

10. System Evaluation

10.1. Introduction

Evaluation of software and databases is a practice established for so long as that of developing software itself. (Martyn & Unwin 1986). Not only evaluating in the sense of exercising the code using examples, but also evaluating in the sense of making some development step and then reflecting on the results of the step to check whether they were really what were intended. The idea of evaluating as an integral part of the software and database development process which takes place at all stages in order to meet the needs of the users is adopted in this thesis.

This chapter presents the qualitative techniques used for the evaluation of the system, and the obtained results.

10.2. Evaluation in software development

Software evaluation entails running software products under known conditions with defined inputs and documented outcomes that can be compared to their predefined expectations.

Generally, there are two methods of evaluating any web based system: verification and validation. (Rakitin 1997)

Verification determines whether the software is built correctly and does not contain technical errors. Verification also involves the review of the requirements, to verify that the right problem is being solved. Verification ensures that the software is syntactically and logically correct and that it performs functionally as specified.

Validation, on the other had, involves the more deceptively difficult task of insuring that the meaning and content of the rules meet some carefully defined criteria of adequacy. Defining such criteria is the key to successfully conducting a validation procedure and demonstrating the level of acceptability of the system. A thoroughly formal validation consists of a process in which validation requirement and specifications are previously established in terms of some sort of prerequisite requirements and design specifications.

Verification and Validation issues are also clearly fundamental to Databases, since they concern establishing and maintaining the correctness of a Database, by ensuring that artifacts of a Database collectively provide a plausible representation of the universe of discourse and its behaviour.

10.3. Criteria for software evaluation

10.3.1. Criteria for software verification

The development of any new system requires firm foundations which, in this case, are the *System's requirements*. These requirements need to be expressed as correctly as possible to avoid any failings in the system that is to be produced. We refer requirement as a condition or capability needed by a user to solve a problem or achieve an objective. System's Requirements are a description of the environment in which the system will work, plus a list of requirements that the system will have to fulfil. The verification criteria are intended to prove that the System's Requirements cover as many requirements as are realistically within the scope of the system.

When talking about Databases, integrity is the most familiar verification. In natural language, integrity is the property of honesty, uprightness, or original perfect state (Collins Dictionary). In the DB world it has a weaker meaning, i.e., plausibility. **Integrity** constraints are rules or conditions, Boolean-valued functions that are required to be true, and thus limit the set of valid instances of a schema. They are defined to ensure that an instance is self-consistent, complete and plausible with respect to known generic properties of the universe of discourse.

Intensional **correctness** is also of concern for Databases and is addressed by methodologies for database design and schema evolution management. However, the focus has been on avoiding data errors and inconsistencies. The theory of relational databases includes a number of semantic features domains, keys, dependencies between the columns, attributes of tables relations, etc., which are a consequence of the meaning of the data represented, rather than the whims of the designer.

Normalization allows verification of database designs through application of design constraints, called normal forms. Normal forms are formulated in terms of table keys, and dependencies, functional, multivalued, join, template, etc., between columns. These constraints ensure the

removal of redundant data and thus avoid update anomalies. Update anomalies occur when an update creates inconsistent data or inadvertently removes data.

Referring to software, verification should be carried out through consistency, robustness and accuracy to analyze the correctness of the system.

A system is said to be **consistent** if repeated executions with the same data lead to the same results or conclusions.

Robustness is measured by testing rules and control strategy (inference) to extreme conditions, to determine which input parameters are least, and most, significant in the form of the interim and final results and output. At least one test should be developed for each of the documented system limits. Such tests are designed to investigate how the system reacts to data which is maximal or minimal in the sense of attaining some limit. During the system testing, the system should also be tested beyond the limits specified for it. The purpose here is to find any situation where insufficient safety margins have been built in.

Accuracy is measured by comparing the number of correct predictions with known data. This is to compare the conclusions made by rules with historic data, and to observe the correctness of the outcome and how these outcomes reflect reality.

10.3.2. Problems in software verification

The correctness of a Database is verified through Normalization. This is the last step of the creation of a database so it was already done as an integral part of it when designing the database.

When verifying a Database System, robustness is not possible to be measured. There are no extreme conditions because all the parameters to choose are previously defined by the system. In the 'Life cycle Document Management System', all the parameters such as Actor, Type of contract, Type of document, etc., should be selected from a list of defined aspects.

Accuracy is another aspect that cannot be measured in this Database System. The results obtained from the system are very subjective and can only be verified comparing them to the previous requirements of the system.

10.3.3. Criteria for software validation

Validation criteria can be either qualitative or quantitative. Qualitative criteria employ subjective comparisons of performance, while quantitative criteria employ statistical methods to compare system performance to real world or human expertise. Qualitative validation criteria do not mean informal validation. It is possible to develop a highly formal qualitative validation. Some common qualitative validation criteria are:

Generality: Generality is measured in terms of the capability of a system to be used with a broad range of similar problems. Generality is the range of contexts within which a system can be expected to perform. A system is broad when the range of conditions and contexts, within which it should operate reliably, is wide.

Adaptability: Adaptability is measured in terms of the capability of the system to be customized for a particular user needs and its possibility for future development and application. The system must also be adaptable to different work environments, and integrate with existing or proposed hardware or software systems.

Compatibility: Tests are made to probe where the new system does not subsume the facilities and modes of use of the old system where it was intended to do so. In other words, a search is made for system incompatibilities.

Visual Interaction: (or user interface) Visual interaction is measured by testing the user friendliness of the system and how the human user interacts with the system. The interface must be evaluated for ease of understanding and use, as defined by the quantified measures that will have been used in the *System Specification*.

Usefulness: Usefulness is proven by ensuring that the knowledge base contains necessary and adequate parameters and relationships for use in the problem-solving area.

Help information: Such help information must be sufficient to enable a naïve user to use the system. Test must be devised to generate help information which must then be evaluated for the appropriate level of detail and presentation. The *Users Guide* must be searched for areas that ought to have help information but where none is available, or where help exists but is not documented. The Users Guide must be reviewed for clarity, ease of use, details of all required system facilities and details of all system error messages, together with sufficient information to allow the rectification of errors.

Security testing: Tests are performed that attempt to contravene the system's security, such as the access of database-held by unauthorized users.

Maintenance: In some cases, a system will be specified to have particular maintenance requirements, for instance, those changes of a given level of complexity should take no longer than a certain time.

Installation: Here we are concerned with systems that can be installed on a variety of machines or models, or that allow a number of installation options such as different peripheral devices.

Storage: In the case of any system with specified storage requirements such as a maximum amount of main or backing storage occupancy, we need to devise test that seek to detect instances when the system exceeds the specified limits by, for instance, processing or supplying large amounts of data in a similar way to the volume tests.

Convenience: Convenience to the users or compatibility with the existing web software is a very important issue in the acceptance and usage of the prototype system. When the prototype system is designed, the convenience to the user must be carefully measured during the initial stage of development by designing a prototype system which uses simple facilities, does not require any complicated assembly, and does not require substantial effort on the part of the users to either carry or use.

Methods for performing these reviews are detailed in Fagan's paper (Fagan, 1976).

10.3.4. Problems in software validation

When validating the Database System, security will not be validated because for the moment it's only a prototype and there is free access to the system to allow all interested parties have access to it.

When validating a web based system, installation and storage is not needed. Users only have to access the web page with their correct password and download the needed information.

10.4. Methodology for System evaluation

Before validating the system in construction companies, the system was verified in the development environment to determine the accuracy of the robustness of the system and the correctness of the output provided by the system.

10.4.1. Methodology for System verification

Verification of contents:

After writing the *System Requirements* (see Chapter 9), the information and relations of the Database was checked and verified for any discrepancies and errors. This was achieved by the submission of the system data content for criticism and evaluation by several academics. The academics were asked to read and comment on the principles of organizing the flow of information into Phases/Stages, Activities/Subactivities, Document metadata, etc.

The following questions were used to verify the requirements to see that the right problem was being solved.

1. Is the system's scope well-defined?
2. Have all the users of the system been identified?
3. Are there any general areas omitted?
4. Are the System Requirements understandable?
5. Are Phases/Stages well defined?
6. Are Activities/Subactivities well defined?
7. Are Actors well defined?
8. Are types of contract well defined?
9. Are types of documents and their metadata well defined?
10. Are the relations between the previous aspects well defined?

Once the system was developed, technical errors and inconsistencies of the software were also verified through Normalization to remove redundant data and avoid anomalies.

Verification of consistency:

All parts of the system that were subsequently built had their consistency checked using sets of input data to test the logic.

10.4.2. Methodology for System validation

Informal validation by domain experts has been used to test the system. The validation focuses mainly on the performance issues specific to the design and application of the system.

Acceptance validation consists of checking whether the prototype system has reached a reasonable level of quality at the end of the development stage. Several qualitative validation criteria were chosen, rather than a single one, in order to provide an analysis, as independent and as exhaustive as possible.

Before validating the system in different construction companies, it was necessary to have information of the company referring to the size, use on IT technologies, WPMS, etc.

The following questions were used to get previous information of the company.

1. Type of company: Client/Designer/Contractor
2. Number of employees
3. Average number of participants in a project
4. Have you got any Quality System?
5. What type of repository of information do you have? Central/Local
6. Who creates the system organization (folder, archives, etc) of each project?
7. Steps when starting a project
8. Do you have a well structure file organization?
9. Do you have well defined formats?
10. Are you satisfied with the organization of documentation?
11. Have you ever used WPMS for the management of any project?
 - a. If Yes: Why? How many? Advantages and Disadvantages
 - b. If No: Why?
12. Have you ever used any DMS?
 - a. If Yes: What type? Advantages and Disadvantages
 - a. If No: Why?

For the validation of the 'Life cycle Document Management System', the survey was done part face-to-face and part giving the access to the system, the user's guide and the questions by mail. The questions used in the validation were:

1. What do you think about the system interface with the users?
2. Are the questions asked by the system comprehensible?

3. Did you find the explanation facility helpful?
4. Is the organization of files well structured?
5. Could it be useful for all the people of the company?
6. Could it be useful to organize the information to deliver in each phase/stage?
7. How can you valorate the system?
8. What do you miss?

Validation experimentation was conducted to examine the acceptability of the prototype system as a tool for improving the document management of construction projects.

10.5. System evaluation results

From the previously defined methodology the observed results and findings were the following:

10.5.1. System verification results

Verification of contents:

From the definition of the *System Requirements* (Chapter 9), the system's scope was verified. All the academics stated that the scope was well-defined but that it had to be clarified. They also stated that not all SMEs are in conditions of using the 'Life cycle Document Management System' but only those endowed with a minimum IT infrastructure such as interconnected PCs or server, Internet access, etc.

From their point of view all the general areas were identified and the system requirements were understandable.

The first proposal for document organisation was based on the life cycle of a project (phases and stages). All the information related to a specific area was grouped together in activities. In this first proposal, each stage had different activities with different names. Different academics suggested defining the same activities for all the stages and adjusting the definition of a document by introducing the 'subactivity' field. This proposal was considered and adapted after evaluating the differences between the initial idea and the proposal.

The terminology to be used in each phase/stage and activity/subactivity was also discussed. Some modifications were carried out but the terminology issue was solved by providing the definitions of all the fields for each aspect.

Referring to the metadata, and specially the 'type of document', several types of documents, such as 'Generic document', were added, so that they could be used when there is no other type of document that fits the specific information being stored.

In relation to the type of contract, all the academics thought it appropriate to define just three types of contract and then choose the one that best fits the real project.

The problem of defining actors or roles was discussed and we reached the conclusion that it was impossible to define all the possible roles in a project because each project is different and some kind of agreement should be reached.

These and other less relevant comments and suggestions were compiled and implemented wherever possible.

Verification of consistency:

This process was achieved by running the system more than 100 times each time using a mixture of inputs.

Each time the output was observed and the content of rules and their logic were changed accordingly until the system produced the intended results.

10.5.2. System validation results

Participants in the validation experiment were 30 SMEs from the construction sector with a number of employees from 3 to 200.

The participants were asked to use the system in hypothetical cases and feedback the experience. The points of the test were to submit the prototype system to criticism and to highlight what difficulties were encountered by the user.

Some companies were contacted by phone and the access to the system was given by e-mail.

Those companies who were contacted face to face gave higher punctuations of the system than those that were contacted by phone and e-mail. Some of the companies contacted by phone or mail didn't answer. The main reason might be that, due to their traditional way of working, they currently don't perceive the value of the system.

Taking into account those companies who accessed the system and answered the survey the following results were extracted:

The companies provided the following information:

The average number of employees taking part in a project was five, but each employee had his own role within the project.

Seventeen companies had fewer than 20 employees, eight of them had from 20 to 50 employees and only five had more than 50 employees.

The majority of the companies did not have quality assurance systems, but those that did, generally complied with the ISO 90002 standard.

Nearly all the companies had a central repository for information but this tended to be a simple server where they could store the information. Only a few companies had servers with templates, document structures, etc. based on 'Explorer' files.

Half of them had no folder structure templates; if employees need new folders, they have to create them themselves.

In terms of folder organisation or document management, those companies that had folder structure templates followed the steps below when starting a project:

1. Copy the template folder structure into the server or PC.
2. Use the document templates when necessary.
3. Store the documents in the folder structure.
4. Create the necessary folders.

Very few companies consider their files to be well organised and most are not satisfied with the organisation of documents within the company. The companies that were most dissatisfied with the organisation of information were those with fewest employees.

However, the majority of the companies felt that they had well-defined formats. They nearly always use and create the same type of documents, so they basically have predefined formats for these documents and for their working methods. Even if they do not have a certified quality assurance system, they work to their own standards of quality.

In reference to the use of WPMSs, 50% of the companies had used a WPMS and nearly all of them were designers. Of these companies, 50% had used a WPMS just once and 50% had used one twice; in general, they had been driven to do so by a party higher up in the value chain. None of these companies are currently using a WPMS in their work. Only two companies view the use of a WPMS as a future development within the firm and just three companies are aware of the benefits of similar tools. Design companies seem to be most interested in these kinds of tools.

The companies that have used a WPMS at least once consider the main advantage to be improved communication management, but for the moment they continue to use the telephone, fax and e-mail.

With regard to document management, they nearly always use paper documents in delivering projects or information. Some companies argued that they felt they were working in a very innovative manner because they use e-mail for communicating and exchanging documentation. These companies argued that the main disadvantages are the need for training and the problem of introducing new working practices in the office.

Those companies that have never used a WPMS argued that insufficient understanding and training are the main barriers to using them. They also argued that, for the moment, they have not been exposed to them.

In relation to DMSs, the majority have used one, but basically as a server within the company. Some of them (mainly those companies with more than 50 employees) also have an extranet that provides access to specific company information. Six of the companies have used a WPMS to store and exchange information with other stakeholders in a project, although only as pilot projects.

According to those companies who have used a DMS, the main advantages are efficient information access and document management capability. Although they have used these tools, training is still a big handicap for them.

The companies who have never used a DMS (12%) are companies with fewer than 10 employees, in which each employee has a PC containing his or her own data. Information is not centralised because each of the employees is in charge of different aspects of the company.

The findings indicate that it is architects and engineers who are most likely to use WPMSs, because they have the necessary infrastructure to support them and the employees are more familiar with technological solutions such as CAD. Subcontractors show the most resistance to

adopting these tools because they do not currently perceive them as adding value, they have not been exposed to them and they lack education about them.

Nearly all the companies centralise documentation in a server but they are not keen to have all their documentation on the Internet because of security considerations. AEC SMEs in Spain generally have document templates and they are satisfied with them, but the main problem is the organisation of documentation.

All the companies with more than five employees argued that they need to redesign their working processes and they need better document management. The majority do not use WPMSs but they are very interested in improving communication and document exchange with other stakeholders in projects; they argued, however, that they cannot afford to invest in IT tools.

Once the main information of the company was obtained, the validation of the *Life cycle Document Management System* was done. The following are the findings of the test:

Visual Interaction: User friendliness is the most important criteria for winning acceptance and overcoming any disincentives in using the system and in the acceptance of the system. With menus, windows and explanations, the prototype is easy to access and convenient to use. Screen colours, typefaces, figures, etc. were selected to provide an attractive design and visual interaction. Based on the input from some companies, the visual interaction was partially modified and improved but generally all companies found the system's interface with the user good. Initially, the information provided in each screen was difficult to understand, but after some changes, all the users agreed that they could understand the system easily. (The results for evaluating the visual interaction and user-friendliness were obtained from *Question 1. What do you think of the system interface with the users?*) Figure 43 shows the results of the survey. In general, all the companies agreed that the system had good visual interaction.

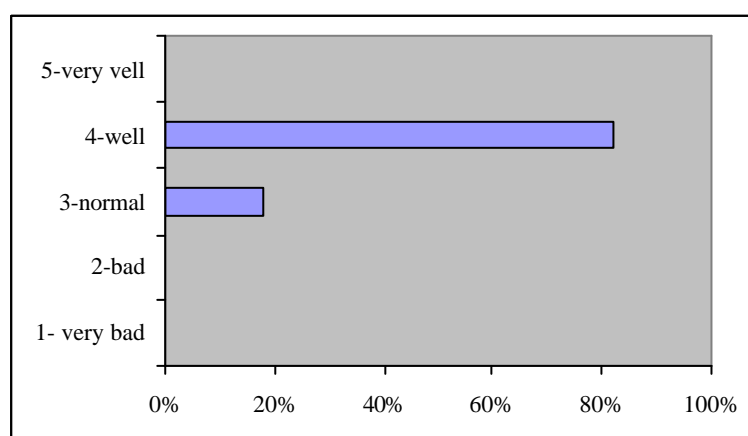


Figure 43 Answers of the survey for the validation of visual interaction

Help information: Two types of help information are provided in the system. On one side, each screen has *Instructions* to select data, download information, etc. All the users found these instructions very useful and sufficient to use the system. There is also a *User's Guide* with complementary information that can help users when they have a specific question. The guide was reviewed for clarity, ease of use, details of all system requirements and details of all system error messages, together with the provision of information that was sufficient for rectifying errors. (The results for evaluating the visual interaction and user-friendliness were obtained from *Question 2. Are the questions asked by the system comprehensible?* and *Question 3. Did you find the explanation facility helpful?* See Figure 44)

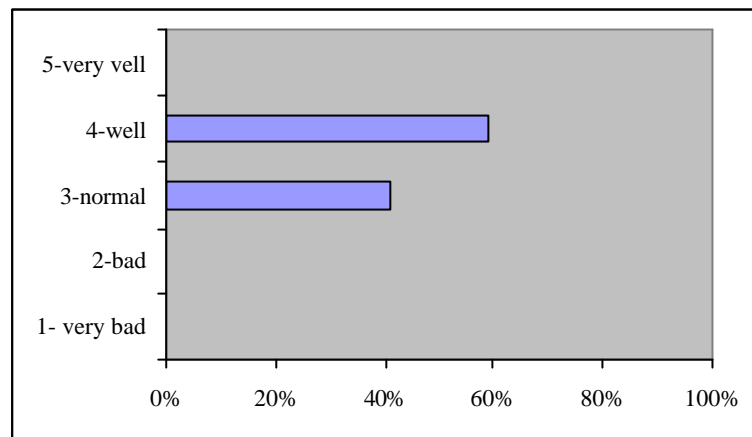


Figure 44 Answers of the survey for the validation of help information

Usefulness: Nearly all the companies found the system useful in their day-to-day work. For the moment, those companies with less than 10 employees do not feel the need for either such a system or Internet access. They work on small projects and with other companies similar to themselves, so they are used to communicating with each other face to face and delivering information in a paper-based format. Some of these companies were contractors that do not even have a server because their office work is basically reduced to one person. Companies with more than 10 employees found the system useful for managing their internal documentation, because they generally do not have a well-structured file organisation and they think this system could improve their internal management. At present they are worried about the system's security and they think interactions with other companies can be done using other means. Finally, those companies that have previously used a WPMS found it a necessary tool for the correct functioning of a WPMS. They argued that the main usefulness of the system lies in organising company information and providing the same folder structure as that of the WPMS. (The results for evaluating usefulness were obtained from *Question 5. Could it be useful for all the people in the company?*, *Question 6. Could it be useful to organise the information to be delivered in each phase/stage?* and *Question 7. How would you evaluate the system?* See Figure 45)

The results show that currently many companies think that the system might be useful for internal document organisation (Question 6). However, with regards to the organisation of

information to be delivered at each phase/stage (Question 7), they do not see the benefits for the moment, as they do not trust web services because of security issues and, in some cases, because of a lack of exposure. On the other hand, the majority of them evaluate the system positively (Question 8) and think that they could gain an advantage over other companies if they were to incorporate IT tools like this system at their company.

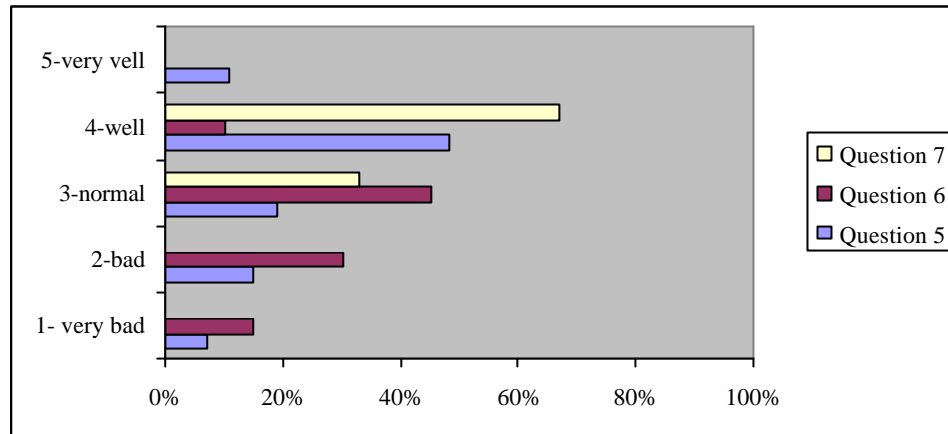


Figure 45 Answers of the survey for the validation of the usefulness

Generality and Adaptability: This was only measured for those companies who at least had a server on which to store information. They found that the input data (type of contractual arrangement, actor, etc.), the output information and the structure of the files covered all type of projects and could be used in any context. The only problem was trying to evaluate generality in companies that are not interested in document management and in improving their general management. (The results for evaluating the generality and adaptability were obtained from *Question 4. Is the organisation of files well structured?* and *Question 7. How would you evaluate the system?* See Figure 46)

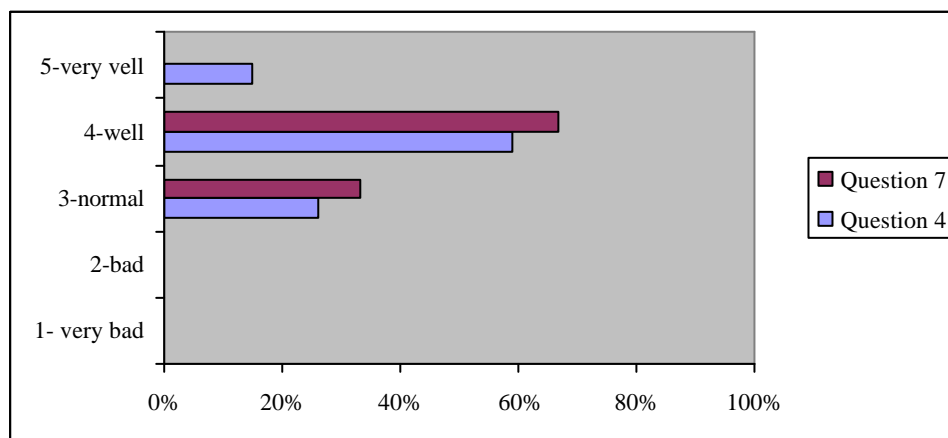


Figure 46 Answers of the survey for the validation of the generality and adaptability

Convenience and Compatibility: Convenience was also evaluated by contacting some WPMS developers and asking them how to export the folder structure created by the Life Cycle Document Management System. They revealed that it is technically very easy but we were unable to test it due to data protection measures. The system is intended to be used by all the companies taking part in a project and before using the WPMS. The user does not require any specific software to run the system and download the folder structure (just an Internet connection). Moreover, the folder structure is generated in 'Explorer Visualisation' to make it easy to understand. Employees are used to working with 'Explorer' so they found the proposal for document organisation using this structure very useful and easy to understand.

Maintenance: There is no maintenance by the users but the system administrator is able to modify and update information, migrate all the contents, input data, etc. These functions are easily performed using MySQL.

Some feedback was obtained from *Question 8. What do you think the system is missing?* from the companies.

Initially evaluation took place at Spanish companies, although in English. All the companies complained about the language. The system was quickly translated and they were provided with access to the system with the aim of evaluating the system again. After this, all the employees said that the language facilitated their understanding of the system.

From *Question 9. What do you miss?*, some feedback from the companies was obtained.

At the beginning the validation was done in Spanish companies but in English language. All of them complained about the language. The system was quickly translated and the access to the system was provided to them with the aim of validating the system once again. After that, all the employees ensured that the language made the understanding of the system lighter.

One of them suggested that it could be interesting to have a direct link or automatic update of documents from the companies' server to the WPMS of the project and vice versa. This means that when a document is updated in the company's server, it might be useful to synchronize it and automatically update it in the WPMS.

Another suggestion was to link the documents to the deadlines related to the programming. The survey is summarized in a table and can be consulted in Appendix 3.

10.6. Discussion

The evaluation of the system is the most important part of the thesis. To strengthen the information presented in Chapter 5 (Web Based Project Management Systems) and, concretely, the data from the utilization of WPMS in the Construction Industry, a survey was carried out with the aim to obtain a slightly vision of the real situation in Spain.

Effectively, the results and findings confirmed the initial considerations:

AEC industry is in the early days of adopting WPMS and only big companies are using them. The majority of them deliver paper-based formal documentation but the working documents are exchanged basically via e-mail. Ten years ago it would be unthinkable to be using e-mail by this so traditional sector. This means that even modifications are very slow; the construction sector adopts the improvements of the market.

Designers (Architects and Engineers) are the ones who believe on the benefits of WPMS. They are most likely to use these tools because they have the necessary infrastructure to support them.

On the other hand, contractors are the least likely users of these tools and those who don't see the necessity of adopting them. Their reasons for non adopting these tools are basically the need for training and the difficulties of understanding such changeable tools.

Basically, those SMEs who have ever used any WPMS were driven by a party higher in value chain.

In conclusion, in a near future, big companies from the AEC industry will solidify on the use of these tools and will drive SMEs that will have to move to this new way of working.

With this new scene, the 'Life cycle Document Management System' will play a very important role in the exchange of information between partners of a project and the central repository of the whole project (WPMS).

The verification of the system was carried out to assure the correctness of the system in terms of requirements and technical errors. Different academics from the construction area were asked to check and verify the contents and principles of the database. Following their different opinions, the system was improved and reorganized.

Validation was also conducted to get feedback from the potential users of the system. Successfully, all the companies with sufficient infrastructure to support WPMS considered that the system was very useful, not only for the interaction with other companies when working with WPMS but also for their internal document management. The suggestions and ideas they proposed were taken into account and introduced into the system.

Concluding, the verification and validation of the system was carried out during all the stages of the development of the system, reflecting the results of the verifications in the final system.