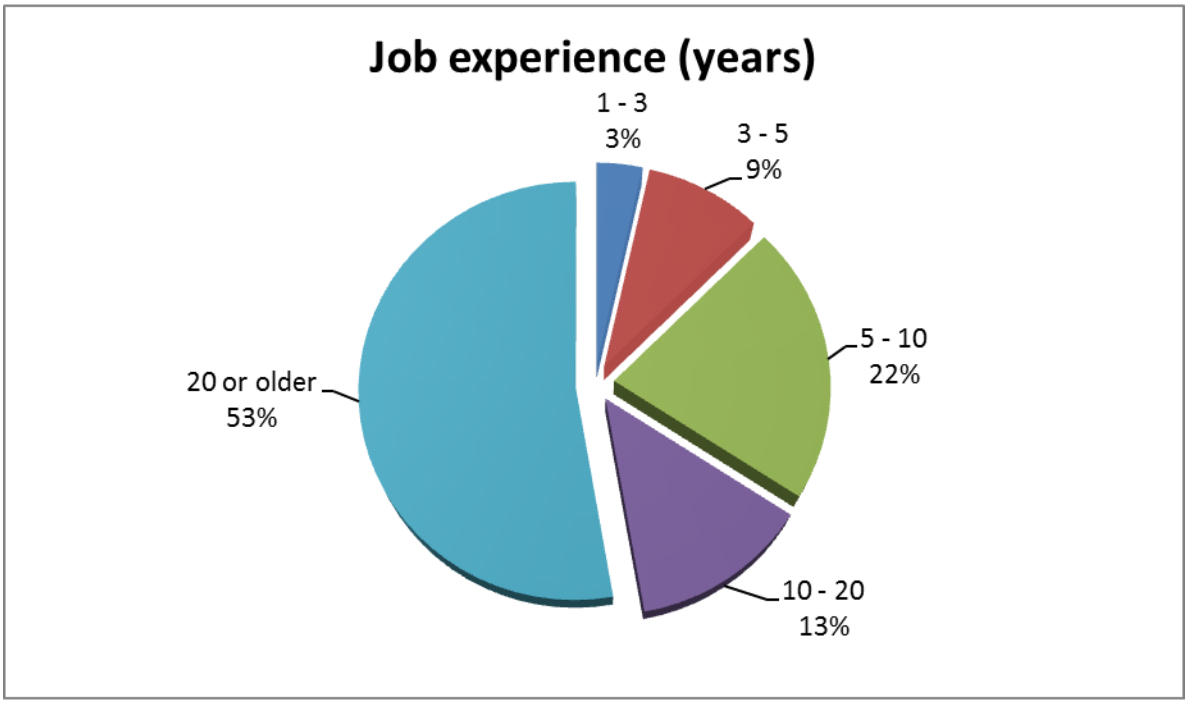


Figure 67 Distribution of the participants according to the years of experience in their job

The academic role of the participants presented high variations. 32% were professors, 24% readers, 16% professors with chair, 12% research fellow associate, 2% students and 5% did not have academic roles (Figure 68).

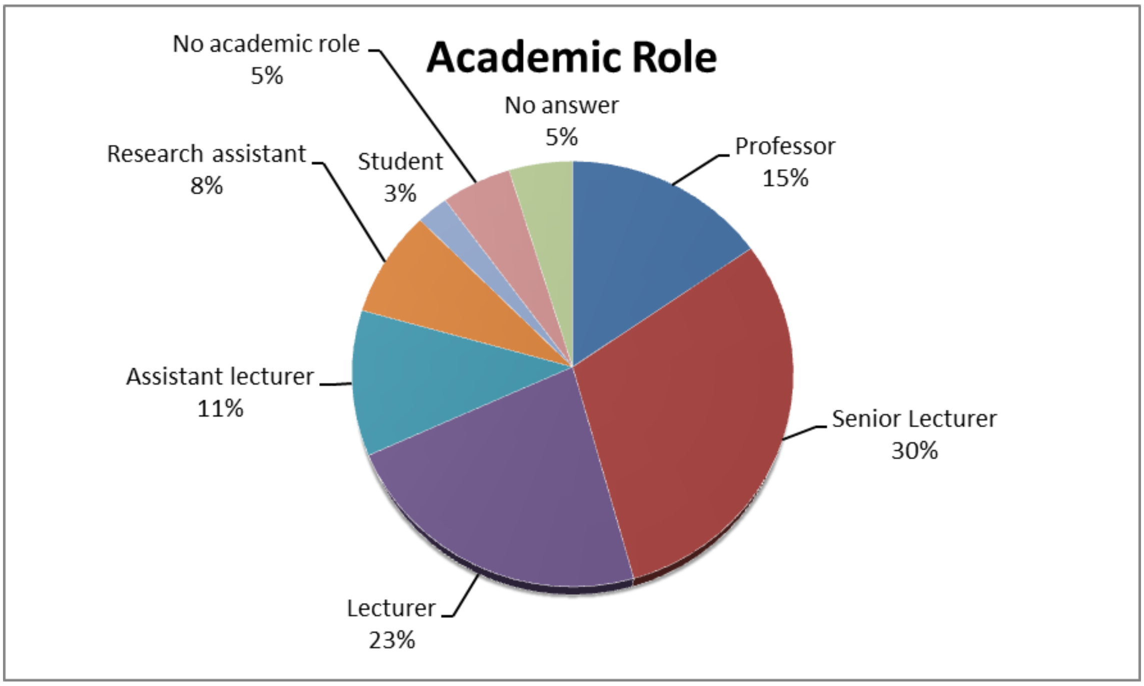


Figure 68 Distribution of the participants according to the academic role.

The general research field of the participants was biology (60%), medicine (26%) and others (14%) (Figure 69).

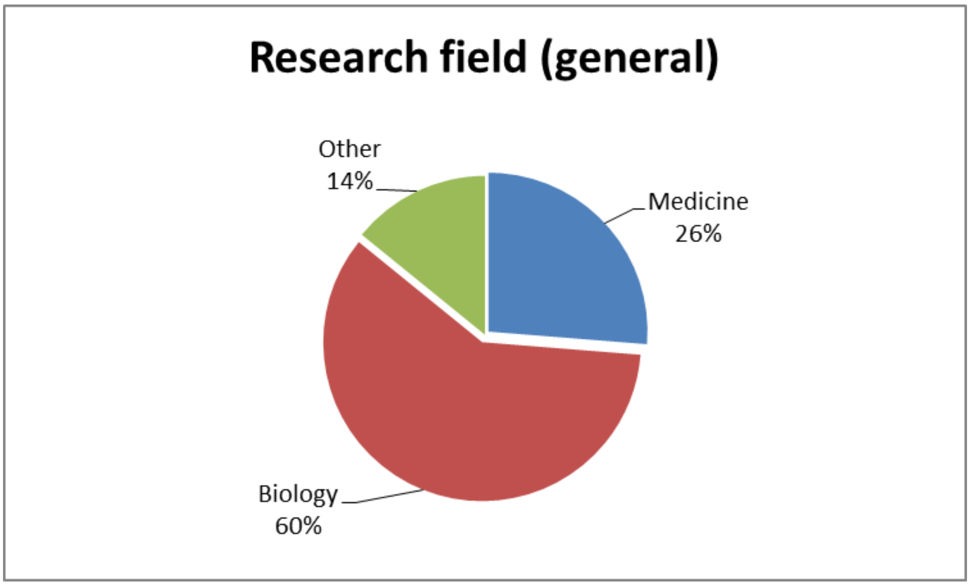


Figure 69 Distribution of the participants according to the general research field

In the medical research area, there was a high variety in specialization field, as showed in Figure 70 - Table XLVIII.

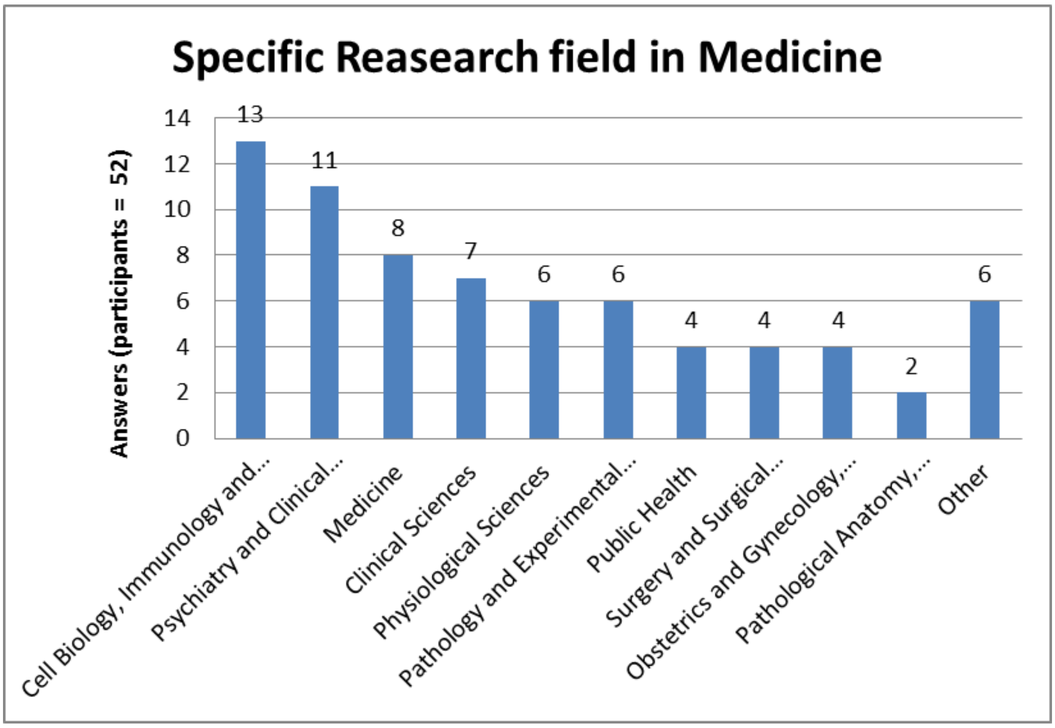


Figure 70 Distribution of participants belonging to the medical area according to the specific field research field in Medicine. The field "Cell biology." includes Immunology and Neuroscience, "Psychiatry…" includes Clinical Psychobiology; "Pathology…" includes "Experimental Therapy"; "Surgery…" includes Surgical Specializations; "Obstetrics" includes "Gynecology", Pediatrics", Radiology and "Anatomy"; Pathological Anatomy…" includes "Pharmacology" and "Microbiology";

“Other” includes “Nephrology” (three participants), “Hematology”, “Evolutionary Biology”, “Nephrology” and Neuropsychology (one participant each). More than one answer is allowed.

Research Field (medicine)	Answers
Cell Biology, Immunology and Neuroscience	13
Psychiatry and Clinical Psychobiology	11
Medicine	8
Clinical Sciences	7
Physiological Sciences	6
Pathology and Experimental Therapy	6
Public Health	4
Surgery and Surgical Specializations	4
Obstetrics and Gynecology, Pediatrics, Radiology and Anatomy	4
Pathological Anatomy, Pharmacology and Microbiology	2
Other	6

Table XLVIII Specific research field of the participants in the area of Medicine

In the research area of Biology, a high variety in specialization field was also found, as showed in Figure 71 - Table XLIX.

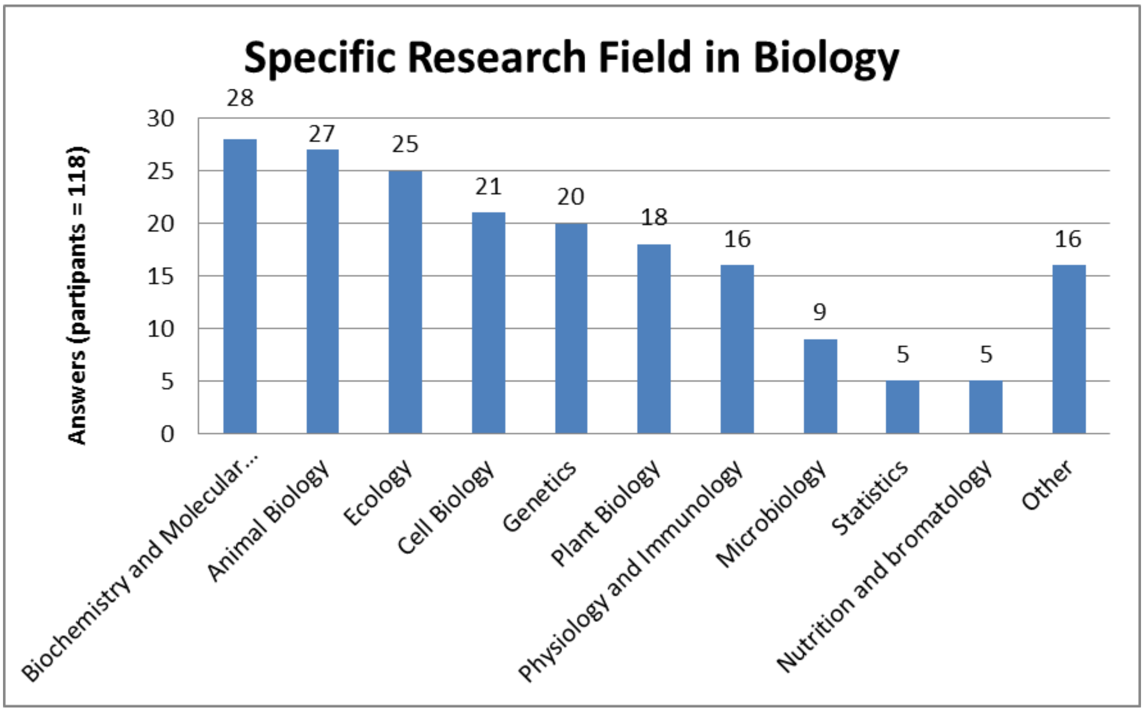


Figure 71 Distribution of participants belonging to the biological area according to the specific research field in Biology. “Biochemistry” includes “Molecular Biology”. “Other” includes: Neurobiology (three participants), Anthropology (two participants), Oceanography, Genomics,

Training in Experimental Sciences, Bioinformatics, Oncology and Developmental Biology (one participant each). More than one answer is allowed.

Research Field (biology)	**Answers**
Biochemistry and Molecular Biology	28
Animal Biology	27
Ecology	25
Cell Biology	21
Genetics	20
Plant Biology	18
Physiology and Immunology	16
Microbiology	9
Statistics	5
Nutrition and bromatology	5
Other	16

Table XLIX Specific research field of the participants in the area of Biology

Other research areas not included in the fields of Medicine and Biology are showed in Figure 72 - Table L:

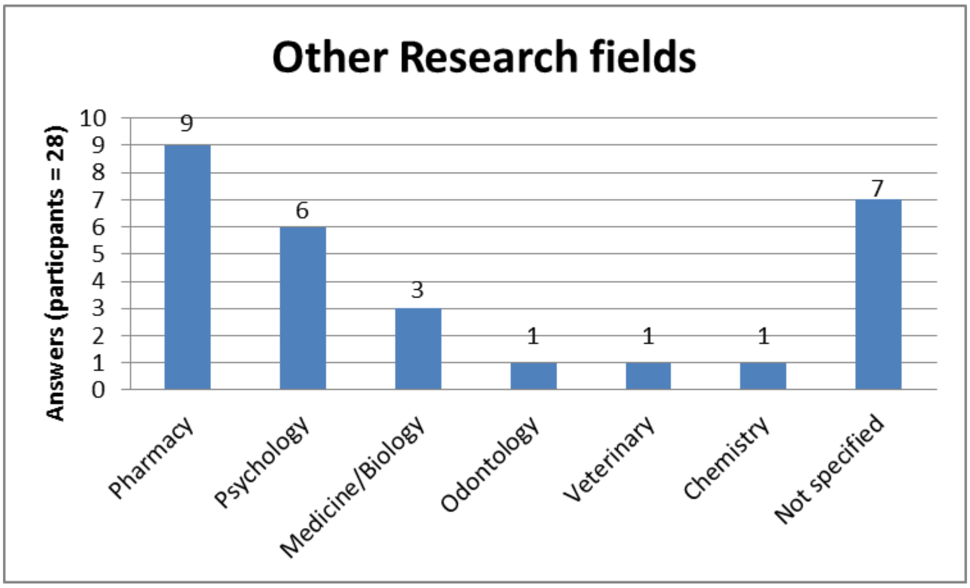


Figure 72 Distribution of participants who declared belonging to main research fields different from Medicine and Biology. "Medicine/Biology" is an exception and indicates participants that responded "Medicine and biology " as "other research fields".

Research field (other)	Answers
Pharmacy	9

Psychology	6
Medicine/Biology	3
Odontology	1
Veterinary	1
Chemistry	1
Not specified	7

Table L Main research fields of participants different from Medicine and Biology

Use of images in publications (P7 – P11)

P7

89% of the participants declared to use images in publications regularly (Figure 73).

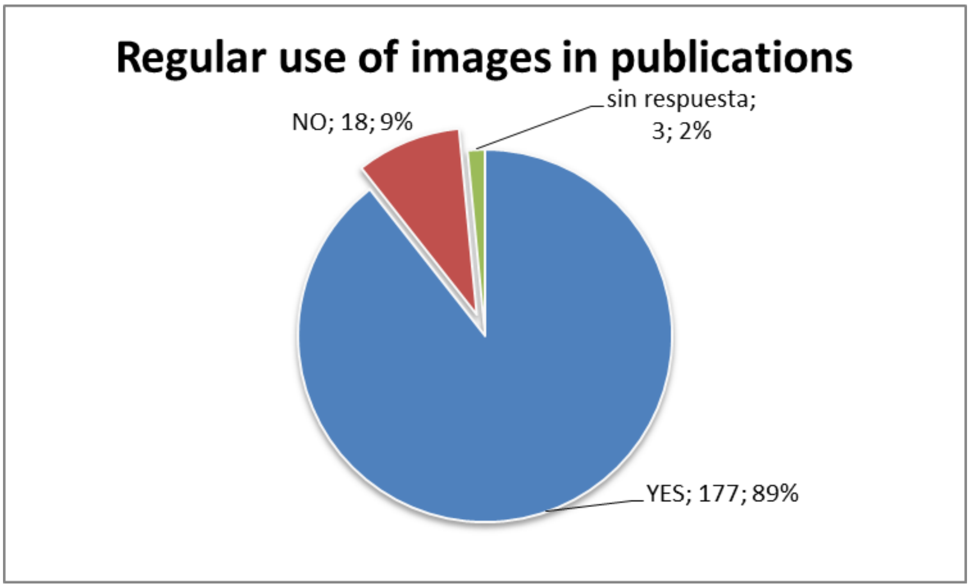


Figure 73 Percentage of participants that regularly use figures in their publications

P8

Most cited publications declared by participants to be illustrated by images regularly were articles, teaching materials and conference proceedings (Figure 74 - Table LI).

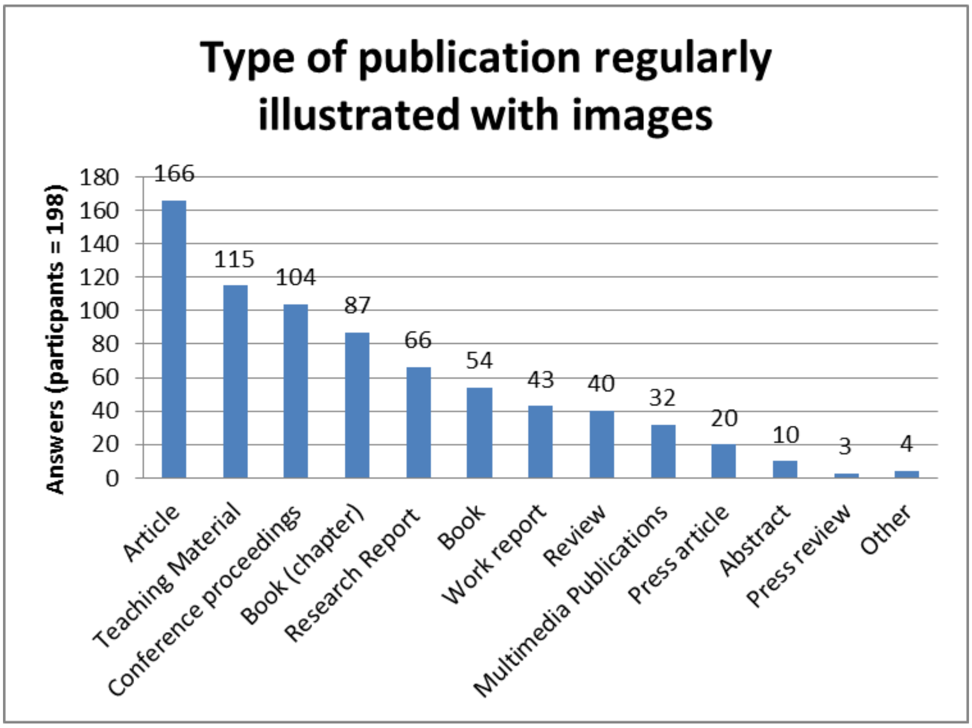


Figure 74 Type of publication regularly illustrated with images by the participants. "Other" includes two posters, one web and one clinical session.

Which type of publication do you regularly illustrate with images?	Answers
Article	166
Teaching Material	115
Conference proceedings	104
Book (chapter)	87
Research Report	66
Book	54
Work report	43
Review	40
Multimedia Publications	32
Press article	20
Abstract	10
Press review	3
Other	4

Table LI Publications regularly illustrated with images by the participants

P9

This Images were used in publications mainly to illustrate the content of the article or because the visual representation was a way to present information better than tables or in-text textual descriptions. Other reasons for using images were to support the argument by a visual example, to emphasize essential detail and to show numerical data (Figure 75- Table LII). Two answers included in the "other" option emphasized the role of images as a mean that allow explaining concepts easily, avoiding the use of large paragraphs of text. One answer specified that the figures of structures in Physiology display the function of the organism better than text.

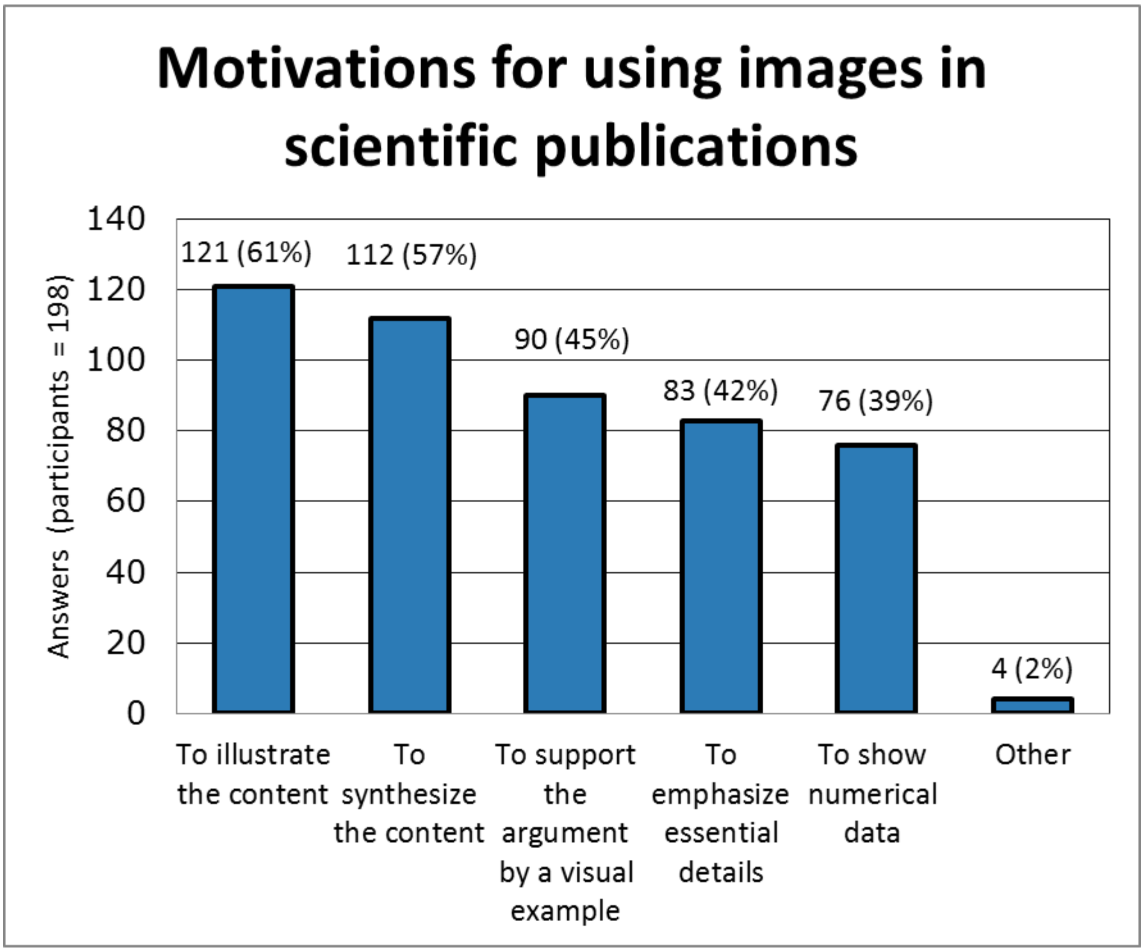


Figure 75 Function of images in academic articles according to the participants' answers.

Motivations for using images in scientific publications	Answers
To illustrate the content of the article	121
Information can't be easily or clearly presented in other formats such as a table or an in-text description	112
To support the argument by a visual example	90
To emphasize essential details	83
To show numerical data	76
Other	4

Table LII Function of images in academic articles according to the participants' answers.

P10

Most used images in medical publications were graphs (which include other categories of images such as charts, diagrams and illustration) and photographs. A high variety of typologies among the medicine-related image were used, as showed in Figure 76 - Table LIII. “Other” category included drawings, Northern Blot, Western Blot.

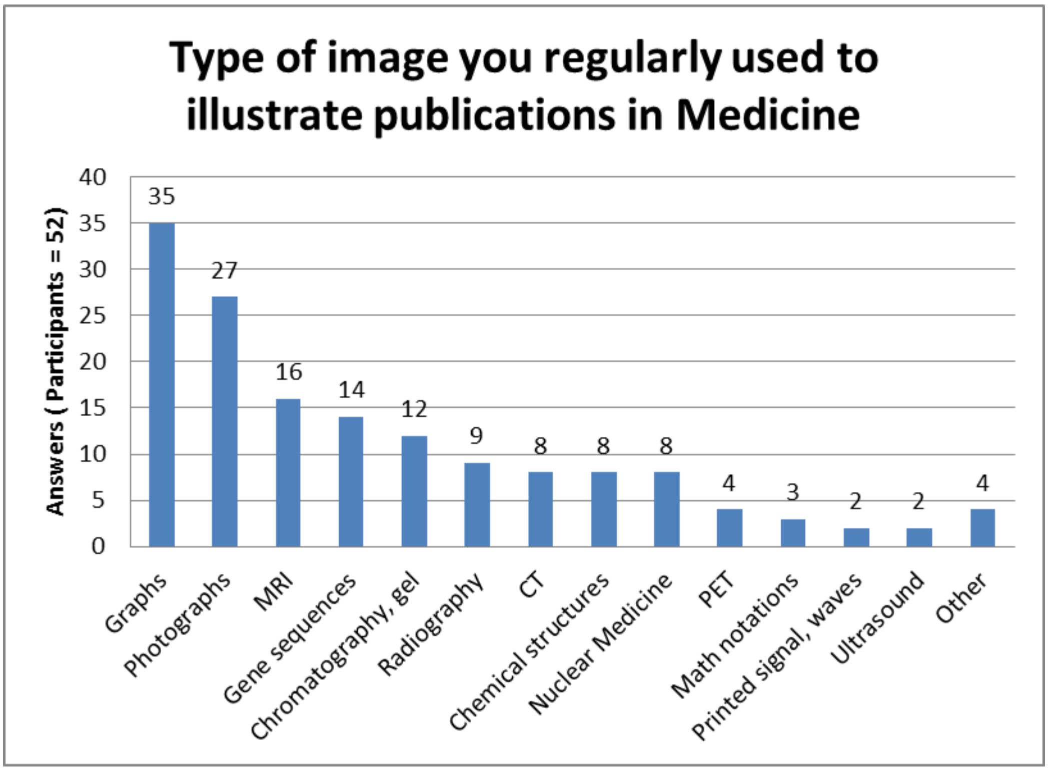


Figure 76 Most used types of image regularly included in medical scientific publications according to the participants belonging to the medical field.

Types of image regularly used to illustrate medical scientific publications	Answers
Graphs	35
Photographs	27
MRI	16
Gene sequences	14
Chromatography, gel	12
Radiography	9
CT	8
Chemical structures	8
Nuclear Medicine	8
PET	4

Math notations	3
Printed signal, waves	2
Ultrasound	2
Other	4

Table LIII Most used types of image regularly used in medical scientific publications

Most used images in publications in Biology area were graphs (which include other categories of images such as charts, diagrams and illustrations), photographs (not including microscopic images) and images representing cell types (Figure 77- Table LIV). “Other” included images such as gel, morphology of animals and organism, Immunohistochemistry, metabolism, fluorescent protein.

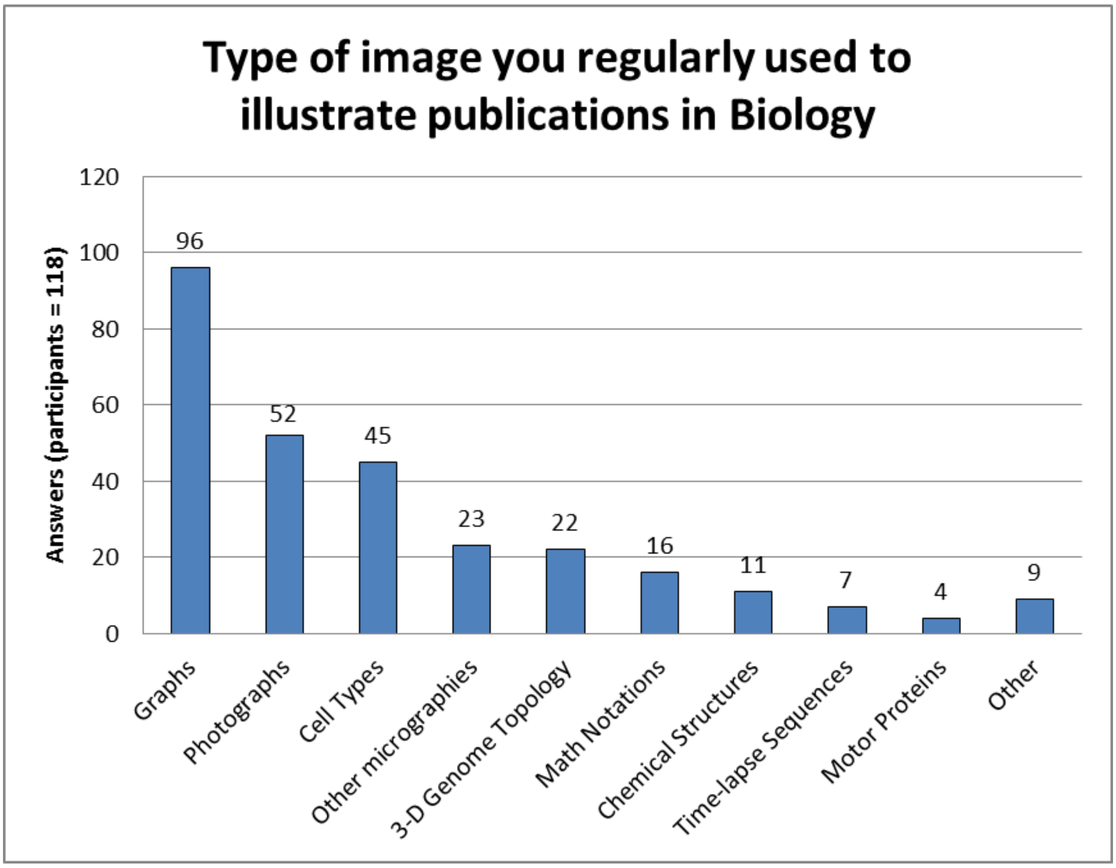


Figure 77 Most used types of image regularly included in scientific publications in biology according to the participants belonging to the medical field.

Types of image regularly used to illustrate scientific publications in biology	Answers
Graphs	96
Photographs	52
Cell Types	45

Other micrographies	23
3-D Genome Topology	22
Math Notations	16
Chemical Structures	11
Time-lapse Sequences	7
Motor Proteins	4
Other	9

Table LIV Most used types of image regularly used in biological science publications

P11

Most used image formats were JPEG and TIFF, as Figure 46 and Table LV show. "Other" category includes Adobe Illustrator Format (AI) and formats not specified. No one declared to use DICOM format.

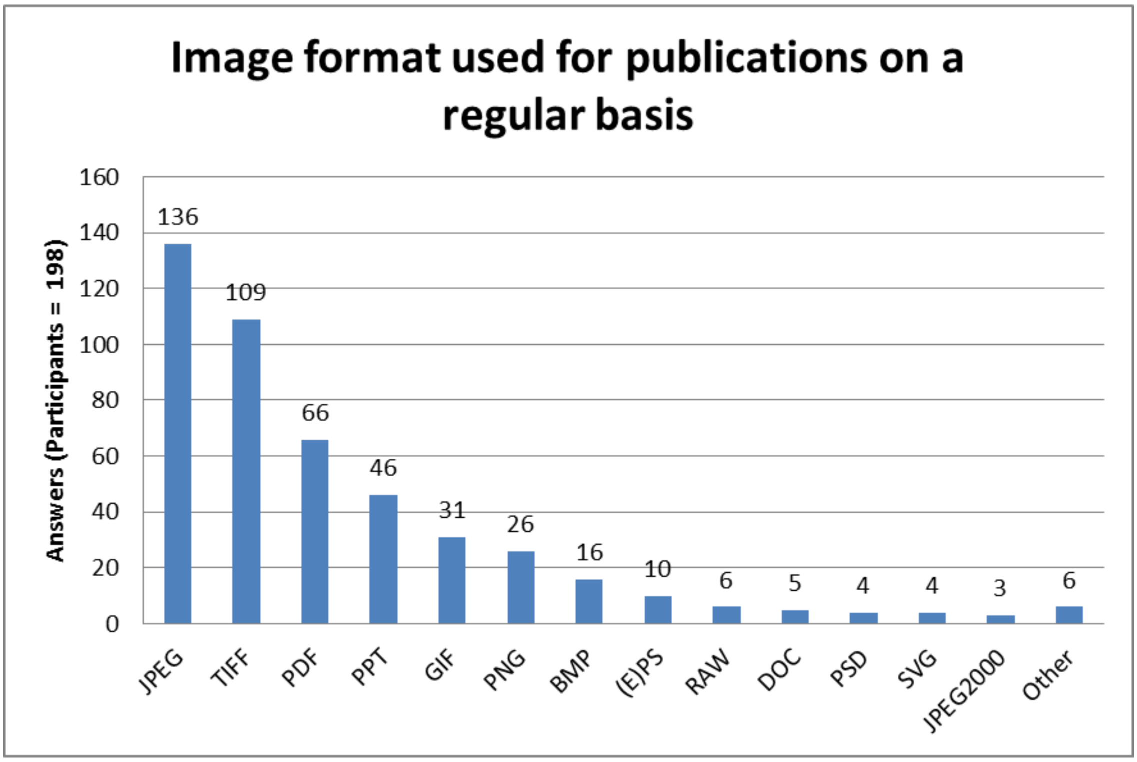


Figure 78 Most used image formats. Other includes Adobe Illustrator Format (AI) and formats not specified.

Which image format do you use on a regular basis?	Participants
JPEG	136
TIFF	109
PDF	66

PPT	46
GIF	31
PNG	26
BMP	16
(E)PS	10
RAW	6
DOC	5
PSD	4
SVG	4
JPEG2000	3
Other	6

Table LV Most used image formats. Other includes Adobe Illustrator Format (AI) and formats not specified

Participants were asked to specify the reason for using each format. The most cited reasons for choosing a specific format were:

- It is required/recommended by publishers for the academic paper submission.
- It has good balance between quality and compression (such as JPEG) or it maintains high quality (such as TIFF).
- It is commonly adopted, generated by different software and it is portable.
- It is useful for a specific action: image manipulation, sharing, reviewing, etc.
- It is useful for a specific type of image: photo, diagram, vector graphic, drawings, chemical structures, math notations, microscope images, etc.
- It is useful in a specific context: in a paper for an academic publication, in a presentation as teaching material, etc.

Image retrieval (P12 – P14)

P12

Most of participants searched images for publications monthly (Figure 79- Table LVI Frequency of search for images used in publications.

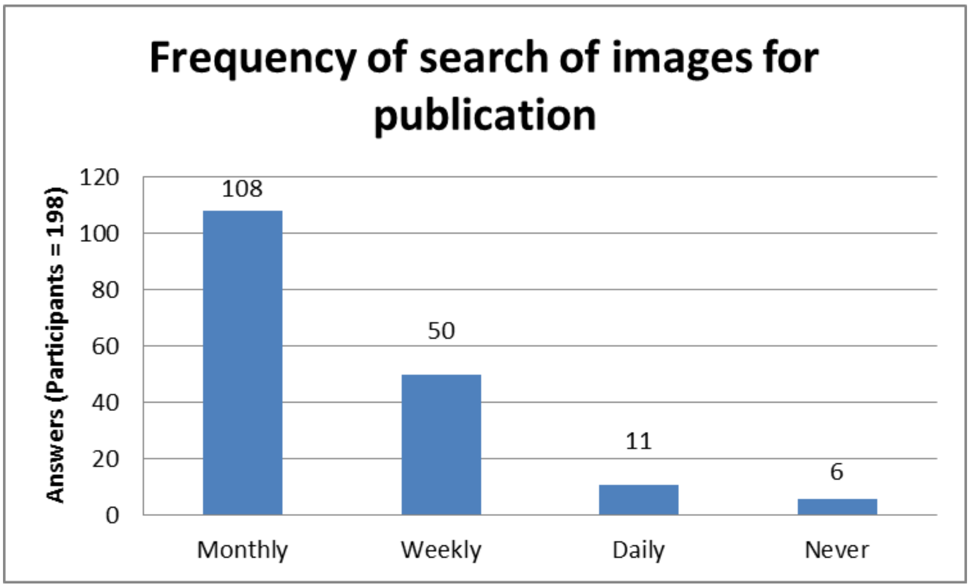


Figure 79 Frequency of search of images used in publications

Frequency of search of images for publication	**Answers**
Monthly	108
Weekly	50
Daily	11
Never	6

Table LVI Frequency of search for images used in publications

P13

For image search and retrieval needs, visual annotations - such as arrows, lines and Regions Of Interest - and textual annotations - such as symbols and letters – are relevant for half of the participants (Figure 80 Table LVII). “Other” category includes answers such as “I don’t look for images, I create them”.

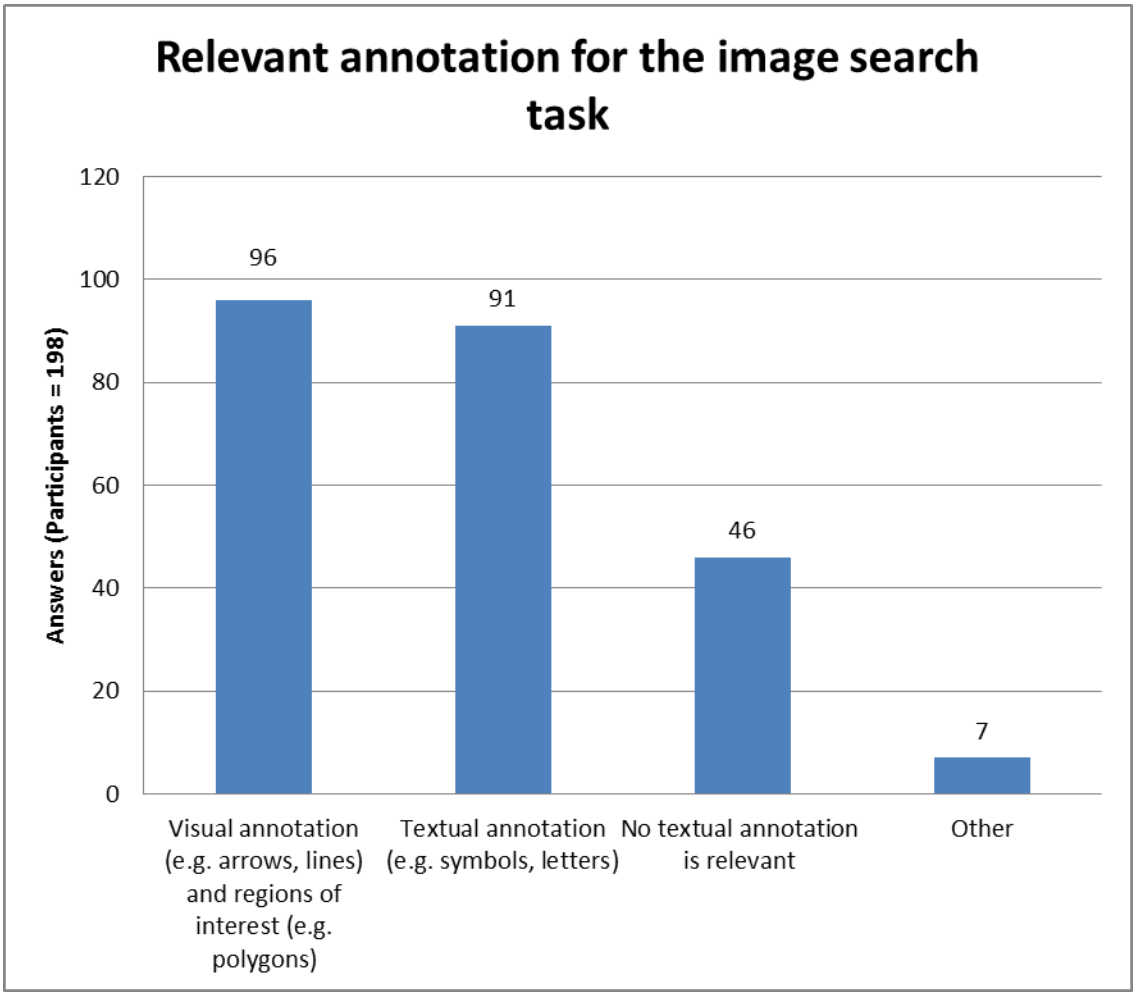


Figure 80 Annotations relevant for the image search according to the participants.

Relevant annotation for the image search task	**Answers**
Visual annotation (e.g. arrows, lines) and regions of interest (e.g. polygons)	96
Textual annotation (e.g. symbols, letters)	91
No textual annotation is relevant	46
Other	7

Table LVII Annotations relevant for the image search according to the participants.

P14

Most used resource for image search are academic articles (108 responses), personal collections (91 responses) and web based resources (82 responses) (Figure 81 - Table LVIII). The "Other" category includes answers in which participants declared to create and generate images themselves or answers in which any resource in particular were specified.

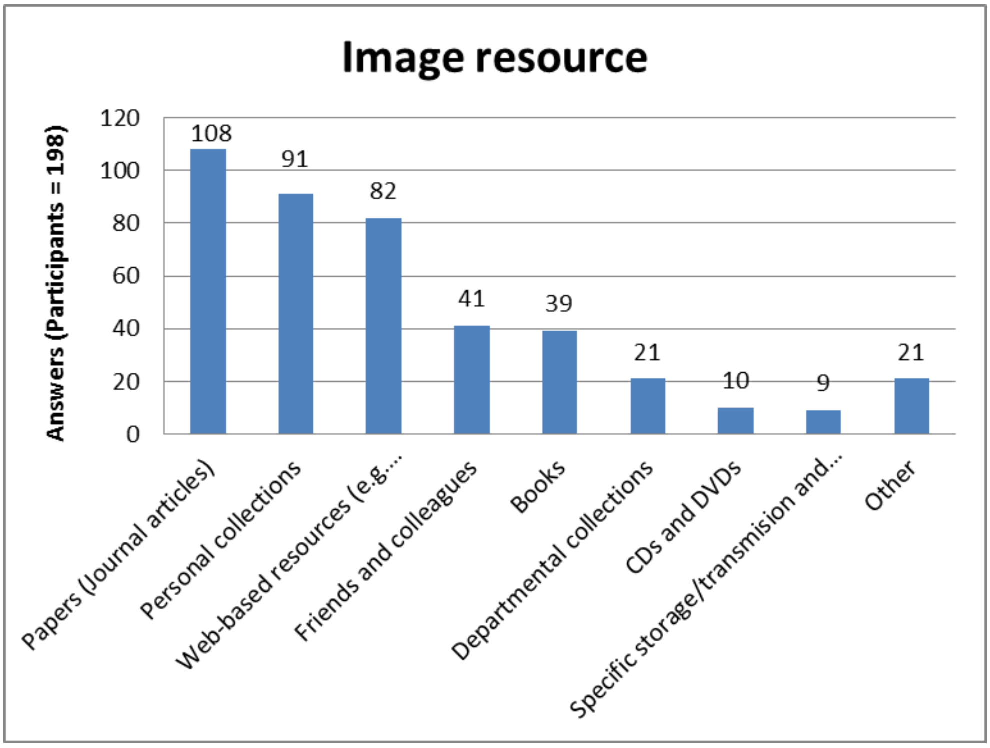


Figure 81 Relevant resources of images used by the participants. "Web-based resources" includes web services such as Google Images. "Specific storage…" includes digital image download systems, such as PACS.

Image resource	Answers
Papers (Journal articles)	108
Personal collections	91
Web-based resources (e.g. Google images)	82
Friends and colleagues	41
Books	39
Departmental collections	21
CDs and DVDs	10
Specific storage/transmission and digital image download (such as PACS servers)	9
Other	21

Table LVIII Relevant resources of images used by the participants.

Most used web-based resource for image retrieval was Google Images (Figure 82 - Table LIX).

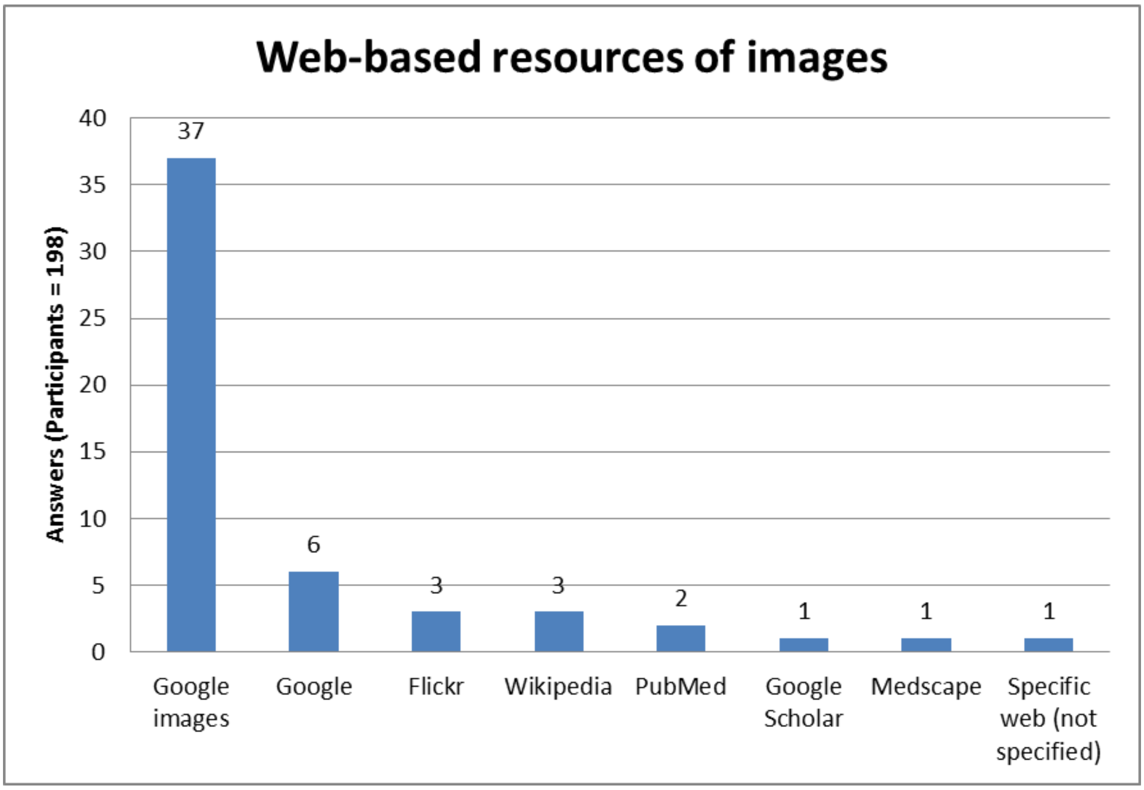


Figure 82 Web-based resources used by participants for image search

Web-based resources (e.g. Google images)	Participants
Google images	37
Google	6
Flickr	3
Wikipedia	3
PubMed	2
Google Scholar	1
Medscape	1
Specific web (not specified)	1

Table LIX Web-based resources used by participants for image search

Image-related text in reading (p15)

P15

During the reading of scientific articles, most relevant information source related to images was image caption, chosen by the great majority of the participants (Figure 50 - Table LX). Image annotations and image title are also considered relevant image-related text for more than half of the participants. “Mentions in the article” were the options choosen by 70 participants. Only 23 participants considered the technical information related to the image (such as DICOM metadata) as relevant information for understanding the image. “Other” category includes answers such as “the same image”.

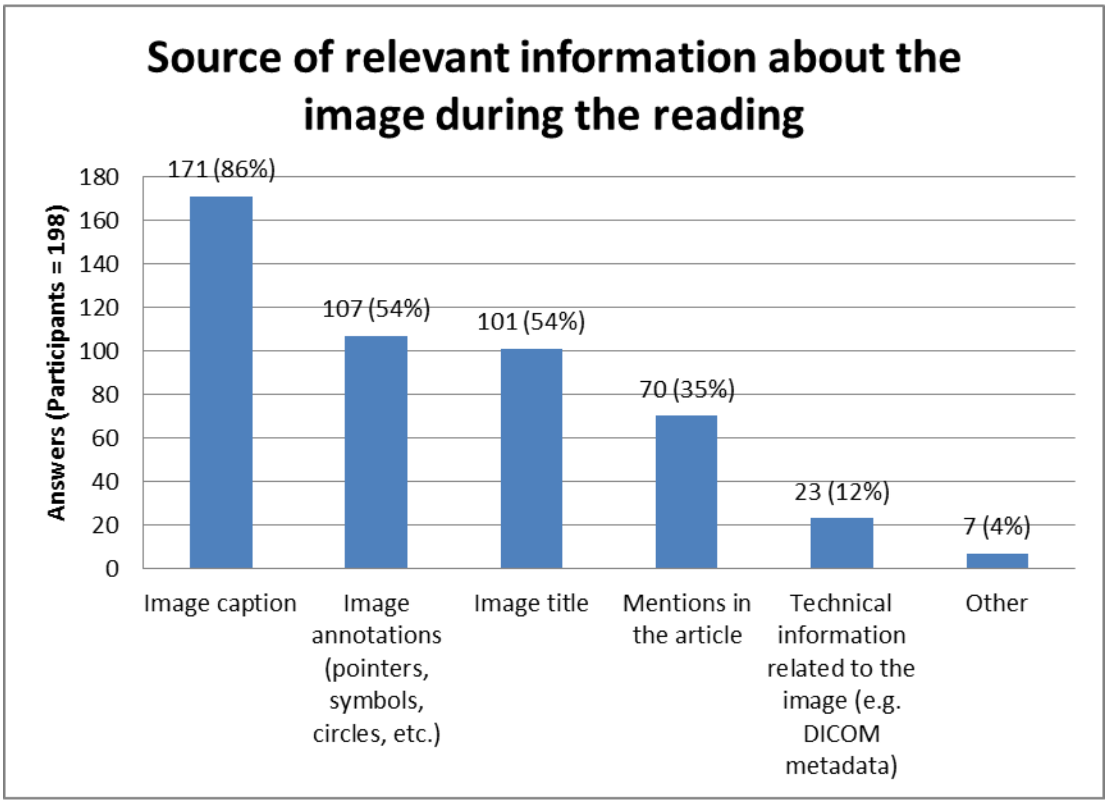


Figure 83 Textual elements related to images in scientific articles considered a source of relevant information of the image content during the reading of academic articles by the participants.

Source of relevant information about the image during the reading	Answers
Image caption	171
Image annotations (pointers, symbols, circles, etc.)	107
Image title	101
Mentions in the article	70
Technical information related to the image (e.g. DICOM metadata)	23

None of the previous answers	3
Other	4

Table LX Textual elements related to images in scientific articles considered a source of relevant information of the image content.

Use of textual descriptions of the images (p16 – p21).

P16

More than half of the participants declared to use visual annotation and textual annotations in combination with images in publications regularly (Figure 49 - Table LXI).

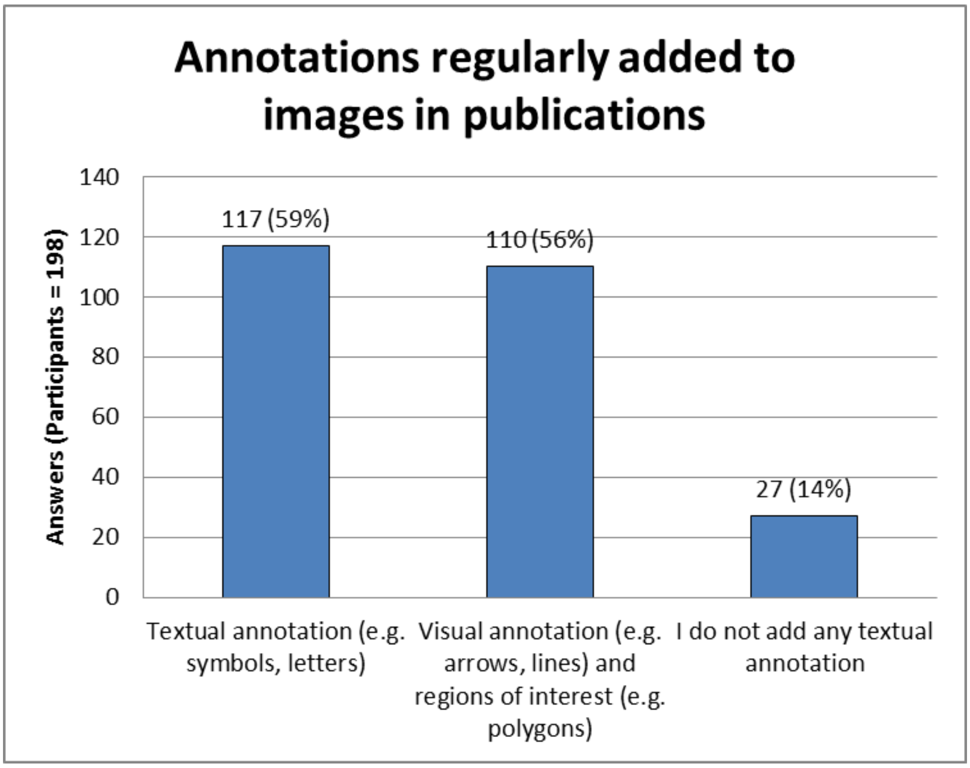


Figure 84 Annotations regularly added by the participants to the images used in publications

Annotations regularly added to in images in publications	Answers
Visual annotation (e.g. arrows, lines) and regions of interest (e.g. polygons)	117
Textual annotation (e.g. symbols, letters)	110
I do not add any textual annotation	27

Table LXI Annotations added regularly by participants to the images used in publications

P17

72% of the participants declared to describe images used in academic article by a related description text (Figure 48 - Table LXII).

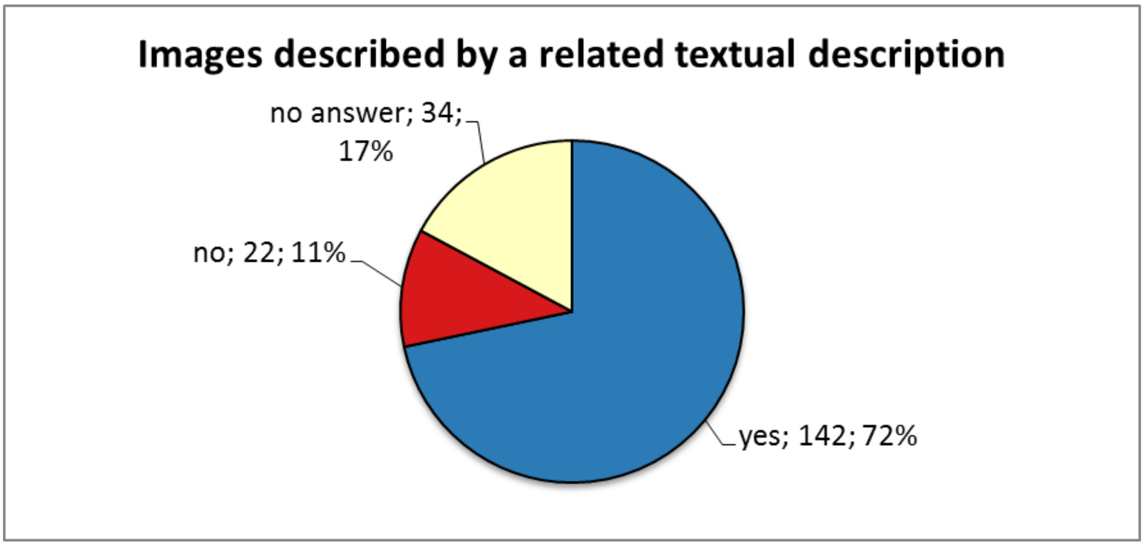


Figure 85 Percentage of participants that associate a description text to the images they use in academic articles

Do you describe images you use in academic articles by a related textual description?	Answers
Yes	87
No	13
No answer	13

Table LXII Participants that associate a description text to the images they use in academic articles.

P18

Among the participants that declared to add a textual description to images (142), half of them declared they usually create the description when they send the image to a publisher for a scientific paper submission (Figure 86 Table LXIII). 61 participants declared to create the description at the moment of the creation of the image. 35 participants said to create the description during the process of image editing. “Other” include the answers “When I write the article” and “Image and its description are separated”.

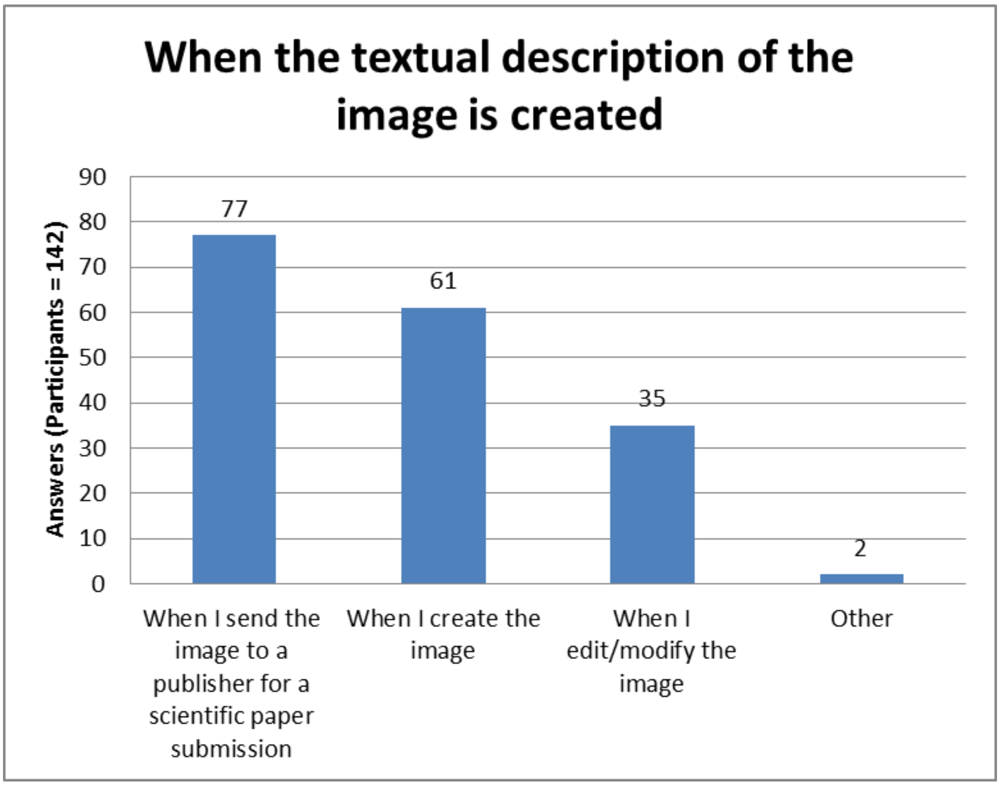


Figure 86 When the participants create a textual description of the image

When the participants create a textual description of the image	Answers
When I send the image to a publisher for a scientific paper submission	77
When I create the image	61
When I edit/modify the image	35
Other	2

Table LXIII Time when participants create a textual description of the image

P19

Among the participants that declared not to add a textual description to image (22), half of them did not add a textual description because images are already described in the article and in the caption. Other reasons are detailed in Figure 87 - Table LXIV . "Other" category includes the answer "I don't write articles".

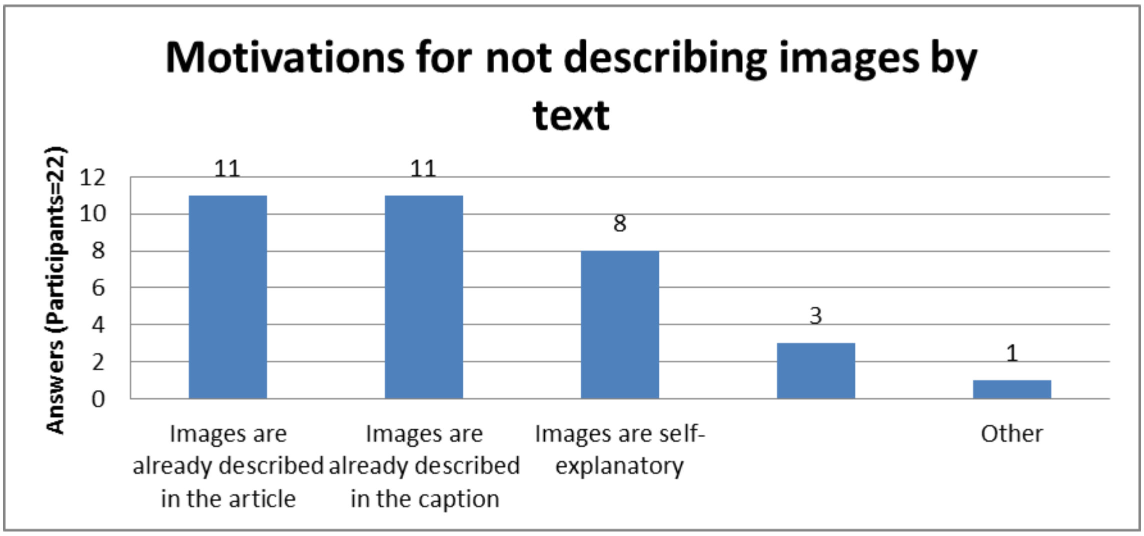


Figure 87 Motivations for not describing images by text according to the participants.

Why don't you describe images by textual information?	Answers
Images are already described in the article	11
Images are already described in the caption	11
Images are self-explanatory	8
There is no reason to describe it by textual information	3
Other	1

Table LXIV Motivations for not describing images by text according to the participants

P20

The most used tools selected by participants for image editing and annotation were Microsoft PowerPoint and Adobe Photoshop (Figure 47 - Table LXV).

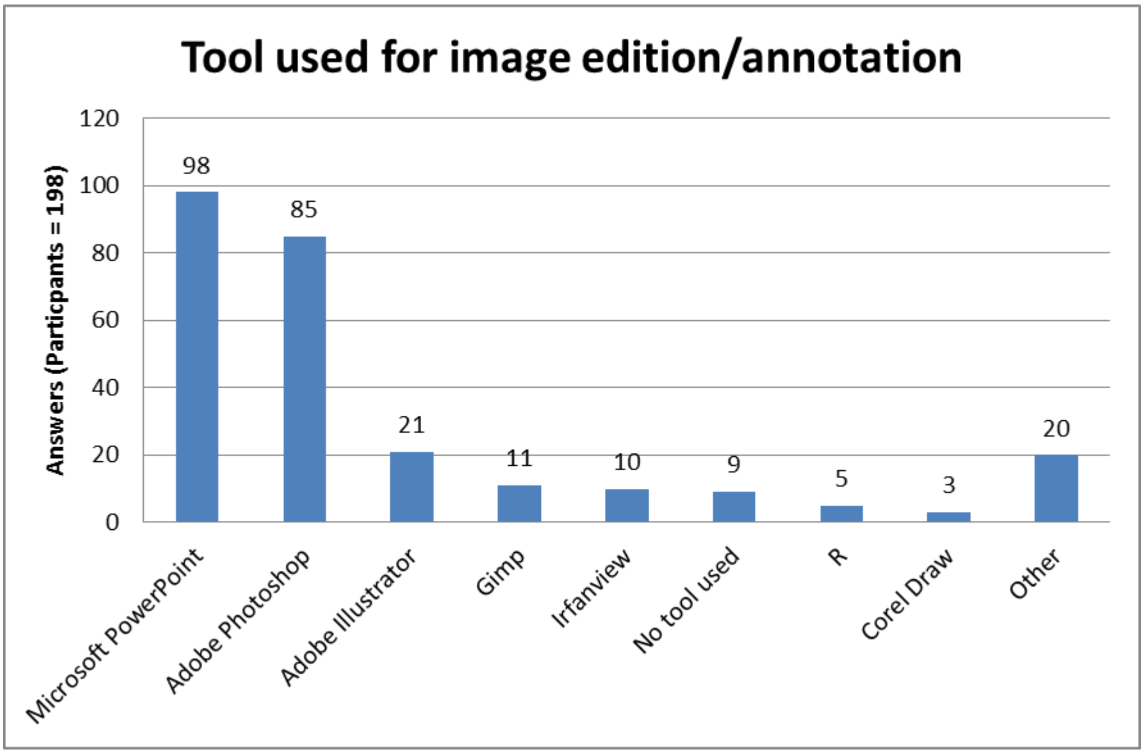


Figure 88 Tools used by participants for image edition.

Tool used for image edition/annotation	Answers
Microsoft PowerPoint	98
Adobe Photoshop	85
Adobe Illustrator	21
Gimp	11
Irfanview	10
No tool used	9
R	5
Corel Draw	3
Other	20

Table LXV Tools used for image edition by participants

"Other" tools are detailed in Table LXVI.

Tool	Participants
Canvas	2
Corel PhotoPaint	2
InkScape	2

Keynote	2
Prism	2
Corel	1
Corel Designer	1
Corel Presentation	1
Excel	2
Microsoft Expression Design	1
FastStone Image Viewer	1
FIJI	1
Graph Pad	1
ImageMagik	1
LibreOffice Draw	1
LightRoom	1
Microsoft Office Picture Manager	1
Picasa	1
SPSS	1

Table LXVI Other tools used for image edition by participants

P21

Participants were asked sending an example of a typical image they use in their publications with a title and a description of the image. 69 images were sent. The following Table LXVII shows the results of the word frequency analysis of the images' titles.

Word frequency in Titles	
Word	**Count**
figure	6
pereopod	4
figura	3
domain	2
estimates	2
expresión	2

months	2
results	2

Table LXVII Word frequency in the images' titles (minimum frequency = 2)

10 titles (out of 55 titles) included the word "Fig", "Figure" or "Figura". Among them, 5 title corresponds to the word "Figure" followed by a number (Figure 89).

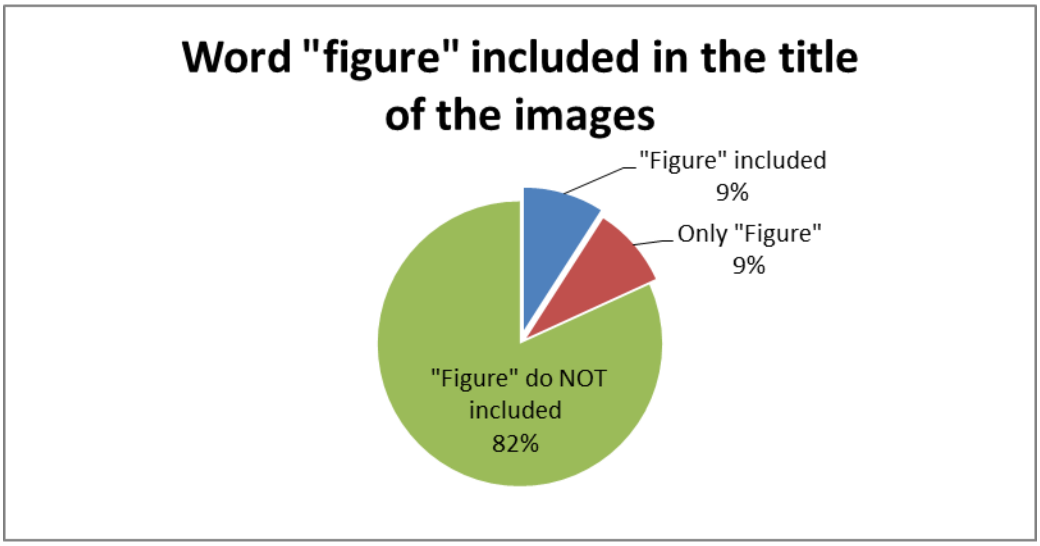


Figure 89 Use of the word "Figure" in the titles of figures sent by the participants

Table LXVIII shows the results of the word frequency analysis of the images' descriptions.

Word frequency in Descriptions	
Word	**Count**
htda	10
cells	7
tern	6
aag1	5
drr27	5
gull	5
indicated	5
little	5
right	5
show	5
strains	5

view	5
area	4
audouin	4
bars	4
co	4
cultures	4
dna	4
fig	4
figure	4
grey	4
left	4
mean	4
nyfp	4
pcr	4
que	4
r27	4
seconds	4
shown	4
white	4

Table LXVIII Word frequency in the images' descriptions (minimum frequency = 4)

In the description field, two participants declared that the caption was the description of the image. One participant declared that description was not necessary.

Table LXVIII shows the results of the word frequency analysis of the images' file name.

Word	**Count**
figure (figure1, fig, figura, figure7, figures)	**23**
pdf	**21**
tif	**20**
jpg	**13**
png	**7**
15	**2**

2011	**2**
2013	**2**
doc	**2**
ejemplo	**2**
eps	**2**
imagen1	**2**

Table LXIX Word frequency in the images' descriptions (minimum frequency = 2)

Figure 90 - Table LXX show the file extensions of the sample images sent by the participants.

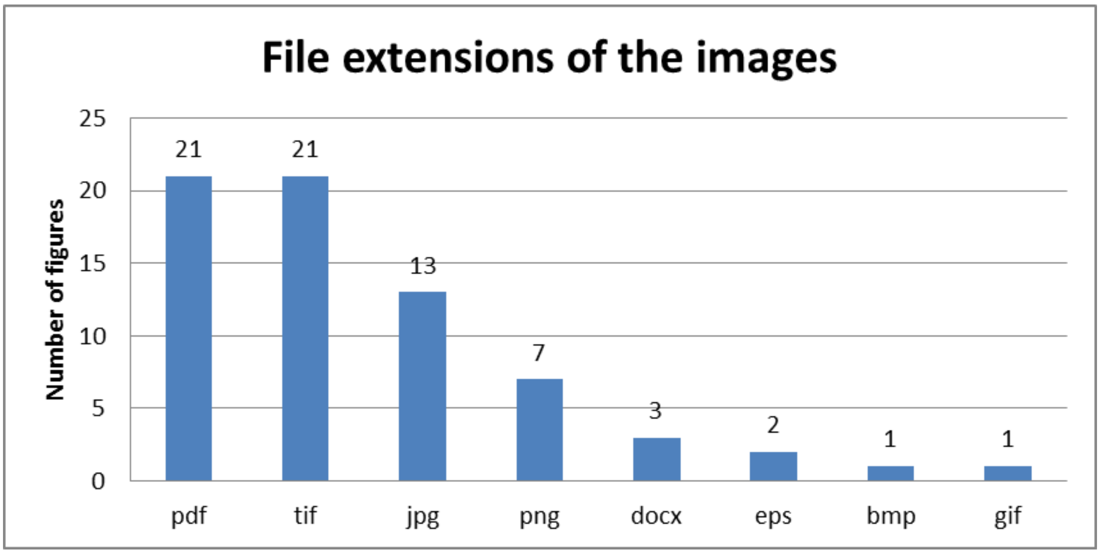


Figure 90 File extensions of images sent by the participants

Image file Extension	Number of figures
pdf	21
tif	21
jpg	13
png	7
docx	3
eps	2
bmp	1
gif	1

Table LXX File extensions of images sent by the participants

Figure 91 - Table LXXI show the file size of the sample images sent by the participants. The size of half of the images sent was below 500 KB.

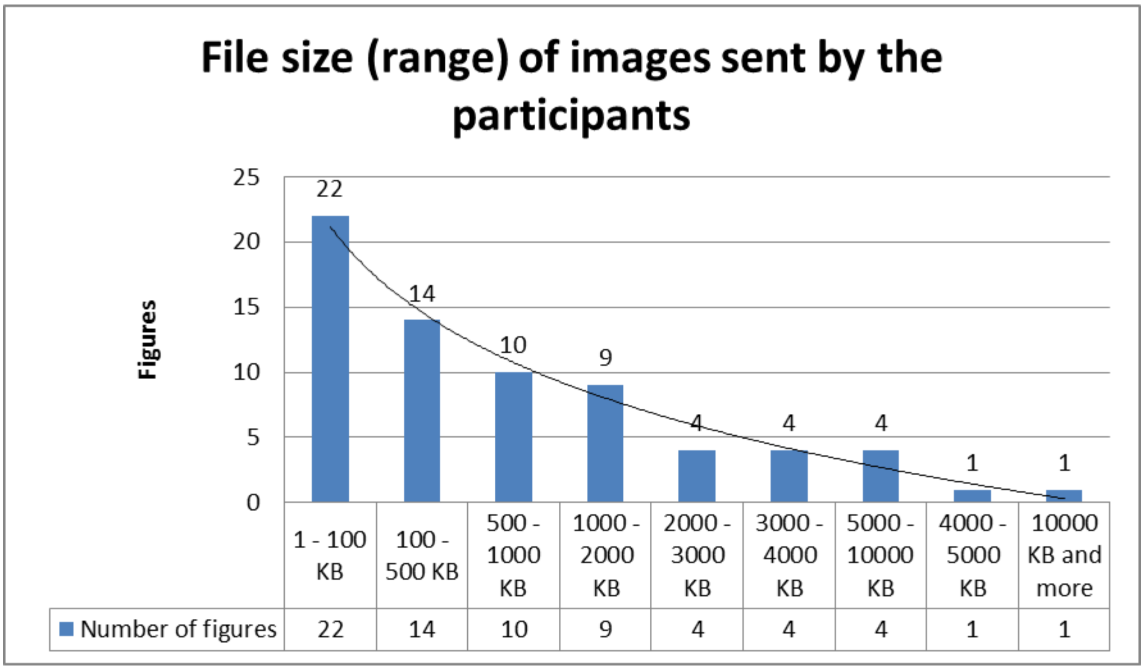


Figure 91 File size (range) of images sent by the participants

File Size (range)	Number of figures
1 - 100 KB	22
100 - 500 KB	14
500 - 1000 KB	10
1000 - 2000 KB	9
2000 - 3000 KB	4
3000 - 4000 KB	4
5000 - 10000 KB	4
4000 - 5000 KB	1
10000 KB and more	1

Table LXXI File size (range) of images sent by the participants

Accessibility-related issues in image publications (P22 – P25)

P22

Only 15% of the participants declared to know the meaning of “alternative text” for images (Figure 52 - Table LXXII).

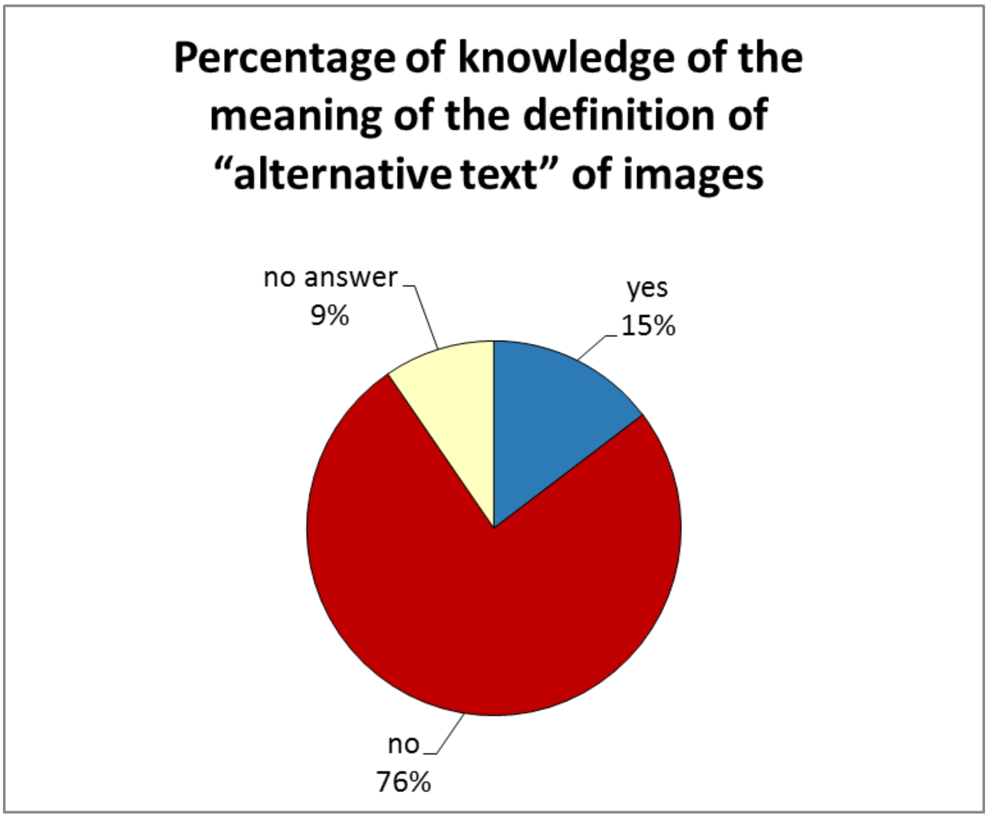


Figure 92 Percentage of participants that declared to know the meaning of “alternative text” for images

Do you know the meaning of “alternative text” for images?	Participants
yes	29
no	150
No answer	19

Table LXXII Number of participants that declared to know the meaning of “alternative text” for images

P 23

Only 5% of the participants declared to be aware of the accessibility policy for people with visual impairments in scientific journals (Figure 51 - Table LXXIII).

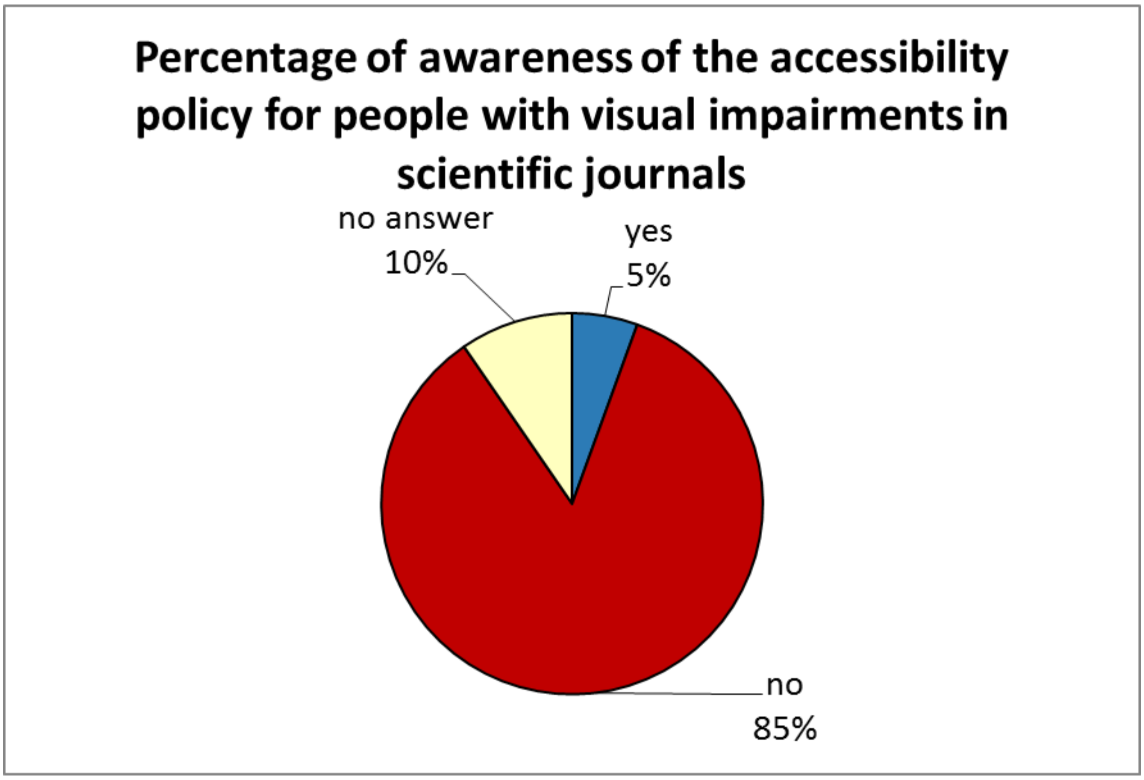


Figure 93 Percentage of participants that declared to be aware of the accessibility policy for people with visual impairments in scientific journals.

Are you aware of the accessibility policy for people with visual impairments in scientific journals?	**Participants**
YES	11
NO	168

Table LXXIII Participants that declared to be aware of the accessibility policy for people with visual impairments in scientific journals.

P24

32% of the participants declared to have experienced limitations in accessing the visual content of images in scientific articles (Figure 53 - Table LXXV).

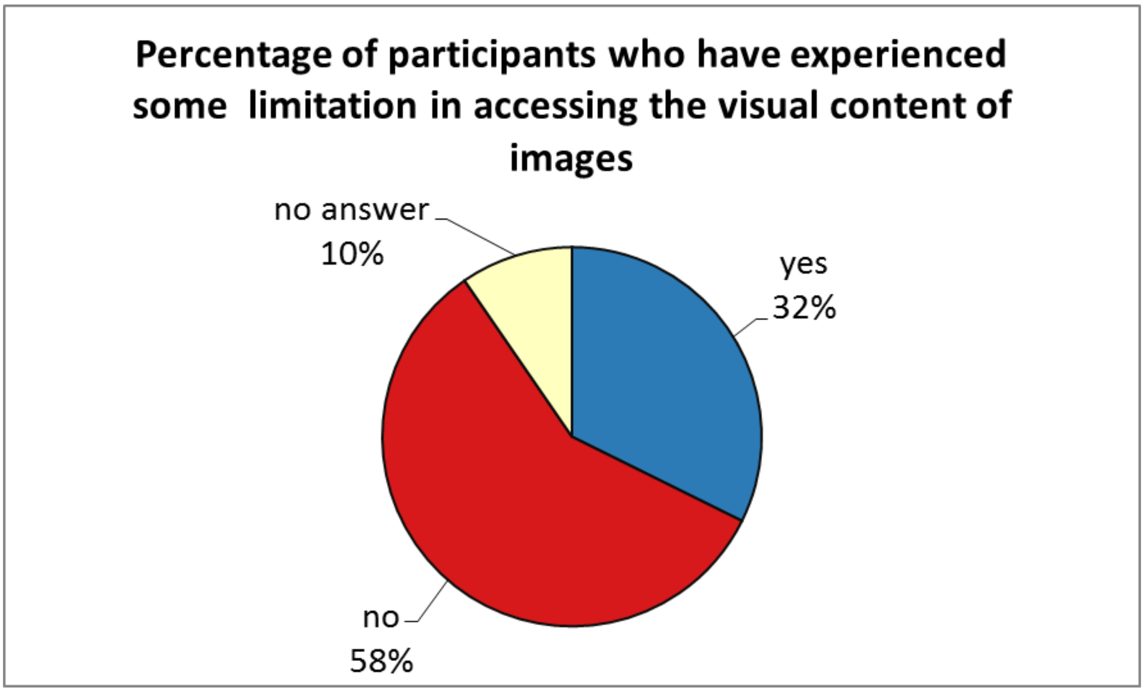


Figure 94 Percentage of participants that declared to had experienced limitations in accessing the visual content of images in scientific articles

Percentage of participants who have experienced some limitation in accessing the visual content of images	**Participants**
yes	64
No	115
No answer	19

Table LXXIV Number of participants that declared to had experienced limitations in accessing the visual content of images in scientific articles

P25

The participants that experienced limitations in accessing the visual content of scientific articles were asked to specify the type of limitation they experienced. The most cited limitations (almost the half of the total citations) were the low quality of the image, the low resolution of the image and the small font of textual annotations inside the image. Other relevant limitations cited were the small size of the figure and blurring effect. Also the incorrect use of color and contrast were cited as problems in accessing images' content. One participant declared to be color-blinded and to have limitations in distinguishing among colored lines in line graphs. The limitations cited are detailed in Figure 54 - Table LXXV

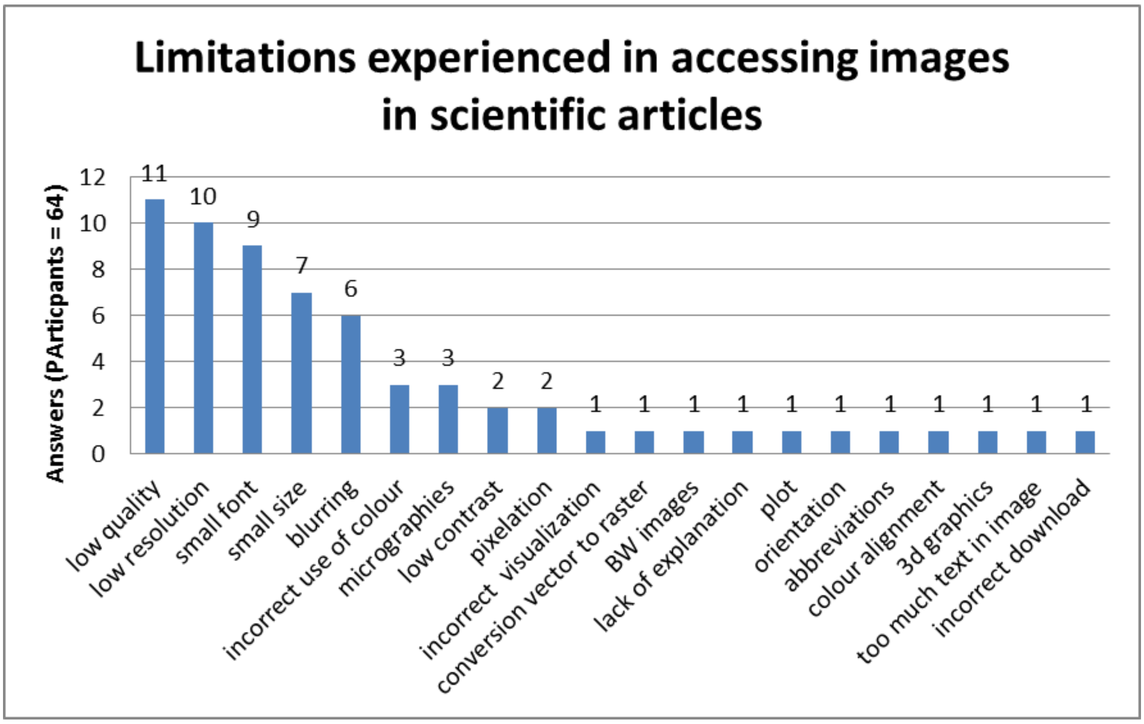


Figure 95 Limitations experienced in accessing the visual content of images in scientific articles according to general categories (minimum percentage of citations = 1%).

Limitations experienced in accessing the visual content of images in scientific articles	**Answers**
low quality	11
low resolution	10
small font	9
small size	7
blurring	6
incorrect use of colour	3
micrographies	3
low contrast	2
pixelation	2
incorrect visualization	1
conversion vector to raster	1
BW images	1
lack of explanation	1
plot	1

orientation	1
abbreviations	1
colour alignment	1
3d graphics	1
too much text in image	1
incorrect download	1

Table LXXV Limitations experienced in accessing the visual content of images in scientific articles according to general categories.

List of tables

List of figures