

**LA GESTIÓ DE LA INFODÈMIA I LA
DESINFORMACIÓ EN SITUACIONS
D'EMERGÈNCIA. UNA PERSPECTIVA DES
DE CATALUNYA**

Sergi Xaudiera Borrell
Tesi Doctoral

Directora: Ana Sofia Cardenal, PhD.

Presentada pel
Doctorat en Societat de la Informació i el Coneixement

Universitat Oberta de Catalunya

Abril, 2024

PARAULES CLAU

Ciberseguretat

Ciències socials computacionals

Comunicació de crisi

Covid-19

Desinformació

Gestió d'emergències

Infodèmia

Mitjans socials

Xarxes socials

RESUM

L'auge de la desinformació que ha acompanyat les situacions d'emergència en la darrera dècada ha generat un ecosistema informatiu d'una complexitat increïble. No existeix una solució simple per combatre aquesta amenaça, i les estratègies convencionals no han resultat eficients posant en risc la salut i la seguretat de les persones. Per aconseguir mitigar-ne els efectes negatius, cal un enfocament multidisciplinari, aprofitar les dades disponibles de les plataformes socials i un conjunt de mètodes adaptat a les noves tàctiques de disseminació.

Aquesta tesi assumeix part d'aquest repte a partir de l'estudi de la propagació de relats alternatius durant la pandèmia de la COVID-19. Mitjançant mètodes computacionals estudia el contagi entre diferents regions europees d'es d'una perspectiva regional, el paper de les fonts creïbles i el tipus de contingut visual més compartit. A més, es completa la visió amb l'experiència de gestió de la infodèmia d'un govern regional. Aquest enfocament integral és essencial per abordar aquesta problemàtica, ja que permet entendre els mecanismes que influeixen en la disseminació d'informació i dissenyar estratègies governamentals més eficaces per combatre-la.

ABSTRACT

The rise of misinformation that has accompanied emergency situations in the last decade has generated an incredibly complex information ecosystem. There is no simple response to this threat, and conventional strategies have proven ineffective, putting people's health and safety at risk. To mitigate its negative effects, a multidisciplinary approach is needed, harnessing the data available from social platforms and a set of methods adapted to new dissemination tactics.

This thesis undertakes part of this challenge by studying the spread of alternative narratives during the COVID-19 pandemic. Using computational methods, it studies the contagion between different European regions from a regional perspective, the role of credible sources and the type of visual content most shared. In addition, it completes the picture with the infodemic management experience of a regional government. This holistic approach is essential for tackling this problem, as it allows us to understand the mechanisms that influence the spread of information and to design more effective government strategies to tackle it.

RESUMEN

El auge de la desinformación que ha acompañado las situaciones de emergencia en la última década ha generado un ecosistema informativo de una complejidad increíble. No existe una solución simple para combatir esta amenaza, y las estrategias convencionales no han resultado eficientes, poniendo en riesgo la salud y la seguridad de las personas. Para mitigar sus efectos negativos, se necesita un enfoque multidisciplinario, aprovechar los datos disponibles de las plataformas sociales y un conjunto de métodos adaptados a las nuevas tácticas de diseminación.

Esta tesis asume parte de este desafío a partir del estudio de la propagación de relatos alternativos durante la pandemia de la COVID-19. Mediante métodos computacionales, estudia el contagio entre diferentes regiones europeas desde una perspectiva regional, el papel de las fuentes creíbles y el tipo de contenido visual más compartido. Además, se completa la visión con la experiencia de gestión de la infodemia de un gobierno regional. Este enfoque integral es esencial para abordar este problema, ya que permite entender los mecanismos que influyen en la diseminación de información y diseñar estrategias gubernamentales más eficaces para combatirla.

TAULA DE CONTINGUTS

| | |
|---|-----------|
| Paraules clau | iii |
| Resum | iv |
| Abstract..... | v |
| Resumen..... | vi |
| Taula de Continguts | vii |
| Llista d'imatges..... | ix |
| Llista de taules | x |
| Llista d'abreviacions..... | xi |
| Declaració d'autoria original | xii |
| Agraïments..... | xiii |
| Prefaci | xiv |
| 1. Introducció general | 18 |
| 2. Estat de l'art..... | 22 |
| 2.1 El paper dels canals institucionals en la propagació..... | 24 |
| 2.2 L'àmbit regional i de les plataformes | 26 |
| 3. La tesi..... | 29 |
| 3.1 Disseny de la recerca | 30 |
| 4. Ibuprofen Narratives in Five European Countries During the covid-19 pandemic | 33 |
| 4.1 Research questions..... | 33 |
| 4.2 Essay summary | 34 |
| 4.3 Argument & Implications..... | 34 |
| 4.4 Findings | 38 |
| 4.5 Methods | 45 |
| 5. Analysing Spanish Visuals on covid-19 Misinformation: Contrasting Spanish and English Visual Narratives | 49 |
| 5.1 Introduction | 49 |
| 5.2 Literature review..... | 50 |
| 5.3 Methods | 54 |
| 5.4 Results | 55 |
| 5.5 Discussion and conclusions | 60 |
| 6. La gestió de la infodèmia a Catalunya durant el primer any de la pandèmia de la covid-19 | 63 |
| 6.1 Introducció..... | 63 |
| 6.2 Marc teòric..... | 64 |

| | | |
|-----------|---------------------------|-----------|
| 6.3 | Mètodes..... | 67 |
| 6.4 | Descobriments..... | 69 |
| 7. | Conclusions | 77 |
| 8. | Treball futur..... | 79 |
| | Bibliografia | 83 |
| | Annexos | 91 |

LLISTA D'IMATGES

| | |
|--|----|
| Figure 1. French health minister, Olivier Veran, tweet against ibuprofen | 35 |
| Figure 2. Tweets per language: Absolute (Panel A), Percentual (Panel B) | 39 |
| Figure 3. Top tweets by language, truthfulness and author type. | 41 |
| Figure 4. Frequency by: Frame (Panel A), Platform (Panel B). | 59 |
| Figure 5. Email simulating an official communication from the Colombian Government..... | 91 |
| Figure 6. False statement attributed to the Minister of Health of Ecuador. | 91 |
| Figure 7. Photo in which the author tries to convince us that these are people falling from the coronavirus in China..... | 92 |
| Figure 8. Image allegedly showing a doctor in a Madrid hospital who died of coronavirus..... | 92 |
| Figure 9. Infographic supposedly showing how to cure the coronavirus. | 93 |
| Figure 10. Alleged adverse reactions to Pfizer's vaccine in a female volunteer. | 93 |
| Figure 11. Bottle of wine sent by drone to quarantined cruise passengers. | 94 |
| Figure 12. Image showing a lion in Russia to keep people from leaving their homes. | 94 |
| Figure 13. Alleged prophecy in a book by Dean Koontz..... | 95 |
| Figure 14. Alleged prophecy on a DVD movie. | 95 |
| Figure 15. Image showing Asian people supposedly eating a baby. The author criticises them for allegedly being linked to the coronavirus. | 96 |
| Figure 16. Image showing a group of young people of Arab origin who have allegedly taken advantage of the pandemic confusion to steal from shops. | 96 |
| Figure 17. Alleged nurses criticising government party voters who applaud them at 8pm on the balcony (a practice of thanking the medical staff)..... | 97 |
| Figure 18. False information about an alleged economic measure of the Spanish Government. | 97 |

LLISTA DE TAULES

| | |
|---|----|
| Table 1. Number of tweets containing misinformation by author type | 43 |
| Table 2. Number of tweets containing misinformation by author type detached by language | 43 |
| Table 3. Number of Catalan users by infected and following official accounts. Below we include the percentage for rows | 44 |
| Table 4. The Most Common Frames..... | 56 |
| Table 5. The Most Common Frames Spanish and English Compared | 58 |
| Table 6. Persones entrevistades i càrrec en el moment de l'entrevista | 68 |

LLISTA D'ABREVIACIONS

API (Interfície de Programació d'Aplicacions)

IFCN (International Fact Checking Network).

OMS (Organització Mundial de la Salut)

OTAN (Organització del Tractat de l'Atlàntic Nord)

DECLARACIÓ D'AUTORIA ORIGINAL

El treball d'aquesta tesi no s'ha presentat prèviament per complir els requisits per a cap premi en aquesta o qualsevol altra institució d'educació superior. Segons el meu coneixement i creença, la tesi no conté material prèviament publicat o escrit per una altra persona excepte on es faci referència corresponent

Signatura: Sergi Xaudiera Borrell

Data: Abril 2023

AGRAÏMENTS

Per començar, vull agrair a l'Ana Sofia Cardenal tot el suport que m'ha donat durant aquests anys. Imagino que no ha estat fàcil adaptar-se a tots els imprevistos d'aquest temps, però sempre ha estat allà, motivant-me quan més desmotivats estava i exigint-me quan tocava. També m'agradaria fer un agraïment especial a la Sílvia Majó-Vázquez per ser una referent des d'abans de començar el doctorat. A les persones que han format part del comitè i el tribunal de la tesi, Axel, Mariluz, Javier i Carlos. També als revisors dels articles que he publicat, sou una part important del procés d'aprenentatge. Un agraïment als companys de la Generalitat de Catalunya amb qui he descobert i après sobre les plataformes socials Miquel Àngel, Jordi, Eva, Pilar, Eila, Xènia, Maria, Ester, Esther i totes les companyes que gestionen canals institucionals a les xarxes socials i hem coincidit des d'aquella nevada de març de 2010. No voldria oblidar-me dels amics i companys de bàsquet, m'heu ajudat en molts moments, encara que no en sigueu conscients.

També donar les gràcies a la meua família, Carles, Marina i Miquel; Jordi; Marta, Isabel, Jan, Roser, Xevi, Txell, Roser, Joan; Marta i Joan; i cosins per tota l'estima i suport incondicional en cada pas, literalment sense vosaltres no hauria trobat els moments per escriure la tesi. Un especial agraïment a la Dolo, qui més ha patit l'elaboració d'aquesta tesi, gràcies per estimar-me i ser la millor parella de viatge. Finalment, un agraïment a la Berta i en Mateu per ensenyar-me tantes coses importants de la vida.

PREFACI

Sento com la meva experiència professional durant el període de 2009 fins al 2024 m'ha empès a l'elaboració d'aquesta tesi. En aquest temps he pogut viure en primera persona i treballant pel Govern de Catalunya (Generalitat de Catalunya) l'aparició, auge, maduresa i declivi de les xarxes socials com a canals de relació entre la ciutadania i les organitzacions. En un inici la tesi estava pensada dins el programa de doctorat industrial. Malauradament, els canvis en les persones responsables per la part de la Generalitat, afegit a la burocràcia i la manca d'ajut econòmic, incompatible quan es tracta de treballadors públics, va fer que es convertís en un doctorat convencional a temps parcial. Tot i això, la temàtica seguia enfocada en com les institucions públiques podien fomentar un context informatiu més saludable i quines accions es podien dur a terme.

A la primavera de 2010, just una setmana després de començar, va caure a Catalunya la nevada més gran dels darrers 20 anys. Vaig tenir un inici mogut en la meva recent posició laboral. Tot el país va quedar col·lapsat, persones colgades a la carretera i els sistemes de comunicació aturats. Des de la part institucional no podíem actualitzar el web i l'única via de comunicació amb la ciutadania era a través de trucades telefòniques als mitjans de comunicació. En aquest context vam identificar que podria ser una bona oportunitat obrir un canal de Twitter, almenys ens serviria per informar a tots els mitjans tradicionals a la vegada en comptes de trucar-los un a un. El nombre d'usuaris encara era reduït, la majoria periodistes i treballadors del màrqueting, però aquests tenien una gran influència mediàtica. Així doncs, la creació del canal @emergenciescat va ser un èxit i les xarxes socials un gran descobriment per la societat en general i sobretot per la comunicació de crisi. A partir d'aquest moment vam desplegar l'estratègia del Govern a les xarxes social. Era una etapa d'expansió, apareixien noves plataformes i totes les unitats volien el seu canal. De mica en mica les xarxes socials van guanyar usuaris a Catalunya i guanyaven rellevància com a canal informatiu. El 2012 amb els incendis que van afectar la regió de l'Alt Empordà ens vam adonar que el nombre d'usuaris era molt superior i que podien convertir-se en grans amplificadors dels missatges institucionals. Les xarxes socials, especialment Twitter, es van consolidar com un canal imprescindible per la comunicació

d'emergències; actualitzant l'estat i de les novetats, donant indicacions i aprofitant els retuits i *hashtags* per arribar a persones que ni tan sols seguien els canals oficials.

En el següent lustre no va deixar de créixer el nombre de canals institucionals alhora que cada vegada més persones hi accedien per estar informats. Les xarxes socials es van convertir en la font d'informació principal i en una eina de treball imprescindible. Va canviar la relació amb els mitjans de comunicació que sabien publicaríem primer a Twitter quan teníem una nova informació. Això feia la gestió comunicativa més eficient i ràpida, ja que permetia arribar als mitjans i a la ciutadania a la vegada. Aquest podem considerar que va ser el punt de maduresa de les plataformes socials, els missatges oficials arribaven als seus destinataris directament o aprofitant les connexions de l'audiència que els amplificava.

L'evolució o les fases del cicle tenen una gran relació amb grans tragèdies. L'estiu de 2017 Barcelona va patir un atac terrorista amb diverses víctimes mortals. Un cotxe va atropellar diverses persones a la Rambla i els executors van començar una fuga per la ciutat. De seguida es va informar a través de Twitter de la tragèdia i es va demanar l'ajuda ciutadana per capturar els terroristes. Durant aquesta primera fase de la crisi, de persecució dels autors, identifiquem un gran nombre de relats alternatius; per una banda, per a intentar explicar què és el que estava passant i de l'altra sobre la gestió de l'emergència. En especial són dues les històries que més rellevància van obtenir: un tiroteig dins un centre comercial proper a la Rambla i el segrest amb ostatges en una restaurant prop de la zona. Per sort es van poder desmentir les narratives en pocs minuts, però els relats ja s'havien disseminat de manera extensa, sobretot per l'amplificació ciutadana i la de mitjans de comunicació que per la pressa d'informar els primers van publicar la informació sense contrastar-la abans. Aquí vam ser per primera vegada plenament conscients que les xarxes socials també tenien un gran poder amb efectes negatius per la societat. Si es propagaven missatges amb informació no veraç, o directament desinformació durant una emergència podia tenir un impacte enorme en la seguretat i la salut de les persones.

Durant els mesos següents començo a plantejar estudiar el fenomen de la desinformació en situacions d'emergència. Primer com a línia del màster i després com a eix del doctorat. La desinformació com una amenaça especialment en situacions de desastres. A més, estàvem en una etapa en què les plataformes socials oferien accés a les seves dades per estudiar les converses que hi tenien lloc. Era el moment de

l'aparició d'àrees d'investigació com crisis informàtics o ciències socials que aprofiten l'activitat d'usuaris a les xarxes socials com a sensors del comportament social.

A la primavera de 2020 amb l'explosió de la pandèmia provocada per la covid-19 aquest camp va entrar en una nova dimensió. Fins al moment els desastres que s'estudiaven tenien un impacte i durada més curta: atemptats terroristes, accidents provocats per l'acció humana o desastres naturals com incendis o inundacions. La gran pandèmia global va provocar des de l'inici una onada de relats alternatius i informació no veraç que va obligar l'Organització Mundial de la Salut (OMS) a declarar en paral·lel la infodèmia pel mal que aquests relats causaven a la salut de les persones.

Finalment, a 2024 la desinformació és una amenaça assenyalada per les institucions públiques d'arreu. La seva capacitat per perjudicar la salut i seguretat de les persones la converteix en un problema que qualsevol política pública ha de tenir en compte des de la seva ideació, execució i sobretot comunicació. Per aquest motiu aquesta tesi pretén ajudar el Govern de Catalunya a dissenyar un pla d'acció per estar preparat per combatre la desinformació per a quan aparegui una nova situació d'emergència.

1. Introducció general

Aquesta tesi doctoral està centrada en l'estudi de la desinformació durant situacions d'emergència. La conceptualització i disseny de l'estudi es va fer a finals de 2019 amb un pla de recerca inicial pensat per desastres d'origen natural – com incendis o inundacions; com per aquells provocats per l'acció humana com atemptats terroristes o accidents de transport. En aquell moment ningú podia imaginar l'impacte que tindria la desinformació en una pandèmia global ni els efectes en les nostres vides. La pandèmia del 2020 és un punt d'inflexió en la percepció pública d'aquest fenomen. La societat és per primera vegada del tot conscient de la seva amenaça. La declaració de la infodèmia per part de l'Organització Mundial de la Salut (OMS) és el punt àlgid d'aquest ascens en la rellevància mediàtica i acadèmica. Des de llavors diverses organitzacions i acadèmics han estudiat el fenomen per entendre'l, combatre'l i mitigar-ne els efectes negatius alhora que ha despertat l'interès de l'opinió pública i mitjans de comunicació.

Tot i tractar-se d'un fenomen de recent auge mediàtic la desinformació és una pràctica que existeix des de la socialització de l'espècie humana. El que hem viscut en la darrera dècada amb la popularització de les plataformes socials és un creixement en l'abast i la velocitat de disseminació dels missatges. Amb aquestes eines disponibles per qualsevol persona amb accés a internet s'han multiplicat els canals informatius. Alhora la naturalesa de les xarxes socials ha facilitat que qualsevol contingut, sense importància de la seva autoria, pogués circular i ser consumit per audiències ingents en qüestió de segons. Aquests ingredients fan que l'ecosistema informatiu tingui un gran volum de missatges, alguns amb informació contrastada i d'altres no. La dificultat de distingir entre la informació veraç d'aquella que no ho és un problema molt destacat en especial durant les situacions d'emergència en què rebre informació veraç pot ser clau per preservar la salut i la seguretat de les persones. En aquestes situacions és d'especial rellevància el paper dels canals públics oficials. Per la seva credibilitat són els canals que la ciutadania consulta per estar informats. Aquest poder implica una gran responsabilitat sobre la informació que publiquen. En el context en què el missatge en si no ens demostra la seva autenticitat, la intel·ligència artificial generativa fa molt fàcil replicar continguts en qualsevol format, vídeo, imatge o fins i tot webs,

que són indistingibles de l'original, la veracitat del canal és l'únic indicador que ens pot confirmar la verificat del contingut. L'autoria del missatge és l'únic recurs que les persones poden recórrer per verificar l'autenticitat d'un contingut.

Així doncs, un dels principals objectius d'aquesta tesi és **estudiar com es propaga la desinformació en situacions d'emergència posant especial atenció en el paper que juguen les fonts oficials en aquest procés**. En contextos tan volàtils i d'alt risc com les emergències (inundacions, atacs terroristes, pandèmia...) és de vital importància que els ciutadans –particularment les poblacions afectades— es puguin recolzar en canals i fonts d'informació "creïbles" per orientar el seu comportament amb les màximes garanties. Aquest és el paper que juguen els canals i fonts oficials en situacions d'emergència. La seva credibilitat serveix de drecera informativa als ciutadans per identificar la informació veraç i seleccionar els comportaments adequats. Si les fonts oficials donen per bona una informació això serà un senyal fort (i.e., altament informativa) pels ciutadans de la seva versemblança, el que de bon segur impactarà en el seu comportament. No obstant això, la credibilitat de les fonts oficials també pot tenir conseqüències no desitjades. Precisament en contextos de desordre informatiu que exigeixen respostes ràpides per part dels poders públics, les fonts oficials poden caure també en la confusió, contribuint involuntàriament a transmetre informació no verificada, amb el resultat no buscat d'amplificar la difusió de la desinformació i augmentar el desordre informatiu. En aquest sentit, com veurem, la seva credibilitat pot contribuir a agreujar i no a resoldre el problema.

Un altre objectiu d'aquesta tesi ha estat estudiar la propagació de la desinformació més enllà de la comunitat de parla anglesa. La major part de la recerca sobre desinformació s'ha centrat en la comunitat de parla anglesa, amb les seves singularitats. Això ha limitat el coneixement a només una petita part de la conversa que es genera a les plataformes digitals. Aquest és un dels pocs treballs que coneixem que se centra en les comunitats de parla castellana (Almansa-Martínez, Fernández-Torres, i Rodríguez-Fernández 2022; Salaverría et al. 2020) i catalana per comprendre com es propaga la desinformació a través de les diferents plataformes, ampliant així el coneixement global sobre aquest fenomen amb aquestes comunitats del sud d'Europa.

Finalment, hem volgut conèixer la visió dels poders públics respecte al problema de la desinformació i com s'estan preparant per fer front a aquest problema.

Per analitzar aquestes qüestions hem realitzat tres investigacions amb una aproximació diferent en cada una per aconseguir una visió el més global possible de com es propaga la desinformació i del paper que hi juguen els canals oficials. La primera investigació se centra en com es va propagar una narrativa alternativa (i.e., no verificada) sobre l'efecte de l'Ibuprofè en el tractament del COVID-19 a través de les xarxes socials. L'estudi evidencia el paper d'amplificació que van tenir les fonts oficials en la difusió d'aquesta narrativa. El segon estudi investiga quins són els tipus de continguts (visuals) que es propaguen amb més freqüència i a través de quines plataformes en la comunitat de parla castellana. L'estudi posa de manifest que els marcs predominants que adopten aquests visuals a través de diverses plataformes són els de fonts autoritzades, aprofitant la influència que tenen aquestes fonts en la propagació d'informació. El tercer estudi canvia dràsticament d'enfocament i se centra a analitzar com els governs i poders públics poden respondre i preparar-se per fer front a aquestes situacions.

Per portar a terme aquests estudis s'han fet servir diferents dades, mètodes i enfocaments. En primer lloc, hem fet servir dades de les converses públiques a les plataformes socials (concretament, dades de Twitter) per estudiar les dinàmiques comunicatives que apareixen en situacions d'emergència com la pandèmia. Hem estudiat el rastre digital de les converses per captar la dimensió regional, dins d'una mateixa regió, i les interconnexions que sorgeixen entre les persones i relats de regions i països diferents per tal de poder treballar amb previsió dels relats que poden guanyar tracció en un territori. Amb aquesta primera aproximació hem pogut traçar el recorregut d'un relat alternatiu des que apareix en un país i com es contagia en les converses públiques d'altres territoris. A continuació hem estudiat els tipus de contingut que forma aquest relat. Posant especial atenció al contingut visual i des d'un enfocament multiplataforma per tenir una mostra el més realista possible de l'ecosistema informatiu. Finalment, hem fet ús d'entrevistes amb representants del Govern de Catalunya, com a exemple de poder públic, per analitzar com van actuar durant la pandèmia de la covid-19 i poder trobar com estar millor preparats per mitigar els efectes negatius de la desinformació properes crisis.

Tot i la gran quantitat de literatura sobre la desinformació hem intentat cobrir diversos buits que l'acadèmia i organitzacions internacionals han identificat com a rellevants per les futures investigacions. En primer lloc, hi ha pocs estudis que analitzin

el paper dels canals oficials, institucionals (Hagen et al. 2022; Hagen, Keller, et al. 2018; Hagen, Scharf, et al. 2018), canals electes o polítics en la disseminació de la desinformació (Gruzd, Mai, i Soares 2022; Recuero i Soares 2022; Tucker et al. 2018). Els canals oficials tenen un paper molt important a jugar (i una responsabilitat molt gran) en les situacions d'emergència: la seva credibilitat pot contribuir a reduir l'excés de soroll en contextos de desordre informatiu, però també, com hem apuntat, el poden augmentar si involuntàriament difonen informació no verificada. En segon lloc, no molts estudis han monitorat l'impacte de la desinformació en situacions d'emergència des d'un enfocament multiplataforma, que s'ajusta millor a la realitat de com circula la informació en el nou entorn informatiu híbrid i multimèdia. En tercer lloc, la major part dels estudis sobre desinformació s'han centrat en continguts textuais, sent molt pocs els que han estudiat els continguts visuals (Brennen, Simon, i Nielsen 2021). No obstant això, el contingut visual s'ha convertit en el més consumit a totes les plataformes socials (DataReportal 2022). Aquest biaix en els d'estudis s'explica perquè els continguts textuais són més fàcils de capturar i analitzar que els visuals, que requereixen un major esforç. Finalment, hem volgut cobrir el buit d'investigació acadèmica fora de la comunitat dels Estats Units i de països de parla anglesa on es concentren la majoria dels centres i investigadors que estudien la desinformació. Aquesta tesi té l'origen a la regió de Catalunya, en primer lloc, volem contribuir a ajudar la mitigació en aquesta regió europea amb una mirada global per entendre les interconnexions i compatibilitat amb la recerca realitzada per altres acadèmics.

2. Estat de l'art

La popularització de les xarxes socials a partir de la primera dècada del segle XXI augmenta el volum d'informació disponible. En aquest moment ja hi ha investigadors que identifiquen la infodèmia com un possible problema provocat per l'excés d'informació, verificada o no verificada, que dificulta l'accés a la informació veraç (Eysenbach 2002). L'auge de les xarxes socials i eines de publicació social com els blogs permeten a qualsevol persona crear un canal de comunicació propi i influenciar a grans audiències (Bakshy et al. 2011) sense la intermediació dels mitjans de comunicació, que fins llavors havien vehiculat la comunicació de masses. L'aparició d'aquestes eines permet l'aparició de models de difusió entre iguals amb capacitat d'arribar a grans audiències en poc temps (Bakshy et al. 2012). L'ecosistema informatiu i les relacions socials pateixen una transformació (González-Bailón 2017). Es passa d'un entorn on la publicació de missatges estava controlat per pocs actors, amb un alt cost econòmic i, per tant, un llarg procés de validació, a un altre en què qualsevol pot publicar sense cap filtre previ. Aquest canvi provoca un ecosistema amb una major diversitat de continguts (Bakshy, Messing, i Adamic 2015), però alhora fa que el volum i l'abast dels missatges amb informació no veraç arribin a un nombre més gran de persones.

L'increment d'usuaris que publiquen a les plataformes socials les converteixen en un entorn mediàtic cada vegada més rellevant. Les organitzacions públiques adopten aquestes eines i comencen a crear canals per gestionar la seva presència de manera activa. Les plataformes socials es converteixen en eines especialment útils per la comunicació de crisi i per gestionar les emergències (Palen et al. 2020). Tot i això, a mitjans de la primera dècada del segle XXI també es comencen a veure, o a prendre consciència, dels usos menys favorables que se'ls hi pot donar. A partir de l'escàndol de Cambridge Analítica durant el referèndum del Brexit (Rosenberg, Confessore, i Cadwalladr 2018) i la posterior interferència política a les eleccions dels Estats Units d'Amèrica de 2016 es pren una consciència global de la dimensió i capacitat que poden tenir les xarxes socials per propagar informació no veraç per tal d'influenciar l'opinió pública d'una regió. No tan sols per afavorir un candidat o altre sinó també per crear divisió dins de comunitats locals i fomentar el conflicte (DiResta et al. 2018). Veient

la dimensió de l'amenaça la Unió Europea agafa la iniciativa i realitza el primer estudi per intentar definir el problema. Engloba la desinformació dins el *desordre informatiu* (Wardle i Derakhsan 2017) i fa unes primeres propostes per estudiar-lo. És especialment rellevant la classificació que es fa del contingut no veraç publicat i identificat segons la intenció de l'autor. Agafant el terme en anglès per la manca de traduccions fidedignes en català parlem de: (1) *misinformation* quan es desconeix si hi ha voluntat d'influenciar; (2) *disinformation* quan existeix voluntat d'influenciar a través d'informació no veraç; i (3) *malinformation* quan es publica informació veraç amb voluntat d'influenciar. En aquesta tesi fem servir el terme desinformació a escala genèrica. A partir d'aquest moment s'accepta les definicions de les formes de l'amenaça i es disposa d'un primer marc d'investigació comú (Altay et al. 2023).

En els darrers anys la desinformació ha aparegut a l'agenda pública. Organitzacions públiques d'abast internacional de seguretat, com l'OTAN (North Atlantic Treaty Organization 2022), o d'abast local, com governs (Gobierno de España 2022) han assenyalat l'amenaça i han creat grups de treball per entendre i combatre-la. No només en l'àmbit institucional s'ha pres consciència de l'amenaça, sinó que també ha transcendit a altres capes de la societat i l'opinió pública assenjala la desinformació com una de les seves principals preocupacions (Unión de Televisiones Comerciales en Abierto 2022). A l'estat espanyol ocupa el 4t lloc per darrera de la vulnerabilitat energètica, el ciberespai i la inestabilitat econòmica. A més, és la segona amenaça amb major previsió de creixement els pròxims cinc anys després dels ciberatacs. Tot i aquest despertar social, quan es pregunta sobre qui són els grups de població més afectats per la desinformació tots els grups assenyalen altres franges com a públics més sensibles. Aquesta és una de les característiques principals de la desinformació; totes les persones es creuen immunes. El pensament general és que afecta a les altres persones menys a nosaltres. La realitat és que qualsevol persona i grup d'edat és un objectiu de les campanyes d'influència dissenyades amb molts recursos per impactar els públics seleccionats (Starbird 2017).

Les campanyes d'influència informativa són una gran amenaça per al benestar de les societats. Aprofiten les eines i canals disponibles per crear i propagar missatges per buscar un profit propi. Aquesta capacitat d'adaptació i integració de les novetats tecnològiques fa que calgui una constant actualització per a identificar i combatre com es duen a terme aquestes eines. A més, en situacions d'emergència es converteixen en

una amenaça encara més gran. Tant durant crisis generades per la naturalesa, incendis, inundacions o pandèmies, com les generades per l'acció humana, conflictes armats, atemptats terroristes o accidents de transport, es multiplica la capacitat de la informació d'afectar a la seguretat i salut de les persones (Castillo 2016; Eriksson 2018; World Health Organization 2022). És el moment quan les societats són més vulnerables i és quan cal una especial protecció en tots els entorns de relació social.

2.1 El paper dels canals institucionals en la propagació

Entendre com i per què es propaga la informació en les plataformes socials és el primer pas per combatre la desinformació. Les notícies falses arriben a una audiència més gran, a més velocitat i durant un període de temps més llarg que les notícies verídiques (Vosoughi, Roy, i Aral 2018). Hi ha diferents aspectes que fomenten la propagació dels missatges no veraçs. Científics de múltiples disciplines han investigat la propagació des de perspectives complementàries. S'han estudiat els factors psicològics, especialment els biaixos humans que ens fan més vulnerables. Entre ells el biaix de confirmació, quan acceptem relats que confirmen les nostres creences, la dissonància cognitiva, quan no volem acceptar una idea que va en contra de les nostres creences, o l'efecte de tercera persona, quan pensem que la desinformació no ens afecta a nosaltres i sí, en canvi, a les altres persones (Shane 2020). També s'ha estudiat l'impacte de l'alfabetització mediàtica i la manca de competències digitals. S'ha demostrat que les persones amb més competències digitals, que entenen com funcionen les plataformes socials, són menys vulnerables a la desinformació (Guess et al. 2020). D'altra banda, les persones amb menys competències digitals són dels grups de població que més comparteixen, i fomenten, la propagació d'informació no veraç (Guess, Nagler, i Tucker 2016). En l'àmbit psicològic també s'ha documentat un impacte negatiu quan l'usuari visualitza respectivament un contingut no veraç, haver estat prèviament exposat a la mateixa notícia falsa incrementa les possibilitats que et creguis aquesta notícia quan et torni a aparèixer (Pennycook, Cannon, i Rand 2018). Finalment, dues predisposicions psicològiques que tenen un origen en la poca confiança en les institucions, el rebuig de la informació provinent d'experts i figures amb autoritat i la inclinació a interpretar els esdeveniments com productes conspiranoics, a trobar en tot el que passa motivacions ideològiques partidistes i a veure com a finalitat de tots els esdeveniments la cerca d'un benefici pel partit que està en el govern (Uscinski et al. 2020). Tot i la

falsa creença popular, el paper dels bots en la propagació és residual, ja que els missatges que comparteixen els bots acostumen a tenir un índex de lectura molt inferior en comparació amb els missatges publicats per canals publicats per humans (Broniatowski et al. 2018). A més, les plataformes han dut a terme mesures per tal que cada vegada sigui més complicat dur a terme campanyes d'influència coordinades amb bots sense ser identificades (Roth 2020).

Els canals oficials i dels representants institucionals juguen un paper clau en la propagació dels missatges. A causa de la seva posició disposen de gran influència dins la comunitat i una visibilitat que acostuma a ser amplificada pels mitjans de comunicació tradicionals. D'altra banda, obtenir informació de les fonts oficials és una de les mesures bàsiques que es recomana per obtenir informació veraç i combatre la desinformació. Sobretot durant les primeres hores d'una emergència (Innes 2018), quan de sobte passen fets inesperats, canvia la realitat i la incertesa no deixa de créixer (Buntain et al. 2016), és quan augmenta la necessitat de buscar en els canals de confiança una explicació que ajudi a entendre els fets que estan passant (Palen et al. 2020). Però què passa quan són els canals de confiança els que disseminen informació no contrastada? Això pot tenir unes conseqüències catastròfiques, especialment en el context d'una emergència. Quan són les fonts creïbles qui comparteixen la desinformació hi ha dues vies que s'han demostrat eficients per aturar els efectes expansius. En primer lloc, que la mateixa font faci una publicació rectificanc la informació (Wardle i Derakhsan 2017). Tot i això, és possible que el relat amb la informació falsa segueixi el seu recorregut aprofitant la naturalesa de les xarxes socials. En segon lloc, en funció de les habilitats en competències digitals de la població afectada la informació no contrastada podrà arribar més o menys lluny. Per exemple, quan hi ha una massa crítica que valida la informació en diferents fonts (Guess et al. 2020) i que té una visió crítica amb la dieta informativa l'ona expansiva arribarà més lluny. Tot i que hi ha força estudis que analitzen el paper de les fonts creïbles en la propagació de la desinformació, són molt menys els estudis que se centren en el paper dels canals oficials, i molt menys encara els que estudien el paper d'aquests canals en situacions d'emergència. Aquesta tesi té com un dels seus principals objectius analitzar el paper dels canals institucionals en la propagació de la desinformació en situacions d'emergència.

2.2 L'àmbit regional i de les plataformes

Com hem vist la desinformació és una amenaça global que impacta a totes les regions per igual. Tot i el gran component regional de la comunicació social, la investigació relacionada amb la desinformació ha estat centrada en la comunitat de parla anglesa, especialment centrada en els Estats Units d'Amèrica. Diversos acadèmics (Nguyễn et al. 2022; Reuter, Hughes, i Kaufhold 2018; Tucker et al. 2018) han indicat com cal aprofundir en l'estudi de la desinformació en l'àmbit regional, especialment en aquelles comunitats de parla no anglesa (Calo et al. 2021). Cal tenir en compte que en diversos països es parlen diferents llengües i que si estudiem només la conversa en una llengua es deixa de tenir en compte una part important de la població. A més com s'ha vist (Xaudiera i Cardenal 2020) els relats alternatius es contagien d'una comunitat de parlants a una altra. Aquesta tesi enfoca el fenomen en la regió de Catalunya i les llengües catalanes i espanyola. Cal tenir en compte que la llengua espanyola no només es parla a Espanya, sinó que es parla en regions de l'Amèrica Llatina a més d'altres països en què no és llengua oficial. Per exemple en els Estats Units d'Amèrica és la segona llengua més parlada. Les plataformes socials, majoritàriament amb seus als Estats Units no tenen en compte la diversitat cultural a l'hora de distribuir els recursos per combatre la desinformació i centren els esforços en la comunitat de parla anglesa (Avaaz 2020). Això significa que les persones que parlen altres idiomes diferents de l'anglès són més vulnerables per la manca de recursos destinats per les plataformes i es converteixen en objectius dels promotors de relats alternatius (Knuutila, Neudert, i Howard 2022).

Una altra limitació dels estudis existents sobre desinformació és que són uniplataforma. La major part s'ha centrat en Twitter per la facilitat amb la qual es pot accedir a les seves dades (Theocharis et al. 2021). Per fomentar aquesta visió més àmplia diversos acadèmics (Watts, Rothschild, i Mobius 2021) i institucions públiques (European Commission 2023) han demanat a les plataformes una majora transparència i compartició de les dades d'activitat dels usuaris pe fins acadèmics. Molta de la investigació acadèmica que estudia la propagació de desinformació només s'estudia el pas en una sola plataforma encara que s'ha demostrat que de la mateixa manera que els relats es contagien entre comunitats de diferents llengües també es transmeten entre diferents plataformes socials. Per entendre i resseguir la propagació cal un enfocament multiplataforma (Starbird i Wilson 2020) que ens permeti seguir el camí que segueixen

els relats. És necessari doncs ampliar la investigació realitzada tenint en compte com es contagia la informació amb un enfocament més ajustat a la realitat en què es tingui en compte la diversitat lingüística i tecnològica dels relats (Tucker et al. 2018). Aquesta tesi vol cobrir alguna d'aquestes mancances abordant el tema de la propagació de la desinformació en comunitats lingüístiques de parla no anglesa i fent servir un enfocament multiplataforma.

3. La tesi

Aquesta tesi es compon de tres estudis. La primera investigació *Ibuprofen narratives in five European countries during the COVID-19 pandemic* se centra en la difusió dels relats alternatius i el seu component multicomunitat. S'estudia com una història falsa circula a través de 5 estats europeus diferents i com són les seves dinàmiques. També s'analitzen les particularitats dels usuaris que promouen aquesta disseminació i el paper de les fonts creïbles, en aquest cas particular amb especial atenció al paper del ministre de Sanitat de França en la propagació de la desinformació en l'àmbit paneuropeu. Finalment, en aquest primer estudi estudiem l'efecte en els usuaris que té seguir canals governamentals i com impacta a la disseminació de desinformació. L'estudi respon les següents preguntes de recerca:

- Com va circular la narrativa sobre l'ibuprofè i la COVID-19 per aquests cinc països europeus? La desinformació es va estendre seguint una dinàmica global o local? On era més freqüent la desinformació en aquests territoris?
- Quin tipus d'usuaris van ajudar a promoure aquesta informació? Quin va ser el paper de les fonts creïbles en la propagació de la desinformació?
- Quin paper va tenir el seguiment dels comptes oficials del govern en la protecció dels ciutadans contra missatges no verificats? En particular, va ser eficaç seguir els comptes oficials del govern per contrarestar aquesta informació?

La següent investigació *Analysing Spanish Visuals on COVID-19 Misinformation: Contrasting Spanish and English Visual Narratives* se centra en l'estudi del contingut de la desinformació. En aquest estudi analitzem el contingut visual compartit a la comunitat de parla espanyola durant la primera fase de la pandèmia de la covid-19. Primer estudiem el contingut classificant-lo a partir dels marcs més compartits. Després comparem els marcs més compartits de la comunitat castellana amb els de la comunitat anglesa durant el mateix període de temps. Finalment, busquem la relació entre els marcs i les plataformes socials on van ser identificats per primera vegada. L'estudi dona resposta a les següents preguntes de recerca:

- Quins són els marcs comuns més compartits en el contingut visual a la comunitat de parla castellana relacionats amb la covid-19?
- Són aquests marcs comuns més concorreguts a la comunitat espanyola similars als de l'anglesa en el mateix període temporal?
- Quina és la importància de les plataformes socials en la propagació de cada marc? Hi ha relació entre els marcs i les plataformes on es comparteix?
- Incloure aquí la part del disseny, dades i mètode

Finalment, a *La gestió de la infodèmia a Catalunya durant el primer any de la pandèmia de la covid-19* hem fet treball de camp en el territori en el qual ens centrem. L'estudi es basa en entrevistes a diferents persones responsables de la gestió de la pandèmia en el Govern de Catalunya per tal de conèixer i analitzar com la institució es pot preparar per combatre la desinformació en futures situacions d'emergència. Les preguntes de recerca que guien aquest estudi són les següents:

- Com les institucions públiques, especialment la Generalitat de Catalunya, van combatre la infodèmia durant els primers 12 mesos de la pandèmia?
- Quins són els aspectes del flux de treball de les organitzacions de *fact-checking* que poden ajudar al sector públic en la gestió de la infodèmia?
- Com ens podem preparar per combatre millor la següent infodèmia?

3.1 Disseny de la recerca

Per respondre aquestes preguntes, aquesta tesi fa servir un ampli ventall de dades i mètodes diferents. El primer estudi fa servir dades de Twitter. A través de la seva api vam generar un conjunt de dades amb els relats alternatius que relacionaven el consum del medicament Ibuprofè amb l'agreujament dels símptomes provocat per la covid-19. A partir d'aquest conjunt de dades vam aplicar mètodes computacionals - anàlisi de xarxes socials, models de difusió i visualització de dades, per explicar i analitzar com els relats van circular a través de 5 països europeus (França, Alemanya, Països Baixos, Itàlia i Espanya) a l'inici de la pandèmia. Tot seguit, una vegada entesa la dinàmica comunicativa, vam crear les visualitzacions de dades que acompanyen l'article. En aquest article en concret vam tenir en compte la necessitat d'estudiar el fenomen de la

desinformació des de la perspectiva multiregional i de com els relats es propaguen per les diferents comunitats lingüístiques.

Pel que fa al segon estudi, vam aprofitar un conjunt de dades creat per la *International Fact Checking Network* (IFCN) en què les diferents organitzacions de *fact-checking* llistaven i classificaven el contingut visual de notícies falses a mesura que l'identificaven a la seva regió. Posteriorment de manera manual vam classificar els marcs més comuns dels continguts filtrats. Amb aquests marcs ordenats vam comparar els nostres resultats amb els d'estudis previs que havien estudiat la comunitat de parla anglesa. Finalment, vam fer una regressió per buscar relació entre els marcs i les plataformes on van ser identificats per primera vegada. En aquest article vam cobrir el buit acadèmic sobre el tipus de contingut visual, les converses en llengua no anglesa, castellà en el nostre cas, i una perspectiva multiplataforma tenint en compte el conjunt de les plataformes socials.

Finalment, el tercer estudi es basa en 12 entrevistes personals amb responsables de diferents àmbits del Govern de Catalunya i organitzacions de verificació d'informació establertes a l'estat espanyol i Catalunya per conèixer la seva actuació durant la gestió de la pandèmia de la covid-19. A partir d'aquestes entrevistes identifiquem els aspectes més rellevants de la gestió ordenats en amenaces, fortaleeses, debilitats i oportunitats. Finalment fem un seguit de recomanacions. En aquest article abordem la necessitat d'estudiar la desinformació des d'una perspectiva qualitativa d'àmbit local i regional amb un enfocament global més enllà de les particularitats del cas concret. La intenció és fer un seguit de propostes enfocades al Govern de Catalunya que alhora siguin extrapolables a altres organitzacions públiques d'altres territoris.

4. Ibuprofen Narratives in Five European Countries During the covid-19 pandemic¹

We follow the trajectory of the unverified story about the adverse effects of using Ibuprofen for treating the Coronavirus disease 2019 (COVID-19) on Twitter, across five European countries. Our findings suggest that the impact of misinformation² is massive when credible sources (e.g., elected officials, mainstream media) participate in its propagation; yet, they also imply that crisis communication management has a local scope given the greater reach and impact of regional channels in the spread and countering of misinformation. These patterns reveal both the global and local dynamics involved in the spread of misinformation. However, they are based on Twitter data, which might cast doubt on their generalizability. We discuss these and other limitations of the study as well as some of their implications for future research in the closing section of this article.

4.1 Research questions

- How did the narrative about Ibuprofen and COVID-19 circulate across these five European countries? Did misinformation spread following a global or a local dynamic? Where was misinformation more prevalent within these territories?
- What kind of users helped promote this information? What role did credible sources play in propagating misinformation?
- What role following official governmental accounts played in protecting citizens against unverified messages? In particular, was following official governmental accounts effective for countering this information?

¹ This paper is co-authored with Ana S. Cardenal and published in Harvard Kennedy School (HKS) *Misinformation Review*. <https://doi.org/10.37016/mr-2020-029>. July 2020.

² We use the term misinformation to refer to false or inaccurate information that is shared accidentally. In contrast, we use the term disinformation when false or unverified information is shared deliberately.

4.2 Essay summary

- We use Twitter data to understand how unverified information related to Ibuprofen was promoted by credible sources and travelled across five European territories for two weeks—from March 11 to March 25, 2020.
- We find that the impact of misinformation in a particular territory is remarkably superior when regional credible sources are key players in promoting unproven narratives.
- Crisis communication management has a local scope. Despite being a global pandemic, regional channels had a greater reach and impact among citizens in their territory.
- The degree of contagion between credible sources is very high, which causes an exponential increase in the reach of their messages. In the case studied, media played an important role as propagators, both within and across territories.
- Following official sources —when they publish verified information— reduces the likelihood of propagating misinformation and is a good means to be better informed.
- During the pandemic, social media companies tailored their policies to heal the information in their environments, but they face new challenges, as credible sources (from verified accounts) are responsible for disseminating misinformation.

4.3 Argument & Implications

In the first months of 2020, the COVID-19 pandemic caused one of the most dangerous global health crisis of our era. As a novel disease, there was great uncertainty about how to fight it. The World Health Organization (WHO) tried to coordinate a global response; the global scientific community conducted multiple research efforts on how to treat and cure the disease. Although information about the preliminary results of the studies soon began to emerge, the outcomes were inconclusive, and different information on disease treatment was reported in certain regions. Knowledge gaps together with an increasing demand for information fostered the appearance of numerous narratives not always based on proven facts. The vast

amount of unproven information circulating during this period is well-captured by the WHO's declaration³ of the infodemic days before declaring the health pandemic.

In this article we focus on one instance of an alternative narrative appearing in Germany and spreading across different European countries from March 11-24. This narrative advocated avoiding the use of Ibuprofen to treat COVID-19 and even warned that its consumption increased mortality in patients, when in fact no existing scientific evidence supported any of this. What made this story distinct is that the French health minister, Olivier Veran, was one of its main promoters. On March 14th, he posted a tweet⁴ (Figure 1) advising patients against the use of Ibuprofen because preliminary research pointed to associated risks. This message triggered an immediate reaction that greatly extended its reach.



Figure 1. French health minister, Olivier Veran, tweet against ibuprofen ⁵.

Comparable narratives in other territories (e.g., US, Brazil) include the use of hydroxychloriquine as a possible cure for the disease, which were also promoted by top elected officials (e.g., Trump, Bolsonaro). Studying misinformation about treatments for any disease is important, especially when reliable sources play a key

3 <https://www.who.int/dg/speeches/detail/munich-security-conference>

4 <https://twitter.com/olivierveran/status/1238776545398923264>

5 Tweet translation: taking anti-inflammatory drugs (ibuprofen, cortisone, ...) could be a factor in aggravating the infection. In case of fever, take paracetamol. If you are already taking anti-inflammatory drugs or in case of doubt, ask your doctor for advice

role in promoting false information, because of the impact it may have on people's safety and health (Starbird, Arif, and Wilson 2019).

In this research we used Twitter data to trace the public path of the Ibuprofen story. We discovered that this story originated in a WhatsApp voice message. Due to the private nature of instant messaging applications, we were only able to trace the narrative from the time it jumped from WhatsApp to Twitter.

Based on digital trace, we were able firstly to study where the story started and how it spread from the country of origin (Germany) to users in the Netherlands, France, and finally Spain and Italy. Next, we analyzed the role that credible sources, such as political representatives and respected media, played as promoters of misinformation. Finally, we used the case of Catalonia (a Spanish region for which we had data) to examine the effect that following (local) official sources had on the likelihood of debunking false information.

Overall, this study has three real-world implications. The first implication relates to the role of trustworthy sources in the spread of misinformation. Using official channels and reliable sources to obtain information is recommended as one of the basic measures, globally, to combat misinformation. But what happens when these sources are the ones that broadcast unverified information? We show that this might have catastrophic consequences, particularly in the context of a health crisis, at least for two reasons. First, because credible sources can have a massive impact on spreading misinformation, as they have the greatest influence in propagating misinformation owing to their extreme visibility and the implicit trust given to the information they convey; second, because turning to credible and expert sources is deemed as one of the few efficient strategies citizens can rely on to combat misinformation (Vraga et al. 2017).

Previous work has shown how difficult it is to correct misperceptions (Lewandowsky et al. 2012, Nyhan & Reifler 2010) and has proposed several ways in which citizens might fight misinformation (e.g., Swire-Thompson & Lazer 2020, Tully et al. 2020). Yet, these studies address debunking in general, not when official sources are the main channels disseminating misinformation. When experts and public officials are involved, there are two paths to combat misinformation. Firstly, it has proved efficient that the same reliable sources correct the information with messages prepared to be shared in social media (Wardle & Derakhshan 2017). In the case of the

false narratives related to ibuprofen, the official sources that spread the misinformation did not correct themselves. However, other trustworthy and expert sources (e.g., WHO) corrected it, albeit without much success, because despite the correction misinformation continued to spread. Secondly, we can also combat misinformation coming from trustworthy sources by better educating the public on information consumption. Improving the information environment by enhancing skills such as critical media literacy, to encourage a critical attitude towards the information we consume (Buckingham, 2019), news literacy, to understand the press business (Tully et al., 2020) and transmedia literacy skills, to figure out how multiple sources fit into the information environment (Scolari et al, 2018).

A second implication of this study is related to the scope of communication in global health crisis. Although our study shows that the degree of contagion among credible sources is very high, it also shows that communication in emergency crises has a regional dimension. This may multiply opportunities to combat misinformation if enough official sources at the local level are not disseminating unproven information. In particular, our study shows that when these sources are not promoting misinformation, following local official channels decreases (increases) the probability of spreading (debunking) misinformation by 60% (finding 4). Thus, in unaffected countries (or regions), citizens may have ways of fighting misinformation by resorting to official sources of information locally. The downside, however, is that in infected countries the local scope of communication may lock citizens in contaminated environments and leave them completely unprotected against misinformation.

A final implication of this study concerns the accountability of digital platforms in correcting misinformation within their ecosystems, as much of the unproven information was spread from verified accounts. During this pandemic, digital platforms took steps to promote healthy conversations, encouraging, among other measures, that information be obtained from credible sources. However, in this case, platforms faced a new challenge for which they had no solution: that seemingly credible sources were the ones promoting misinformation from their verified accounts. This represents a real challenge, because the kind of verification that platforms exercise to promote healthy conversations consists on authenticating the accounts (i.e., verifying that the content comes from the person who purports to come) not on checking the veracity of the content they publish. In the case studied, the platforms did

not take any action against messages from credible sources promoting misinformation until days later, when political leaders in the United States and Brazil disseminated similar unproven narratives, prompting the platforms to take stronger action, blocking these messages from their ecosystems.

In a context where multiple unverified narratives circulate on their platforms, social media companies cannot stay on the sidelines; they need to play a bigger role to guarantee a healthy information environment, especially during a pandemic since inaccurate information can potentially have very harmful consequences.

4.4 Findings

4.4.1 Finding 1: The spread of misinformation in crisis communication has a regional dimension

Our research shows that once the story gained traction, mainly after the French health minister's tweet, it continued to be transmitted between countries, even though different particular users and credible sources across countries denied the information. Figure 2 presents the number of tweets related to the narrative from March 11 to 24, showing that the narrative is prior to the tweet of the French health minister, although it becomes really relevant after this tweet is issued, on March 14. Each color represents a language⁶. In Panel A in Figure 2 we can find the total number of daily publications per language; in Panel B, the percentage of daily tweets per language. Note that Panel B maps tweets by language on an equal scale, which eliminates the huge differences in number of daily tweets between the period previous to the French minister's tweet (11-14) and the period following this tweet (14-24).

This reveals two aspects of the dissemination of information. First, despite the impact of the message of the French minister's tweet (see Panel A), the message is disseminated at different rates in each region. In Panel B, we can observe how the conversation gains traction on different days in different territories, the 14th especially in France, Germany and Spain, later the Netherlands and Catalonia are added, and we see that in Italy it was relevant from March 21st to 23rd.

Second, adding to previous research, we uncover a regional dimension in the spread of misinformation, showing that despite denials, alternative stories continued

⁶ ca for Catalan, de for German; es for Spanish; fr for French; it for Italian and nl for Dutch.

to circulate (Wang & Zhuang, 2018; Vosoughi et al., 2018). In our research we found that, although supragovernmental bodies such as the WHO⁷ ⁸ and the European Medicines Agency⁹ ruled out the information, the speed at which false information spreads was not reduced until official channels in each region denied the information. This indicates a strong regional dimension in the dissemination of information.

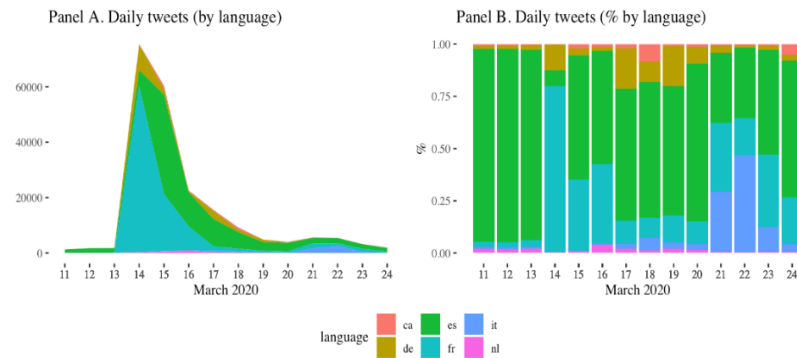


Figure 2. Tweets per language: Absolute (Panel A), Percentual (Panel B)

4.4.2 Finding 2: Misinformation was predominant in the territories where credible sources promoted it

Figure 3 shows the fate of Ibuprofen’s alternative narrative by country. Our set includes five European countries. It also includes one region (Catalonia) that will be used to test the “debunking” effect of following governmental sources. Dots are represented in different colors according to whether the message contains verified information (=TRUE, blue) or not (=FALSE, red). The shape of points represents the authorship according to types of author: credible sources (rounded); fake credible (square); generic (triangle); and profiteers (cross).

In France, where the message was promoted by a credible source, the impact of misinformation was the greatest of all studied. This can easily be seen in Figure 3, where all messages in French are coloured in red, indicating that they were all infected. In this country, the minister of health promoted the message shortly after the narrative first publicly surfaced in its territory. A few instances later, other reliable sources, such as the media, reported this information without denying it, which helped to further convey the message. Such was the effect of credible sources’ posts in this region—

7 <https://twitter.com/WHO/status/1240409217997189128>

8 <https://twitter.com/WHO/status/1240409220916432899>

9 <https://www.ema.europa.eu/en/news/ema-gives-advice-use-non-steroidal-anti-inflammatories-covid-19>

that infected information went completely unchecked with no single popular message denying the unproven information (as all the dots in red indicate).

In contrast, in the country of origin, Germany (de), the story followed a very different path. We traced the origin of the alternative narrative in this country, where it flowed through WhatsApp as a voice message. It was forwarded to different users who helped with its dissemination (the red dot shortly before March 15). As it was not possible to identify the author, the narrative didn't gain credibility and the global conversation essentially consisted of denying and making jokes about it (all blue dots around March 15). Days later, between March 17-19, when the narrative gained strength in other territories, it was debunked again (blue dots during this period). No further outbreaks occur. In short, we observe how misinformation was predominant in the regions, especially France, where reliable sources (squared dots) spread the message. Additionally, the populations' trust in official sources meant that no message refuting the misinformation had much impact.

The other studied territories lie in between these two extreme cases, by combining messages that are infected with messages that are not. The Netherlands was the country where the message first jumped, harmlessly, into the public sphere, due to geographical proximity. In Spain, Catalonia and Italy, the media and journalists were the first to report the misinformation, citing the French minister's statements. In Spain and Catalonia, we observe generic users, with less reach, denying information from the very moment the media covered it. In all the territories on the 18th a second wave appears; especially in Spain, Catalonia, France and Italy the narrative obtains a large reach. In these countries, between 18-19th the message is debunked again. In spite of this, in Italy the story circulated again strongly between the 20th and the 23rd.

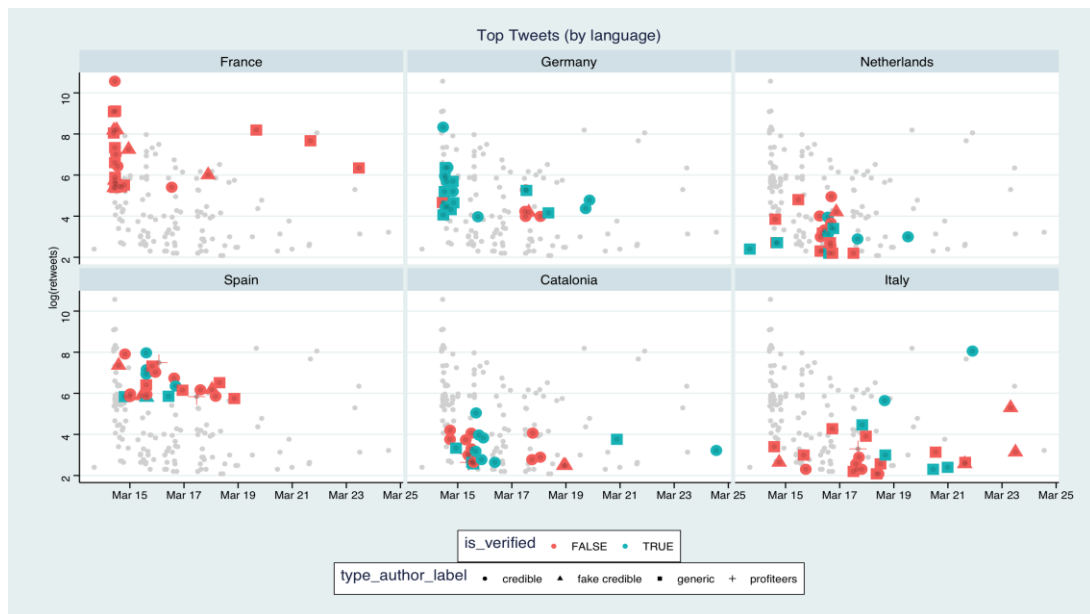


Figure 3. Top tweets by language, truthfulness and author type.

4.4.3 Finding 3: Credible sources have a critical role in information dissemination

To assess the role of credible sources involved in the trail of contagion, we code tweets according to four categories of actors: credible sources, fake credible, generic users, and profiteers. Following other studies¹⁰ we distinguished between credible sources, official and verified accounts where it is advised to obtain security and health information (e.g. official governmental accounts, elected representatives, mainstream media and reputable journalists). Fake credible, in this case, influential accounts in matters other than health and safety, such as celebrities or social media influencers; generic users (common users) and profiteers, users who want to benefit from the misinformation. As expected, each region's behaviour was different depending on whether there were reliable sources promoting the message.

Among the four categories of authors, we found that most misinformation messages were written by credible sources (see Table 1 and Table 2). We found a dangerous information connection in which trustworthy sources quickly infected each other, first with the local media reporting the misinformation of the French Minister, and then with media in other regions reporting based on information from the French media. Within this category it was remarkable how in Germany we did not find any

¹⁰ <https://www.bellingcat.com/news/2020/03/25/the-coronavirus-disinformation-system-how-it-works/>

relevant message from reliable sources promoting misinformation (Figure 3). All messages with impact from credible sources debunked the narrative. Misinformation promoted by truthful sources is especially dangerous because, as trusted sources, people have a lesser degree of reflection and tend to believe the recommendations without question.

Generic users follows credible sources as the author type with most tweets. The fact that these users did not represent the largest category shows the uniqueness of this case, in which credible sources were the main distributors of misinformation. Usually, generic users have a shorter reach and a smaller radius of influence than the other type of users. In spite of this, as they are the largest group, they play an important role in the propagation and influence of information in the closest circles of acquaintances. Following Generic users as author type are Fake credible authors (e.g., celebrities), with 17 tweets. Although this category publishes a smaller number of messages (17) it represents a great threat due to the large visibility and power of connectivity of Fake credible authors, like celebrities. Indeed, other studies (e.g., Nielsen et al.,2020) have found that celebrities have played a significant role in the propagation of unproven information during the Coronavirus pandemic.

At last, we find a small group of profiteers (5), mainly represented by politicians from the opposition, which tried to take advantage of disinformation campaign to attack the national government. The few profiteers concentrate on the more politically divided southern European countries as Table 2 shows. It is especially in Spain that the profiteer action achieves a greater scope promoted by a member of the European Parliament from a far-right party.¹¹

¹¹ <https://twitter.com/hermanntertsch/status/1239872228533186600> tweet was deleted by the author a posterior when it was already retweeted 343 times, with an estimated reach of 665,932.

Table 1. Number of tweets containing misinformation by author type

| Author | Tweets | % |
|------------------|---------------|------------|
| Credible sources | 41 | 39,6 |
| Fake credible | 17 | 16,8 |
| Generic | 38 | 38,6 |
| Profiteers | 5 | 4,95 |
| Total | 101 | 100 |

Table 2. Number of tweets containing misinformation by author type detached by language

| France | | Germany | | Netherlands | |
|------------------|---------------|------------------|---------------|------------------|---------------|
| Author | Tweets | Author | Tweets | Author | Tweets |
| Credible sources | 5 | Credible sources | 4 | Credible sources | 7 |
| Fake credible | 6 | Fake credible | 1 | Fake credible | 1 |
| Generic | 14 | Generic | 1 | Generic | 9 |
| Profiteers | 0 | Profiteers | 0 | Profiteers | 0 |
| Spain | | Catalonia | | Italy | |
| Author | Tweets | Author | Tweets | Author | Tweets |
| Credible sources | 8 | Credible sources | 11 | Credible sources | 6 |
| Fake credible | 3 | Fake credible | 0 | Fake credible | 4 |
| Generic | 5 | Generic | 2 | Generic | 9 |
| Profiteers | 2 | Profiteers | 2 | Profiteers | 1 |

4.4.4 Finding 4: People who follow local official accounts promote less misinformation and debunk it more

As a final step, we analyzed the informational impact of following official sources that shared valid and confirmed information for the region of Catalonia (Spain). For this analysis we have classified tweets according to (1) whether or not users followed any of the most important official accounts¹² participating in the conversation at that time in Catalonia, and (2) whether or not the messages disseminated were truthful (Table 3).

Table 3 presents the results. It shows that among those Catalan users who contributed to disseminating unverified information (1099), the majority (702, or 64%) did not follow official sources at the time of publication. Indeed, 2 out of 3 untruthful messages came from users not following official accounts. In contrast, among those who shared messages contributing to debunk misinformation (302), the vast majority (257, or 85%) was following official channels, compared to a minority who was not following (45, or 15%).

Table 3. Number of Catalan users by infected and following official accounts. Below we include the percentage for rows

| | Follow | No follow | Total |
|---------------------|---------------|------------------|--------------|
| Infected | 397 | 702 | 1099 |
| %% | 36 | 64 | 100 |
| Not infected | 257 | 45 | 302 |
| % | 85 | 15 | 100 |
| Total | 654 | 747 | |

After running a simple logistic regression, we estimated the probability of sharing misinformation when users did not follow official accounts in almost 70%,

12 Channels with crisis and health related information: @gencat (Catalan), @salutcat (Catalan), @emergenciescat (Catalan), @sanidadgob (Spanish). More details on methods section.

while following these official channels decreased this probability by almost 60% (results are reported in the supplementary material). Although basic, this analysis suggests that following official sources, which share credible information, is a good means of encouraging healthier conversations with verified information, which help preserve people's safety and health.

4.5 Methods

In this case study we have explored how alternative narratives about Ibuprofen travelled through different European territories in the context of the COVID-19 (or Coronavirus) global pandemic. To investigate this case, we have relied on data collected from Twitter API using R and the rtweet library (Kearney, 2019). Starting on March 11, and until March 28, 2020, we captured (N= 809,072) tweets using one of the three following words: “ibuprofen,” “ibuprofeno,” or “ibuprofene.” Thus, our data has been selected using relevant keywords (Jungerr, 2016,p. 82). Applying social network analysis, we identified the center of the conversation in five European territories: Germany, France, Spain, Italy and the Netherlands. More details related to data gathering and filtering can be found in the supplementary material.

Once the main sample was established and the trace of the story understood, the next step was to measure the social impact of reliable sources and other categories of authors in the spread of alternative stories. We used a reduction technique to focus on the main conversation (Borge-Holthoefer & González-Bailón, 2017). We filtered the top 25 messages by reach for each language. Then we manually checked that they were written in the corresponding language and we classified them into messages that contained misinformation, messages that debunked misinformation and those that were unrelated. We proceeded by excluding the unrelated messages, selecting the next message according to its impact written in the same language, until we obtained the 25 most relevant messages for each of the six territories to be studied. We were left with 152, as there were two languages with 26 tweets, since two messages had the same impact.

In the next step, we further reduced the scope of study. We focused our attention on the Spanish region of Catalonia and measured the impact of misinformation in this region. The reason for including Catalonia is twofold. First, it allowed proving further our point that the dynamics of misinformation in global health crisis have a local scope;

second, we were interested in measuring the influence that official government channels had on misinformation demystification in this region. Catalonia had the additional advantage of having a regional language of its own¹³, allowing to study the dynamics of misinformation for a sample of users located at the sub-country level. To conduct this analysis we gathered Twitter followers from official accounts managing the crisis in the region¹⁴. We applied epidemiology techniques and classified the Catalan users of the main conversation according to whether or not they followed the official channels, and their role in spreading misinformation (promoters or debunkers). From this classification, we were able to estimate the behaviour of each group of users.

The use of these techniques has helped uncover interesting patterns concerning the global and local dynamics involved in the spread of misinformation. However, to assess the significance of these patterns several caveats are in order. First, this study is based on a convenient sample of Twitter users for most of which –as is common in studies using this kind of data— we lack relevant demographic information (Golder and Macy 2014). This might raise concerns about the generalizability of our findings. Twitter data may be subject to several shortcomings stemming from biases in representativeness and sampling (Barberá and Rivero 2015), which might limit the potential for generalizing findings to the general population. In turn, problems of representativeness cannot be corrected in the absence of demographic information (Barberá and Steinert-Threlkeld 2019). Although this is an important limitation, processes of information contagion count among the many behaviors that are less sensitive to non-representative samples (Barberá and Steinert-Threlkeld 2019).

A second and perhaps more important caveat concerns causality. Although we uncover a strong correlation between the publication of unverified information by credible sources and the spread of misinformation this should not be confused with causation. In other words, we should not conclude from our findings that the spread of misinformation following the tweet of the French Health Minister was the result of “contagion” not of “selection” (for example, many influential sources in France,

13 The official languages of Catalonia are Catalan, Spanish and Aranese. Due to limitations in the data provided by Twitter we could not separate those messages from Catalan users written in Spanish. So we used the language variable to reduce our sample, even knowing that we were not studying all the Catalan users involved in the conversation.

14 <https://twitter.com/genecat> (Catalan), <https://twitter.com/emergenciescat> (Catalan), <https://twitter.com/salutcat> (Catalan), <https://twitter.com/sanidadgob> (Spanish)

including the French Health Minister, may have tweeted about the unverified Ibuprofen narrative reflecting a broad social consensus in this country over this matter). In order to separate contagion from selection and to investigate the causal effect of credible sources' influence in misinformation propagation, future studies should use more appropriate research designs such as randomized controlled experiments in virtual or physical labs (Golder and Macy 2014).

Aside from these, other limitations of this study concern the connections with the offline world and additional data restrictions. Even though social media has been a widely used media during this crisis it has not been the only media citizens have relied on to get information. Citizens have also turned to mainstream media, particularly to broadcasting media (Nielsen, Fletcher, Newman, Brennen, & Howard, 2020) as a widely shared media to get information about coronavirus and thus have had other opportunities to debunk misinformation beyond social media.

Finally, in this study we have only been able to trace the Ibuprofen story once it went public, although we know that much of the misinformation circulates through instant messaging (IM) applications. In this particular case, the story started with a WhatsApp voice message. Due to the privacy characteristics of these applications, wherein the messages are encrypted, it is not possible to quantify their role in the spread of misinformation. Despite this, these campaigns, when they gain traction and reach, jump into the public sphere, usually via Twitter, and it is from this point that we can study their behaviour.

5. Analysing Spanish Visuals on covid-19 Misinformation: Contrasting Spanish and English Visual Narratives¹⁵

5.1 Introduction

Especially since the beginning of the covid-19 pandemic, much research has been done on misinformation in emergency situations. Despite the popularity that the field gained since the WHO declaration, in early 2020, the fields of infodemiology (Eysenbach 2002), crisis informatics (Palen i Anderson 2016) and information disorder (Wardle i Derakhsan 2017) have a long history. The problem of the abundance of information, especially untruthful information, is very present in the academic world since the creation of content in digital environments has been popularised since the 2000s: websites, blogs, social media or instant messaging. Despite this large number of publications there are still aspects that need further study to understand human behaviour in these situations in order to design better strategies to combat this threat. This article aims to contribute knowledge on three of the aspects which academics have previously identified as necessary: analysis of visual content, the study of non-English communities --Spanish in this article, and a multiplatform approach for misinformation dissemination.

This article provides an approach to cross-platform research on misinformation visuals identified by international fact checkers in the Spanish-speaking community. For our research we use the data provided by the International Fact-Checking Network (IFCN), an alliance of international fact-checkers who, during the pandemic, created a hub with all the content they were verifying on the different social platforms. From this information we gather data on the visuals posted in different social networks and instant messaging platforms around the globe. We filter the visuals identified as misinformation in the Spanish-speaking community (N=163) and classify the visuals into categories according to their purpose using the categories used by Brennen, Simon, and Nielsen (2021) in the study of the English language community. We first

¹⁵ This paper is co-authored with Ana S. Cardenal and pending review for publication.

arranged visuals in frames in order to find out which were the main ones in our sample. Then we compare the results obtained in the two communities during the same period, January-March 2020, to look for coincidence among the findings. We find similarities in the most representative frames for the two languages. The main one in the two conversations is the authoritative agency, which includes the impersonation of government messages. As a third step, we investigated the relationship between the frame of the visual and the platform on which it was first identified. Although we did not find a strong relationship between frames and social platforms, we did find some indications that warn of the danger of disseminating this content in instant messaging spaces where the content that is shared can't be moderated.

In summary, this article provides a broader description of the phenomenon of information crises and misinformation by covering three aspects that academia has previously identified as gaps –the study of non-textual content, in this case visuals, shared in non-English speaking communities, comparing the behaviour of the Spanish community with that of the English one, and a cross-platform approach to get closer to real propagation, not limiting the study to a single platform. Exploring this content identified by the fact-checkers, we bring out one worrying finding: the coincidence between the three main frames of visuals misinformation in the two communities. It is especially relevant that the most popular is that of the authoritative agency, in which official messages are simulated to disseminate non-real messages. It is especially relevant for official agencies because with the emergence of generative AI it will be even easier to create visual content, images, audio and video known as deepfakes, which will be indistinguishable from the real thing. In this information context, institutional channels are the only available resource for citizens to verify the authenticity of the content. This implies the imperative for governments to be proactively present on social platforms with an active and known presence for citizens and with an active listening to the conversations that are taking place. Only in this way will citizens know where to go for official information during emergency situations.

5.2 Literature review

Especially since the emergence of the covid-19 pandemic, much research has been conducted on misinformation in emergency situations. Thousands of academics studied the phenomenon from different perspectives: actors (Hagen et al. 2022; Nielsen et al. 2020), propagation tactics (Cotter, DeCook, i Kanthawala 2022),

countering strategies (Royan et al. 2022; Vraga et al. 2023; Vraga i Bode 2017, 2021), and the measurement of its impact on each social platform (Ben Nimmo i Eric Hutchins 2023; Disinfo Lab 2022; Nimmo 2020). Most of the published misinformation studies have some common characteristics: they focus on textual content; they are centred on English-speaking communities, and they tend to study a single social platform.

In this study we wanted to fill research needs, identified by several authors in multiple articles, with which to investigate segments and paths yet to be explored and at the same time broaden the general knowledge about the field. In this study we want to investigate the misinformation crisis in three ways that we think are still under-researched. For this purpose, we have been inspired by *Beyond (Mis)Representation: Visuals in COVID-19 Misinformation* that analyse English speaking content (Brennen et al. 2021) in which they analysed and establish the main frameworks of the content of visuals identified as misinformation in the English-speaking community. Based on the structure of these authors' study, we have adapted their methodology of visual content analysis to the Spanish-speaking community, to compare it with their results from the English community and subsequently extend it to a content study taking into account the platform on which it was first identified.

The first academic gap we want to fill is the lack of research on non-textual misinformation, in our case focusing on visuals. Previous studies of misinformation has focused on textual content, some attribute this to greater ease of data collection and processing (Tucker et al. 2018). Even though currently the most consumed and shared content on social networks and instant messaging platforms is visual (DataReportal 2022), there is a scarcity of research from this perspective. Some researchers were already pointing to this problem long before the pandemic, which has only accelerated the trend towards a prevalence of visual content. The popularisation of access to generative AI, which facilitates and accelerates the production of visual content, makes the research of this format of messages even more relevant.

Most published studies on misinformation content focus on textual fields, such as title, description and tags, and indicators such as shares or likes. In this study, we analyse and classify the content according to the intentionality and the message it intends to convey. This interpretation work requires a manual and individual analysis to interpret and categorise each image within each frame. To facilitate this work point

to visual framing (Brennen et al. 2021), a technique that categorises visual content into common frames to facilitate further analysis. In our case, based on the categories identified in the English community: authoritative agency (balanced claims about actions of public authorities), virulence (over- or understates the spread of the virus, or claims that the disease is not real), medical efficacy (offers medical information, highlights tests, vaccines, equipment, suggests that there exist cures, treatments, or preventatives for the virus), intolerance (expresses racism, xenophobia, sexism, etc.), prophecy (suggests virus was previously predicted) and satire (satirical or humorous content). Examples of each frame can be found in the supplementary material section. We classified the visuals into frames to answer the first research question: **What are the most common frames in COVID-19- related Spanish visuals misinformation? (RQ1)**

The next objective of the research is to broaden the scope of academic knowledge on misinformation by including the Spanish-speaking community. Misinformation as an information disorder is a communication phenomenon of global impact but local in nature. This means that each community and region is governed by its own dynamics with particularities of the people who compose it. Research in this field have mostly been studied in the English-speaking community and there is a need to broaden the scope to understand the behaviour in other communities. This article fills a gap in the study of misinformation crisis by focusing on the Spanish-speaking conversation.

Several researchers (Nguyễn et al. 2022; Reuter et al. 2018) have previously pointed out the need to expand the study of misinformation to non-English speaking regions. Moreover, languages are often not spoken in a single territory and people from different countries may speak the same language. In the case of Spanish, it is spoken in different countries of Latin America and Spain. Furthermore, it is worth noting that Spanish, although it is not the most widely spoken language, also has a large number of speakers in other regions, for example it is the second most spoken language in the United States of America¹⁶. This common language between geographically separated regions creates a connection in which information circulates more easily.

¹⁶ Table 53. Languages Spoken At Home by Language: 2017", Language use in the United States, August 2019, U.S. Census Bureau, retrieved February 19, 2016. <https://www.census.gov/>

Therefore, if we want to study how misinformation is transmitted in a territory and we only analyse the content of a single community of speakers, we are missing a large part of the phenomenon. Having demonstrated the connection between different online language communities, we have seen how misinformation travels from one community to another at great speed and adapts to the language of each region (Xaudiera i Cardenal 2020). The study of the Spanish-speaking community provides insight into a part of the puzzle yet to be discovered. This aspect is particularly sensitive because we know that social platforms spend far fewer resources on verifying misinformation in languages other than English (Avaaz 2020). This means that the actors coordinating misinformation campaigns can more easily reach the most vulnerable people, either because of the lack of resources allocated by the platforms, or because they are less educated and therefore more likely to be targets of misinformation (Knuutila et al. 2022). Once we have the results of the most common frames in the Spanish-speaking community, we want to know if they coincide with the results obtained in the English-speaking community. Therefore, the second question of our research: **are the most common frames in Spanish similar to those in English? (RQ2)**

Another shortcoming of the research identified in the published literature is the focus on data from a single platform. Just as most of the research focuses on textual content type, in terms of platforms most of the research is done with Twitter data because of the ease with which academics can access its content through its API (Theocharis et al. 2021). This means that although it is not the most widely used platform (DataReportal 2022) it is the most studied, causing an over-study effect. This results in a large bias in our academic knowledge about misinformation, since it leaves out some of the platforms most used by the world's users. For this reason, the various platforms have long been called upon to make a greater commitment to transparency and accessibility of the data held on their servers (Watts et al. 2021) for the study of computational social science and related fields. In fact, starting in 2023 with the implementation of the Digital Services Act (DSA), it is an obligation for social platforms in the EU territory to share data and information for the study of user behaviour. This phenomenon is well known in the scientific community, which is why it evokes further study of the interconnectedness between the different platforms, which better represents how messages spread in digital environments. In addition,

other studies have shown how the behaviour and content most shared in each environment is different on each platform. With this cross-platform analysis, we have a more realistic picture of the role that different platforms play in the spread of misinformation (Starbird i Wilson 2020). This indicates that each platform has distinct audiences with unique communities. Taking advantage of the work of the International Fact-Checking Network (IFCN), which has identified the visuals on the different platforms, we have been able to study the phenomenon in all its expansion. This analysis leads us to answer our last research question: **What relevance the social platform has in the expansion of each frame? (RQ3)**

5.3 Methods

To answer our research questions, we have relied on the database provided by the CoronaVirusFacts Alliance¹⁷. This alliance led by the International Fact-Checking Network (IFCN) at the Poynter Institute unites more than 100 fact-checkers in more than 70 countries and includes articles published in at least 40 languages. This collaboration has provided since the beginning of the pandemic a database that gathers all the falsehoods that have been detected by its members.

We accessed the database on 25 June 2021. To adjust data to our interest we first filtered by the language of the visual to Spanish¹⁸. For those images that do not contain any textual element, we inferred the language from the context of the account. As a result of removing all non-Spanish and non-visual content we obtained (N=4330) visuals. Hereafter we filtered the time limit of the identification of the contents from 1 January to 30 March 2020 to adjust our sample to the one used by (Brennen et al. 2021) in his study of the English-speaking community. After manual content verification, we obtained a final sample of 163 misinformation visuals.

For our analysis, we have used only the content provided by the IFCN and the coding of content of each fact checker. Although there are common working guidelines to classify misinformation, we are aware that each organisation has its own criteria for the verification and classification of visuals. We considered these criteria as good. Lastly we manually coded the contents among the most common frames, we have

¹⁷ <https://www.poynter.org/coronavirusfactsalliance/>

¹⁸ <https://chequeado.com/latamcoronavirus/>

taken as a reference the classification the framework used by (Brennen et al. 2021). Then we coded the source where the visual was first identified. After this classification we obtained Facebook (N=83), WhatsApp (N=40) and Twitter (N=39) as source. At this point we run a regression for each frame and each social platform to see the relationship between frame and platform. Regardless of the possibility that the content could circulate in other environments before without going noticed or after the identification.

5.4 Results

We first classified the visual misinformation identified by the International Fact-Checking Network (IFCN) into the most common frames (RQ1). For this purpose, we used as basis the categorisation used by studies in English language (Brennen et al. 2021). Once all the content has been classified into frames, we look at their prevalence. Looking at table 1 we can see that the most common frame is authoritative agency (33%), which proposes content about the action of security agencies and governments. The second frame with the most identified content is virulence (33%), which identifies all misinformation, related to the virulence of the pandemic, and even questions the reality of the virus. The third most popular content is related to medical efficacy (19%), which includes content related to medical treatment, virus prevention, vaccines. The fourth most common frame is the one concerning prophecy (6%), referring to the virus was previously predicted. The fifth frame is the Satire (7%) with satirical or humorous content. The sixth most common frame is Intolerance (1%) which expresses racism, xenophobia, sexism, etc. Finally, we found a new frame that the study of English-language visuals did not encounter. This content related to Politics (1%) referring to content with political partisanship. Such visuals can also be categorised as disinformation.

We have found (Table 4) that the visuals are unevenly distributed among the most common frames. The two main ones, Authoritative Agency and virulence, have a prevalence of 33% of the total and the third one, Medical Efficacy, 19%. These three frames account for 85% of the content studied. The remaining frames also have two clear trends. On the one hand Satire with 7% and Prophecy with 6 have a significant representation. On the other hand, the intolerance and politics frames only have a 1% prevalence.

Table 4. The Most Common Frames¹⁹

| FRAME | DESCRIPTION | EXAMPLE | PREVALENCE (%) |
|-------|----------------------|--|----------------|
| 1 | Authoritative Agency | Valanced claims about actions of public authorities | 33 |
| 2 | Virulence | Over- or understates the spread of the virus, or claims that the disease is not real | 33 |
| 3 | Medical Efficacy | Offers medical information, highlights tests, vaccines, equipment, suggests that there exist cures, treatments, or preventatives for the virus | 19 |
| 4 | Satire | Satirical or humorous content | 7 |
| 5 | Prophecy | Suggests virus was previously predicted | 6 |
| 6 | Intolerance | Expresses racism, xenophobia, sexism, etc. | 1 |
| 7 | Politics | Political partisan content | 1 |

¹⁹ Examples of each visual is available in Annex A

The next step is the comparison of our results with those obtained in the English-speaking community study. Below in table 5 we have compared the distribution of misinformation visuals from our study with Spanish speakers with those of English-speaking. The first thing we observe is a similarity in the order of the most common frames. Thus, we find that the three most common frames are the same in both communities: authoritative agency, virulence, and medical efficacy. On the authoritative agency frame, we found 33% in the Spanish-speaking community compared to 40% in the study among the English-speaking community. Secondly, the virulence frame is almost equally relevant in both communities 33 and 33. In third place is medical efficacy, which is less relevant in the Spanish-speaking community 19 than in the English-speaking community 29.

From the fourth position onwards, we find differences in the order of the most common frames. In the Spanish community, we found prophecy with 6, compared to 10 and fifth place in the English community. In fifth place is Satire with 7, while in the English community it had 6 and was in sixth position. The sixth most named frame in the Spanish-speaking community is Intolerance with 1, while it was the fourth most named with 15 in the English-speaking community. Finally, in our study we have identified a new frame of visuals: politics (1) that was not reported in the English-speaking study.

In this analysis we found a clear similarity between the 3 frames with the highest volume for each language: authoritative agency, virulence, and medical efficacy. The contents classified in these three categories represent 85% of the cases studied in the Spanish language. For the English language, we can see that they also represent a good part of the visuals, but we do not have such an exact precision, as the contents could be classified in more than one frame. For the following frames, we find that intolerance was the fourth most abundant frame in English, for Spanish it is the sixth, followed by prophecy and satire, which has a greater weight in Spanish. In addition, for the Spanish study we have identified a new frame politics - with messages that seek to attack the political party in government -, which represents a very small part of the visuals. Therefore, we can affirm that the thematic misinformation visuals related to covid-19 identified between January and March 2020 are similar in English and Spanish.

Table 5. The Most Common Frames Spanish and English Compared

| | FRAME | SPANISH PREVALENCE (%) | ENGLISH POSITION | ENGLISH PREVALENCE (%) |
|----------|-----------------------------|-------------------------------|-------------------------|-------------------------------|
| 1 | Authoritative Agency | 33 | 1 | 40 |
| 2 | Virulence | 33 | 2 | 33 |
| 3 | Medical Efficacy | 19 | 3 | 29 |
| 4 | Satire | 7 | 6 | 6 |
| 5 | Prophecy | 6 | 5 | 10 |
| 6 | Intolerance | 1 | 4 | 15 |
| 7 | Politics | 1 | - | 0 |

With these data we can conclude that both for the English and Spanish language the distribution of frames and their prevalence are similar. This is related to other research that has looked at the relationship of different language communities and their tendencies in the emergence of alternative narratives.

In the last part of our analysis, we studied the relationship between the most common frames and the social platform where the visual was first identified (RQ3). In cases where a visual was flagged on more than one network, we manually went in to identify the platform on which it was first reported and assigned it as the source. After cleaning our dataset, we found four major platforms where the content was first shared: Facebook (83), WhatsApp (40), Twitter (39) and Instagram (1). Figure 1 shows the distribution of the visuals classified by the 7 frames and the 4 sources.

With the frame and source of the visual labelled we run a linear regression between these two variables for each of the 7 frames we have identified. In general, we found a low relationship between the content and the platforms on which it was first identified. This is the reason why, in this particular case, we have discarded the results as not conclusive.

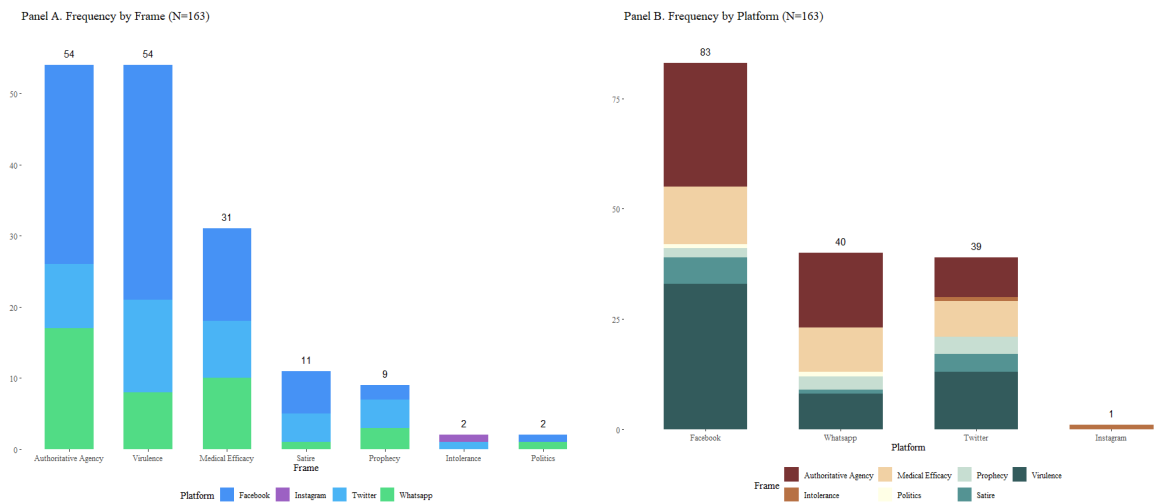


Figure 4. Frequency by: Frame (Panel A), Platform (Panel B).

Studying by platform (Figure 4) we see that WhatsApp is predominant in the authoritative and medical frames. Frames that are more related to the content of official government channels. The promoters of these messages took advantage of the fact that for the moment WhatsApp did not allow an official government presence on its platform. Instead, governments had a presence on Facebook and Twitter, so it was easier for citizens to verify information with the official channels of the institutions. WhatsApp is trying to solve it with the bots of public bodies.

Twitter is dominated by satire and prophecy frame content. On this platform the user publishes content in the open (mostly) and seeks to laugh, to make jokes that lead to public recognition. On Facebook, on the other hand, the frame with the most visuals are virulence, just behind Twitter. This content talks about the virus but without speaking on behalf of official bodies, as could be the case with the authoritative or medical frames that predominate on WhatsApp.

We see how there is a distinction between the most recurrent frames on each platform. The nature of each environment and its capabilities encourage the diffusion of one type of content on each platform. WhatsApp detected this and tried to change its nature by allowing the presence of official bodies on its platform. In the future, we could study the impact of this action and measure whether it encourages a reduction of visuals referring to authoritative and medical frames.

5.5 Discussion and conclusions

Communication is one of the crucial factors in the management of emergencies. In this context, the misinformation that arises represents a threat with great potential to negatively impact the health and safety of the population. Consequently, it is necessary to keep studying this phenomenon. Above all because the emergence of generative artificial intelligence exacerbates this threat, as anyone can create visual content with unauthenticated information pretending to be official and completely indistinguishable from those created by legitimate sources. In this study we have seen that the most common frame is precisely the impersonation of official messages. Bearing in mind that we are dealing with 2020 data today, generative AI makes this content easier to produce and of higher quality.

As regards the format of the content in visuals, images and videos, it is still more complicated to monitor. For the moment the best way is to manually evaluate each content and its veracity. This way of reviewing is not very scalable in terms of the volume of content shared in digital environments, nor is it scalable in terms of speed. In the same way that AI can help to generate new content, it is a field that could also evolve along the lines of interpreting the content of an image more easily and making an assessment. For both governments and platforms, AI can offer a range of tools that facilitate automated monitoring and moderation of visual content on social platforms.

With regard to language, we have found an evident relationship between the stories of the Spanish and English-speaking communities. This confirms the relationship between communities and the transfer of stories that other studies have also revealed. Despite this discovery, it is necessary to continue studying and, above all, to extend local studies. Since most of the research and resources on content moderation platforms are aimed at the English-speaking community. This interconnection teaches that combating misinformation requires action by platforms and governments at the local level for each region and language community. The evolution of global stories on a local scale makes it necessary for verifiers to understand the context of the region in order to do their work well. On the other hand, there also needs to be a proactive response from platforms and governments to monitor globally what is being said and prepare for it to reach their region.

Finally, we have also seen how disinformation must be further investigated with a multi-platform approach. We have seen how some platforms are more prone to one

type of frame than others, but they all feed back into each other as do linguistic communities. Social platforms need to open up their data to complement the task of monitoring and moderating content. Here governments must continue to play an active role. For example, in the European Union, the Digital Service Act (DSA) was introduced, which obliges platforms to cooperate with public authorities under the threat of fines running into millions of dollars. Without the help of the platforms, it is very difficult to study the phenomenon and to act appropriately. Besides the will of the social platforms there is also the threat of content being shared on spaces where the content is encrypted. This disables the platforms themselves from knowing what kind of content is being shared within their infrastructure. An example is WhatsApp where it is not possible to moderate the content being disseminated. Decrypting content does not seem to be a suitable solution because it infringes on people's freedoms. Proactivity is needed here and official messages need to overflow the misinformation in order to combat it. At the same time, it can also be useful to create channels and bots that are in charge of capturing the messages that citizens send to verify and become a tool for active listening. It seems especially interesting in WhatsApp where there is an increasing volume of conversation.

In summary, misinformation in emergencies and especially visual storytelling is a field that needs to be further explored. Both for the language part and for the multi-platform study a *glocal* approach is necessary. Knowing and monitoring on a large scale in order to be able to go down to the local level. Any alternative story that appears in the world has the potential to reach any corner of the world in its original form or adapted to one's own reality. Proactive platforms and governments are the only way to combat misinformation and ensure that it has the least impact on the health and safety of the population.

6. La gestió de la infodèmia a Catalunya durant el primer any de la pandèmia de la covid-19²⁰

6.1 Introducció

El 2020 la pandèmia mundial causada per la covid-19 ens va impactar a tots, independentment de característiques demogràfiques o econòmiques. Mentre estàvem confinats a casa, privats de llibertat com mai havíem imaginat, les xarxes socials i la missatgeria instantània es van convertir en la nostra única via de relació social. En aquest moment d'incertesa i desconcert tots intentàvem fer-nos una idea del que estava passant. A causa de la naturalesa recent del virus poca informació es podia considerar veraç i a vegades “la veritat” canviava d'un dia a l'altre. Diferents actors van aprofitar aquest context per generar i promoure l'aparició de relats alternatius amb informació no veraç o contrastada. De cop, Govern i responsables de la Salut pública van trobar-se amb un repte únic i de grans dimensions: mitigar els efectes catastròfics de la desinformació. Tal va ser la dimensió de la problemàtica que l'Organització Mundial de la Salut (OMS) va alertar i popularitzar el terme infodèmia. Fins i tot el president Obama va alertar amb el clar missatge “*disinformation is killing people*”²¹.

L'amenaça d'origen cibernètic poc tenia a veure amb les anteriors identificades per les agències públiques de seguretat. El seu fons no era tecnològic, com podria ser un accés no desitjat o una caiguda de servidors, sinó que el seu principal component era de relacions socials aprofitant els canals digitals. Per aquest motiu el repte de combatre la desinformació va recaure en unitats de comunicació i atenció ciutadana, en lloc de les unitats més enfocades a la tecnologia. En aquest article, entrevistem els actors més rellevants en la gestió de la infodèmia, principalment responsables d'unitats de comunicació i, de forma addicional, amb unitats de ciberseguretat. Estem d'acord

²⁰ Aquest article va ser publicat a *Quaderns del CAC*, 2022, Vol. 25, Núm. 48, p. 49-56, <https://doi.org/10.34810/qcac48id406106>. Octubre 2022.

²¹ <https://www.axios.com/2022/04/21/barack-obama-disinformation-social-media>

amb altres autors en què la desinformació és una amenaça que cal afrontar des de diferents enfocaments.

L'estudi que esteu llegint el vam fer a partir de set entrevistes semiestructurades durant març i abril de 2021, els dotze mesos posteriors a l'esclat de la pandèmia originada per la covid-19. Vam triar les persones per les seves responsabilitats laborals i àmbits d'actuació dels seus equips. Per complementar la visió de la gestió de la infodèmia a Catalunya vam ampliar les entrevistes amb persones externes a la gestió institucional. Per això ens vam entrevistar amb els representants de les dues principals organitzacions de verificació de fets (*fact-checking*) establertes al territori de parla catalana, Maldita i Verificat, que van tenir un paper rellevant en identificar i verificar informació relacionada amb la infodèmia.

Amb aquest article contribuïm a documentar com es va gestionar la infodèmia a Catalunya durant el primer any de la pandèmia. Expliquem el funcionament dels equips institucionals de la Generalitat de Catalunya, com identificaven la desinformació, quins procediments existien per contrastar la informació, com es coordinava la resposta de mitigació i s'avaluaven les intervencions. Alhora també aportem la visió complementària de les organitzacions de verificació de fets, els seus processos de captació, validació i resposta cap a la comunitat. Finalment, a les conclusions abordem el debat sobre la desinformació a Catalunya i els possibles camins a explorar per afrontar aquest repte. Aquest estudi de cas pot ajudar a altres institucions, públiques i privades, a preparar-se per mitigar de manera més eficient els efectes dels relats alternatius i disposar de més recursos per combatre noves infodèmies.

6.2 Marc teòric

El terme infodèmia²² va obtenir gran popularitat arran de la pandèmia provocada per la covid-19 de 2020. Tot i això, no és un terme nou, quasi vint anys abans ja hi havia literatura acadèmica parlant-ne, encara que les tàctiques i els mitjans disponibles eren diferents (Eysenbach 2002). En aquell moment els canals digitals per on

²² *Difusió ràpida de rumors, informació inexacta i notícies enganyoses amb relació a una malaltia infecciosa o un problema de salut pública paral·lelament a la seva propagació o evolució.* Termcat

circulaven els missatges eren els webs personals. Amb l'avanç tecnològic i l'explosió de web social, també conegut com a web 2.0, es va democratitzar la publicació de continguts a internet, ja no calien coneixements informàtics, sinó que qualsevol persona podia amb poc esforç convertir-se en un canal d'emissió d'informació i fins i tot ser influent (Bakshy et al. 2011). A partir de l'aparició de web social, Governos i institucions públiques adopten noves eines per relacionar-se amb la ciutadania. Es comença a veure la utilitat dels canals digitals, especialment les xarxes socials com a eines per a gestionar i estudiar les emergències (Palen i Anderson 2016), quan és més necessari que la informació veraç arribi el més ràpid possible al màxim nombre de persones. Es valora la facilitat i velocitat amb què la naturalesa d'aquestes plataformes permeten que els missatges circulin a gran velocitat entre els usuaris que comparteixen la informació. Apareix el camp científic de *crisis informàtics* (Palen et al. 2020) que estudia la gestió de les emergències a partir de l'activitat a les xarxes socials.

L'ús massiu de les plataformes digitals va provocar un canvi massiu en el consum d'informació. Sense arribar a substituir els mitjans de comunicació de massa (Nielsen et al. 2020) les xarxes socials es va convertir en un entorn on persones i institucions acudien per obtenir la informació. Al mateix moment l'expansió del nombre de canals emissors també va incrementar la capacitat d'arribar a més persones dels canals que compartien informació no veraç. A mesura que el problema es va anar fent més gran i es va veure el seu efecte en el comportament social, les eleccions als Estats Units d'Amèrica de 2016 marquen un punt d'inflexió, es va anar prenent consciència de la dimensió del fenomen del trastorn informatiu – conegut en anglès com *Information disorder* (Wardle i Derakhsan 2017).

A més dels punts de vista ja comentats cal afegir l'impacte en la seguretat de la ciutadania en els canals digitals. La visió clàssica de la ciberseguretat té una mirada tecnològica, on persones amb capacitats tècniques estudien amenaces i presenten solucions a partir de les seves habilitats. Com hem vist la naturalesa de la infodèmia és de caràcter social i de comunicació, amb un impacte directe al benestar, seguretat i salut de les persones (Cotter et al. 2022). Tot i que, en part, utilitza els canals digitals per aconseguir els seus fins no es pot considerar que s'hagi de resoldre únicament tecnològicament. La literatura acadèmica ens ha mostrat com cal un abordatge multidisciplinari (Briand et al. 2021) per combatre la infodèmia en general i les operacions informatives particulars (Francois i Lin 2021).

Com hem vist la infodèmia és un fenomen social, i com a tal les dinàmiques i comportaments són diferents a cada comunitat local, per això cal estudiar-la cas a cas per millorar l'eficàcia de les intervencions (Calo et al. 2021). Actualment, la majoria de literatura acadèmica sobre la infodèmia i la desinformació estudien la comunitat de parla anglesa, sobretot als Estats Units i Regne Unit. Alguns autors de prestigi internacional (Tucker et al. 2018) i institucions (World Health Organization 2021) han assenyalat com una necessitat important estudiar el comportament en altres llengües i territoris per ampliar el coneixement del fenomen i mitigar-ne millor els efectes. Només cal repassar les referències d'aquest article per adonar-se que hi ha una manca de literatura sobre aquest fenomen que es fixi en comunitats de parla no-anglesa.

Pel que fa la llengua castellana, trobem alguns estudis que han estudiat el fenomen, especialment centrats a l'estat espanyol. Alguns estudis han focalitzat el seu interès en l'anàlisi del contingut identificat com a no verídic (Salaverría et al. 2020) i d'altres des d'una perspectiva qualitativa, a partir de l'experiència i percepcions dels verificadors establerts al territori (López-García, Costa-Sánchez, i Vizoso 2021).

En l'àmbit de Catalunya s'ha estudiat la gestió d'emergències i desinformació en canals digitals en poques ocasions (Espineta, Garcia-Alsina, and Canals 2015; Xaudiera and Cardenal 2020). Pensem que per millorar la capacitat de resposta és necessari conèixer amb més detall l'afectació que té la infodèmia en la nostra comunitat. A partir de la necessitat de conèixer com és el fenomen de la infodèmia a Catalunya, ens preguntem **(RQ1) com les institucions públiques, especialment la Generalitat de Catalunya, va combatre la infodèmia durant els primers dotze mesos de la pandèmia**. Amb aquesta voluntat ens vam entrevistar amb diferents persones, de diferents àmbits, per documentar i entendre com va ser la resposta de l'Administració de la Generalitat. Alhora, per fer-nos una imatge de la situació més general en la gestió de la infodèmia també vam entrevistar responsables de les dues principals organitzacions de *verificació de fets* amb presència a Catalunya. L'objectiu de les trobades és **(RQ2) identificar aquells aspectes del flux de treball de les organitzacions de fact-checking que poden ajudar al sector públic en la gestió de la infodèmia**. La informació d'aquestes entrevistes l'hem estructurat en les 4 fases principals que l'OMS recomana per combatre la infodèmia: detecció, verificació, resposta i avaluació (World Health Organization 2021).

Existeixen diferents propostes per combatre la infodèmia i mitigar-ne els efectes. En aquest article ens centrem en el potencial de les xarxes socials per la seva facilitat d'arribar a grans audiències de manera efectiva gràcies a la capacitat amplificadora de les plataformes (Houston et al. 2015). En altres territoris s'han elaborat conjunts de propostes generals per gestionar la infodèmia (Eysenbach 2020) i sobretot ser eficients en cobrir els buits informatius, fins i tot quan la informació encara no es pot verificar (Gorman i Scales 2022). Aquestes tècniques proposades a escala global són tan diverses com: crear un banc comú de coneixement (Watts et al. 2021), ajudar-nos dels mitjans de comunicació (Marx, Mirbabaie, i Ehnis 2018), o estudiar els trets demogràfics i psicològics que fan que un contingut es torni viral (Pennycook et al. 2020). Com hem dit, en tractar-se d'un fenomen social, cal estudiar el comportament per a cada regió i comunitat. Finalment, a les conclusions donem un seguit de respostes a la pregunta **(RQ3) com ens podem preparar per afrontar en millors condicions la pròxima infodèmia.**

6.3 Mètodes

Entre el 18 de març i el 18 de maig de 2021 vam portar a terme una investigació qualitativa a partir d'entrevistes semiestructurades amb responsables de diferents àrees impactades per la desinformació durant la gestió de la pandèmia. Amb aquestes entrevistes vam obtenir informació de primera mà sobre com es va gestionar la infodèmia a Catalunya, en especial durant els primers dotze mesos de la pandèmia provocada per la covid-19. A partir de la seva experiència vam documentar la gestió de la desinformació, poder-ne dibuixar les línies mestres i mitigar-ne els efectes de manera més eficaç en properes infodèmies.

Cada una de les entrevistes semiestructurades les vam dividir en quatre parts, una part cada fase de l'operativa per combatre-les els efectes de la infodèmia: detecció, verificació, resposta i avaluació. Dins de cada part teníem diferents preguntes i temes a tractar en funció de la naturalesa de l'entrevistat. Així doncs, en la fase de la detecció ens interessava conèixer les formes en què monitoraven les xarxes, les col·laboracions publicoprivades i si disposaven de programari específic per dur a terme la tasca. En la fase de verificació volíem descobrir els procediments interns, de la Generalitat, i externs, dels verificadors, per comprovar la validesa de la informació del contingut. En la tercera fase de resposta, com es coordinen els canals per aconseguir que la comunitat obtingui la informació verídica, aprofitant tots els canals a l'abast, quins

mitjans utilitzen i com recullen tota la informació verificada. Finalment, a la fase d'avaluació, les preguntes estaven enfocades en quins indicadors i mètodes es fan servir per mesurar l'impacte i influència de les operacions.

En primer lloc, vam entrevistar a treballadors de la Generalitat de Catalunya amb responsabilitat en diferents unitats amb relació a la gestió de la comunicació de la pandèmia. Les unitats on treballaven aquestes persones eren: atenció ciutadana, departament d'interior, departament de salut, protecció civil i emergències i l'agència de ciberseguretat. Es van triar aquestes àrees per la relació directa amb el problema a estudiar: la desinformació durant la pandèmia sanitària i la resposta de la Generalitat per assegurar la seguretat de la ciutadania. Aquestes entrevistes ens van permetre aprofundir en la gestió de la crisi i dels riscos associats des de diferents punts de vista interns de l'administració pública.

Tot seguit, vam mantenir entrevistes amb persones que combatien la infodèmia dins l'àmbit català des de fora el sector públic. Vam parlar amb persones de les dues principals organitzacions de verificació de fets (*fact-checking*): Maldita i Verificat. El punt de vista extern ens va aportar una visió complementària més centrada en la identificació i verificació de la informació i ens va permetre analitzar la col·laboració entre entitats públiques i privades en la seva missió d'assegurar un ecosistema informatiu saludable.

Table 6. Persones entrevistades i càrrec en el moment de l'entrevista

| Organisme | Persona | Càrrec |
|----------------------------------|----------------------------------|--------------------------|
| Departament d'Interior | Marc Homedes | Director comunicació |
| Departament de Salut | Rosa Romà | Directora d'innovació |
| Protecció Civil | Sergi Delgado | Subdirector general |
| Atenció Ciutadana | Jordi Graells | Director general |
| Agència de Ciberseguretat | Xavier Panadero | Responsable d'operacions |
| Maldita | Pablo Hernández | Responsable acadèmic |
| Verificat | Alba Tobella i Lorenzo Marini | Fundadors |

Finalment, vam analitzar la informació obtinguda dels protagonistes i la vam agrupar en les 4 fases de gestió de la infodèmia: detecció, comprovació, resposta a la desinformació i avaluació. Aquesta síntesi ens va permetre fer-nos una idea clara de com va ser la gestió de la informació durant els 12 primers mesos de la pandèmia. Totes aquestes experiències i idees estructurades ens van permetre de dibuixar possibles fulls de ruta i intervencions per millorar l'efectivitat en properes infodemies.

Com a darrer punt metodològic volem ressaltar un aspecte terminològic important. En anglès es diferencia entre *disinformation*, quan es detecta una voluntat expressa de l'emissor d'influir l'opinió pública amb contingut volgutament no verídic, i *misinformation* quan es desconeix la voluntat de l'emissor, o bé aquest publica un contingut no verídic sense conèixer la seva autenticitat. En aquest article es parla de desinformació com a concepte general per manca d'un millor terme en llengua catalana, tot i que ens referim a l'accepció anglesa de *misinformation* en què l'emissor publica el contingut sense contrastar-ne la veracitat.

6.4 Descobriments

6.4.1 Detecció

El primer aspecte en què vam centrar la nostra investigació era com les diferents unitats públiques detecten i identifiquen la desinformació. L'escolta activa digital és una activitat bàsica de les institucions per conèixer els relats que sorgeixen entre la ciutadania i organitzacions. Tot i treballar en diferents unitats i departaments, els participants van reportar problemes similars i van assenyalar una manca d'eines i coordinació interna quan es tracta d'identificar relats alternatius quan sorgeixen en les plataformes digitals. Cada unitat feia per ella mateixa un monitoratge reactiu bàsic de la conversa en els entorns digitals, com una tasca més i sense una eina específica ni persona responsable del seguiment. Això significa que el mateix relat pot ser identificat i estudiat des de diferents àrees de la mateixa organització.

Un dels entrevistats va assenyalar una pràctica habitual entre la gent del seu equip “a conseqüència de no tenir suficients recursos, sovint detectem els relats alternatius quan els mitjans de comunicació ens pregunten sobre la seva veracitat”. Els cercles personals a vegades també actuen com a font d'informació “a vegades un conegut m'envia un missatge de WhatsApp preguntant-me sobre l'autenticitat d'un contingut que li ha estat reenviat”. Hi havia, per tant, un treball reactiu en la detecció

i identificació dels relats en què en moltes ocasions el senyal d'alerta prové de factors externs. Un d'aquests factors externs són els grups de voluntaris digitals en emergències (VOST). “Aquests grups, en els quals participem activament, actuen com a sensors socials amplificats. Ens permeten tenir més orelles en grups més diferents”.

D'altra banda, els *fact-checkers* que actuen al territori tenen un flux de treball molt diferent. En el cas de Maldita, a més de tenir en plantilla persones entrenades per detectar i identificar la desinformació, també formen part d'una comunitat d'usuaris en què cada usuari actua com a sensor social per detectar desinformació en els seus cercles socials. “El *crowdsourcing* és la clau de l'èxit de Maldita” per poder identificar grans volums de contingut per verificar. “Llavors, a través dels nostres bots de WhatsApp i Telegram, ens envien el contingut que han rebut, perquè nosaltres comprovem la seva credibilitat. Això ens permet identificar moltes de les principals històries que sorgeixen durant l'emergència”. En especial, com a canal per rebre contingut susceptible de ser no veraç, els va funcionar molt bé el bot de WhatsApp “el bot de WhatsApp és la nostra principal font per detectar contingut per ser verificat”. A Verificat ens van comentar que també disposen d'un xatbot, si bé amb menys activitat que el de Maldita, però que representa un 80% de l'entrada de contingut per ser verificat.

Des de Maldita ens van comentar que feia poc que havien establert un conveni de col·laboració amb Facebook²³. “Quan els usuaris de la seva plataforma reporten un contingut com a desinformació, Facebook l'envia a Maldita per a la seva verificació, i si resulta ser no verídica, l'algoritme penalitza el contingut per reduir la seva exposició i amplificació”. Encara que es tracti d'una canal complementari i el bot de WhatsApp sigui el seu principal canal de detecció “aquest canal representa un volum molt inferior d'entrada, però representa una manera de fer escolta activa”. A més de l'impacte directe que té l'acció en la difusió del contingut. Com a detall a destacar, van comentar com encara es pot treballar en la millora del procés per fer-lo més eficient i que l'impacte en la distribució de contingut sigui més immediat.

²³ Organització actualment coneguda com a Meta, companyia matriu de Facebook, Instagram i WhatsApp.

6.4.2 Verificació

A la segona fase del cicle de gestió de la infodèmia durant les emergències hi ha la verificació del contingut. En aquesta fase és molt important la velocitat amb la qual una vegada es detecta el contingut s'és capaç de discernir si es tracta d'informació veraç o es pot classificar com a desinformació. En aquesta fase vam trobar, de nou, que cada unitat treballa de manera independent. Tot i que el Govern disposa dels operatius sobre el terreny per verificar cada relat; “tenim la capacitat de verificar cada contingut que rebem, però hem de seleccionar i escollir quines narratives verifiquem”. Una vegada la informació ha estat validada, cada unitat dona la resposta que creu més adient. “El problema apareix quan la informació no pot ser verificada pel nostre equip, per exemple quan fa referència a informació d'un altre departament. En aquest cas la verificació és més lenta. La manera de treballar en estancs verticals com treballem fa la transmissió de la comunicació interna més complicada”. En el cas de la pandèmia de la covid-19 existia l'agreujant que no es disposava d'informació consolidada “la veritat ens canviava dia a dia. El que una setmana informàvem com a verídica podia ser que es desmentís a la següent”. Aquest aspecte feia encara més complicat el monitoratge dels relats alternatius que apareixen. “Sobretot durant les primeres setmanes el desconcert era molt gran, la informació que ens facilitaven els experts sobre les mascaretes era molt canviant”.

En ser preguntats per un repositori amb les verificacions, una pràctica emprada per altres organismes públics com el Govern del Regne Unit, tots els entrevistats van estar d'acord en la seva utilitat. “Ens hem trobat en situacions en les quals, tot i treballar en la mateixa unitat, les persones que atenen cada canal: xarxes socials, mitjans de comunicació, atenció telefònica han de buscar i moltes vegades verificar la validesa de la informació per ells mateixos”. A més de la ineficiència del procés “també pot perjudicar la coherència del missatge institucional, com ha passat en algunes ocasions en què des de canals institucionals s'ha donat diferent informació sobre un mateix contingut”.

Els participants en aquesta investigació també van assenyalar la recurrència dels relats alternatius: “hem identificat històries falses similars en diferents crisis. A vegades la mateixa història apareix en diferents onades durant la mateixa crisis. En altres ocasions hi ha relats que es repeteixen en cada emergència i fins i tot alguns relats han sorgit durant crisis en anys diferents”. La creació d'un repositori històric

permetria als operatius de donar resposta ràpidament consultant si un mateix relat ha aparegut en altres emergències i la resposta que es va donar. Aquest conjunt de resposta estructurada facilitaria una cohesió en la resposta institucional. A més a més, “si aquest repositori fos obert i la ciutadania pogués verificar per ella mateixa ens podria ajudar a reduir el nombre de consultes i optimitzar els nostres recursos”.

En l'àmbit extern, els representants de les organitzacions de fact-checkers estan d'acord en el fet que hi ha una manca de comunicació directa amb les institucions públiques per dur a terme les verificacions. En aquest sentit, “un tractament com els dels mitjans de comunicació, o una canal de xat amb un contacte directe seria molt útil”. Actualment, no hi ha cap persona de contacte amb qui verificar els relats alternatius que detecten i han d'assumir de manera independent tot el procés de verificació en paral·lel amb el qual estiguin fent les institucions públiques. En alguns casos aquest doble procés de verificació pot aportar valor, però en la majoria dels casos és un format poc eficient en què es dupliquen esforços per múltiples canals. Per verificar la informació, a Maldita disposen d'una xarxa d'experts en diferents temes.

6.4.3 Resposta

El següent pas, un cop detectada i verificada la informació dels relats alternatius, cal donar resposta i dur a terme un seguit d'intervencions. En aquesta fase trobem altra vegada una manca de coordinació entre les diferents unitats. De la mateixa manera que detecten i verifiquen els relats de manera independent, tampoc hi ha cap estratègia transversal en la resposta de la desinformació en crisis. Això fa que trobem diferents respostes sobre les mateixes narratives en funció de qui respon i de la informació que la unitat té disponible.

La manera més utilitzada de resposta, és informar directament a les persones que han reportat la informació. En cas que es consideri informació rellevant també es publiquen missatges públics perquè arribin a tota l'audiència. En algunes ocasions, algunes unitats han llançat campanyes de publicitat segmentades per fer arribar el missatge a les comunitats més sensibles, però no per respondre directament a un relat alternatiu. Els entrevistats assenyalen diferents tècniques “publiquem un missatge quan identifiquem que una història està guanyant popularitat”, “informem el mitjà de comunicació quan identifiquem una informació que han publicat com a no veraç”, però tots estan d'acord que “hi ha una manca de recursos per respondre a tots els missatges i fer seguiment de tots els relats. Twitter és una eina útil per informar ràpidament a les

audiències, però no per resoldre consultes individuals durant les primeres fases de l'emergència. Simplement, no disposem dels recursos humans necessaris". Sobre els missatges molts van remarcar la importància de crear "peces visuals, infografies amb les idees clau. Aquestes tenien missatges planers i moltes vegades provenien de les preguntes més freqüents (PMF) que rebíem a través dels diferents canals d'atenció ciutadana, digitals, presencials i telefònic". El risc d'aquest contingut era que circulés quan la informació ja no era vàlida, i fins i tot "vam trobar peces que imitaven el nostre disseny amb informació no veraç".

En aquest cas, col·laborar amb els mitjans de comunicació tradicionals ha estat profitós "quan podem verificar un relat alternatiu que ens ha arribat a través dels mitjans, els ajudem a desmentir el missatge i amplifiquem el seu contingut. Hi ha canals institucionals amb audiències més àmplies que les dels mitjans. En alguna ocasió, hi ha hagut intents de coordinar una resposta institucional "en alguna ocasió hem intentat d'aprofitar l'audiència orgànica acumulada, la suma dels seguidors, per amplificar missatges de manera coordinada". També es van establir col·laboracions amb mitjans tradicionals innovadors "vam anar al programa de ràdio més sentit a Catalunya a respondre qualsevol dubte que tingués l'audiència. Aquestes intervencions van resultar molt útils".

Internament també es va treballar amb els prescriptors. "Des del departament de Salut vam preparar els públics interns per convertir-se en prescriptors dels missatges oficials". En tractar-se d'una crisi d'origen sanitari els metges es van convertir en referents i en amplificadors de missatges oficials. "Manteníem reunions periòdiques, els hi facilitàvem missatges i contingut per compartir entre els seus contactes i en els seus canals públics". Des del departament van col·laborar també amb col·legis professionals, diferents influenciadors i periodistes "vam crear un grup de WhatsApp on responíem qualsevol dubte o informació en qüestió de minuts. No podíem deixar que un altaveu tan important com els mitjans de comunicació tinguessin cap dubte sobre la informació que publicaven".

Un altre aspecte molt rellevant en la mitigació de la infodèmia va ser arribar a aquells públics més sensibles, en un doble vessant, per tractar-se dels més sensibles per la malaltia o per tenir menys capacitats digitals i poder discernir entre informació verídica o falsa. Es va dur a terme una campanya massiva informativa amb la gent gran

a través dels missatges de text (SMS) i de material gràfic en centres de salut, farmàcies i qualsevol espai on poguessin arribar a la gent d'edat més avançada.

Com a públic d'interès clau també es va identificar la gent d'origen en altres territoris i llengües. “Aquestes persones tenen altres canals d'informació, que estan lluny dels que més estretament col·laborem”. Per tal de poder fer-los arribar els missatges es va contactar amb diferents líders d'opinió de les comunitats per explicar la situació tal com anava evolucionant. “Volíem que tinguessin l'última informació veraç i que la poguessin compartir entre els seus contactes”.

Pel que fa als fact-checkers la resposta que donaven anava independent de la que donaven les institucions públiques. El seu impacte era gran dins la seva comunitat i en alguns casos complementari respecte a les audiències dels canals institucionals. Tot i això, el seu catàleg de recursos per arribar a diferents públics és molt inferior del qual disposen les institucions públiques.

6.4.4 Avaluació

En la darrera fase, cal avaluar la gestió de la infodèmia i de les accions proposades. Aquesta fase no és pròpiament dins la gestió de l'emergència, però és important perquè les institucions es puguin preparar millorar pel repte de la infodèmia. En aquest sentit, van dir “trobem a faltar una anàlisi posterior dels relats que apareixen a les xarxes. Sens dubte, que conèixer com es propaga la desinformació ens seria útil per preparar-nos per a properes emergències”, “entendre la dinàmica d'una història viral a les plataformes digitals seria molt útil per millorar les nostres operacions”, “segur que ens ajudaria a estar millor preparats”. Un altre entrevistat comentava “tenir identificats canals que propagaven informació no veraç ens facilitaria poder-los monitorar en propers pics informatius”.

La manca de recursos humans per combatre la infodèmia és una constant. És per això que en aquests casos es destinen el màxim de recursos a la verificació i resposta, i queden descobertes altres fases com són la identificació i l'avaluació. A més tots els entrevistats han estat d'acord en una manca de personal amb habilitats digitals especialitzades en aquest camp, inclús des de l'Agència de ciberseguretat van dir que una millor capacitat digital ajudaria en totes les fases de mitigació, “entenem l'amenaça de la desinformació, però la gestió de respostes durant l'emergència consumeix tots els nostres recursos”. La tasca de l'agència de ciberseguretat “se

centrava a actuar com a tallafocs especialment en casos de *phishing*. Vam identificar un bon nombre de canals que intentaven recopilar dades personals a través de les plataformes digitals”.

D’entre els entrevistats només l’agència de ciberseguretat i Maldita tenen en plantilla persones amb capacitat tècnica de programació, un entrevistat va admetre que “seria molt bo disposar de personal amb capacitats tecnològiques per ajudar-nos a millorar la nostra resposta”. Va ser una opinió unànime entre els entrevistats, que els seria un recurs molt útil de disposar de persones amb capacitat d’analitzar i treballar amb la informació que ofereixen les plataformes digitals, el que en el món acadèmic es coneix com a ciències socials computacionals.

6.4.5 Conclusions i discussió

La infodèmia i la desinformació en situacions d’emergència és una gran amenaça. Tot i no tractar-se d’un fenomen nou, les noves tecnologies a l’abast de qualsevol persona han multiplicat exponencialment el seu impacte. A més, en tractar-se d’eines que presenten novetats constantment és complicat mantenir un coneixement actualitzat de les plataformes i dels mètodes per poder-les mitigar. La primera conclusió és que es tracta d’un repte majúscul per la seva naturalesa canviant. La infodèmia que documentem en aquest article, les tàctiques i intervencions, poc tenen a veure amb les que es van dur a terme durant les eleccions americanes de 2016.

Com hem vist, les institucions públiques no estan preparades per combatre de manera eficient aquest repte. Tot i les millores en la gestió i l’experiència que es va guanyant en superar cada cas, els mateixos responsables admeten que les persones que els hi toca gestionar la infodèmia no disposen dels recursos ni les capacitats per fer-ho de la manera més eficient. Seria bo plantejar la creació de nous espais i perfils arreu de l’organització amb la capacitat per treballar amb mètodes socials computacionals i poder treballar en temps real amb la informació provinent de les plataformes socials. En aquest sentit, podria ser una bona idea que les organitzacions de verificació de fets impartissin formació i compartissin el seu coneixement, sobretot en les fases d’identificació i verificació de la informació.

També hem vist com una oportunitat de millora millorar la col·laboració entre institucions públiques i les organitzacions de *fact-checking*. Especialment per la seva capacitat de mobilització entre les seves audiències i els recursos tècnics dels quals

disposen. De fet, algunes plataformes ja han signat convenis de col·laboració amb verificadors per la seva agilitat a l'hora de verificar. Pensem que si es fomenten sinergies de treball col·laboratiu, pot esdevenir un guany incalculable per la societat. A nivell més pràctic seria interessant la creació d'un repositori on es recullin els continguts verificats. Vam trobar que la coordinació interna, entre unitats i departaments, i externa és quasi inexistent, i per això pensem que la creació d'aquest punt de trobada seria un recurs senzill d'implementar amb un gran impacte. Aquest espai podria ser obert, la ciutadania podria consultar els continguts i comprovar ella mateixa l'autenticitat de la informació. Alhora que internament seria molt útil per coordinar els esforços de les diferents unitats de la Generalitat.

Pel que fa a la detecció i identificació de relats alternatius l'administració podria replicar les tàctiques que fan servir els verificadors de fets. Obrir xatbots i bústies des d'on la ciutadania pugui enviar contingut que es trobi i que dubti de la seva veracitat. A més d'un millor sistema de monitoratge que permeti detectar els relats de manera proactiva quan comencen a guanyar popularitat. S'ha demostrat que la velocitat en la resposta és un factor rellevant en el recorregut que tenen els continguts no verificats.

Finalment, cal tenir en compte que durant una emergència el coneixement i la informació contrastada és canviant. Això significa que un missatge que el dia 1 era no veraç, es pot demostrar erroni i convertir-se en veraç el dia 2. Aquest és un repte cabdal en emergències i especialment ho va ser en la gestió de la covid-19, ja que pot provocar una pèrdua de confiança en les fonts oficials. Per combatre es recomana el màxim de transparència, tant per explicar els processos interns com en publicar el contingut tan aviat com es tingui. Mantenir un flux constant d'informació, fins i tot quan cal dir que no es disposa d'informació, també ajuda a establir un vincle de confiança amb la ciutadania.

El repte de la infodèmia no té una solució única sobretot en els entorns digitals que són tan canviants. És per això que cal que les institucions públiques, de manera decidida, emprenguin accions per adaptar-se, de la mateixa manera que verificadors externs i les mateixes plataformes fa temps que fan. En el cas concret de la Generalitat caldria veure a quina unitat correspon el lideratge, especialment en les fases més tècniques de monitoratge i avaluació, i estudi dels relats. Caldria trobar un lloc central des d'on un equip de personal qualificat pugui donar suport a les diferents unitats que gestionen els àmbits més especialitzats.

7. Conclusions

La desinformació és una amenaça que ha entrat a l'agenda política en els darrers anys. En aquest context tant institucions governamentals internacionals com regionals com les mateixes plataformes socials han dut a terme un seguit d'accions per mitigar-ne els efectes negatius. Les plataformes socials no viuen alienes a l'amenaça dels relats alternatius que circulen per les seves plataformes. En els darrers anys han dut a terme diferents accions com crear unitats dedicades a fomentar la moderació de continguts col·laborant amb fact-checkers de diferents regions o la creació d'òrgans centrals per ajudar a decidir el posicionament de la plataforma sobre relats concrets. A més també han establert mecanismes de col·laboració amb l'acadèmia. Per exemple estudiant canvis en l'algorisme d'ordenament de contingut respecte a la desinformació (Nyhan et al. 2023), com mostra el contingut (Guess et al. 2023a) o com afecta la propagació entre iguals (Guess et al. 2023b). Tot i això, la majoria d'estudis interns i en col·laboració amb l'acadèmia tenen una visió centralista. Estan, com les mateixes plataformes sobretot alineades amb els seus objectius comercials, dirigits des dels Estats Units d'Amèrica i estudien principalment relats i converses en llengua anglesa. Meta i altres plataformes no destinen prou esforços en estudiar ni identificar relats alternatius en els idiomes que no siguin l'anglès (Valencia 2021). Aquest biaix provoca una desigualtat en el tractament de relats en funció de l'idioma i l'origen de propagació d'aquests relats. Per evitar aquest biaix i obligar les plataformes a vetllar per la seguretat i la salut d'ecosistemes informatius saludables arreu algunes regions han regulat com les plataformes socials han d'actuar en el seu territori. Per exemple la Comissió Europea va publicar un pla contra la desinformació i va regular com han d'actuar les plataformes en aquest sentit en el *Digital Services Act* (European Commission 2023). Aquest regalament obliga les plataformes a compartir públicament dades i informació sobre l'ús de les seves plataformes, gestionar la desinformació i col·laborar amb investigadors independents de la regió per estudiar l'impacte en un àmbit local.

En aquesta tesi hem analitzat el fenomen de la desinformació a Catalunya des de diferents perspectives. Per entendre com es propaga, quin és el tipus de contingut més predominant. Amb l'estudi de camp fet hem realitzat un conjunt d'entrevistes per

conèixer com es combat des del Govern de Catalunya. A més de les conclusions concretes de cada estudi en l'àmbit de resum hem trobat les següents conclusions:

- Les fonts creïbles com representats públics, polítics o mitjans de comunicació són un factor clau en la disseminació de desinformació. Quan són fonts creïbles qui propaga la desinformació l'impacte negatiu en aquella comunitat és superior i els missatges arriben a una major audiència. Les fonts creïbles juguen un paper fonamental en la visibilitat del contingut.
- Cal que les fonts oficials juguin un paper proactiu en el seu territori. Tot i la baixada en la confiança pública, governs i mitjans públics encara són la millor font des d'on obtenir la informació oficial. En el context actual cal que les institucions tinguin canals coneguts per tal de facilitar que les persones hi puguin accedir per verificar informació quan ho necessiten.
- Els canals dels governs són útils per combatre la desinformació. La gent que té els canals oficials dins la seva dieta informativa promou menys la desinformació i alhora contraresta més quan es troba contingut no veraç.
- Les comunitats lingüístiques tenen particularitats i clares connexions. Els relats alternatius es traslladen entre regions i països de manera ràpida i a vegades amb adaptacions en el seu àmbit local. Per tant, cal una visió *glocal*, una mirada global i una acció local, per combatre la desinformació. Els relats són similars entre diferents comunitats durant el mateix període.
- És clau que els governs engeguin sistemes de monitoratge per identificar els relats alternatius que es propaguen per les plataformes socials. En aquest sentit, pot ajudar a establir col·laboracions amb les organitzacions independents de *fact-checking* regionals per monitorar i contrastar la informació de manera col·laborativa.
- La desinformació en situacions d'emergència és un fenomen en creixement i es preveu que augmenti en els pròxims anys. Tot i els esforços de governs i plataformes socials per detectar i combatre la desinformació, els avanços en tecnologia com la intel·ligència artificial generativa permeten crear de manera més senzilla contingut fals indistingible del verídic. En aquest context la veracitat de la font és l'únic que permet atribuir veracitat en el contingut.

8. Treball futur

Durant les investigacions realitzades hem constatat l'amenaça que representa la desinformació en situacions d'emergència. Es tracta d'un àmbit de recerca en la seguretat i salut de les persones que creix i evoluciona amb les tendències de comportament social digital alhora que pateix canvis constants provocats per l'aparició de les noves eines tecnològiques. És un àmbit en constant evolució i, per tant, cal que la investigació en aquest camp no deixi d'abordar els nous reptes que es plantegen i se continuaran plantejant. Aquesta tesi és una foto estàtica d'un moment determinat que ben segur canviarà els pròxims anys. De fet, la mateixa tesi ha patit una evolució forçosa pels canvis enormes d'ençà que es va idear el 2019 fins avui 2024. Hem recollit els aspectes que pensem que són més rellevants de cara al treball futur en aquest àmbit.

En primer lloc, el paper de les fonts oficials, tant canals institucionals com de polítics o càrrecs electes, en la propagació de desinformació i relats alternatius. Les persones amb més poder són en moltes ocasions les que propaguen rumors i relats alternatius que els afavoreixin políticament més enllà de la seva veracitat. En tractar-se de persones amb grans audiències i exposició mediàtica tenen la capacitat de fer aflorar en el debat públic relats minoritaris. A més acostumen a ser identificades com a fonts creïbles. A més, el seu paper respecte a la recerca de la desinformació també és clau. Sovint alguns càrrecs públics han intentat condicionar a investigadors i les persones que s'encarreguen de gestionar la moderació de continguts de les plataformes socials (Roth 2023). És un alt risc per a la llibertat d'expressió i investigació i els posicionament i la influència dels poders oficials.

Un altre aspecte a tenir en compte és l'accés a les dades generades a les plataformes socials. Tot i que a l'inici d'aquestes compartien a través de les seves apis la informació de manera oberta amb investigadors i desenvolupadors la tendència és restringir l'accés a les dades (Bruns 2019). Aquesta tendència ha provocat que plataformes com Twitter, que en el seu moment era un entorn sobre el qual es realitzaven molts estudis, deixin de ser una font influent per la recerca (Valero 2023). Per tal de garantir la transparència de les plataformes algunes institucions públiques com la Comissió Europea (CE) ha regulat al respecte. En aquest cas el Digital Services Act (DSA) obliga a les grans plataformes socials de compartir dades dels seus usuaris

perquè investigadors independents puguin fer recerca. Tot i això, Twitter ha incomplert diverses vegades²⁴ amb la regulació no facilitant accés a les dades. De fet, algunes de les investigacions fetes en aquesta tesi resulten impossibles de realitzar amb les noves normes que regulen l'accés a les dades de Twitter.

Una altra amenaça del futur és el nou ecosistema de plataformes socials més fragmentat. Els temps en què la major part de les converses digitals passava en unes poques plataformes socials com Twitter, Facebook o fins i tot Instagram ha passat i no tornarà. L'escenari a 2024 és molt més fragmentat amb nous actors com Mastodon, TikTok, o Threads, que guanyen nous usuaris i prometen una internet sense grans audiències centralitzades en un sol entorn (Ben Thompson 2023). La nova realitat digital és, per tant, menys centralitzada i més multiplataforma. Aquesta diversitat d'eines provoca un accés a la informació més fragmentat (Newman et al. 2023) i calgui multiplicar els esforços de monitoratge, moderació de continguts i de resposta per tantes plataformes com calgui. Aquest fet incrementa considerablement el cost de les accions que les institucions públiques duen a terme per mitigar els efectes negatius de la desinformació i els de l'acadèmia per comprendre i dimensionar l'impacte de la desinformació.

En l'ecosistema informatiu digital també trobem la tendència a desplaçar les comunicacions en plataformes socials més privades com la missatgeria instantània, WhatsApp i Telegram, o les comunitats de xat privades com Slack o Discord. Aquestes plataformes no acostumen a facilitar accés a les seves dades a través d'apis. Aquest desplaçament fa que no puguem accedir a la informació que circula per aquests entorns, per això cal que busquem noves maneres per estudiar-ho compatibles amb la privacitat (Malhotra 2020). A més, moltes d'aquestes plataformes encripten el contingut que comparteixen els usuaris. Això fa que no hi hagi manera de saber de manera automatitzada de quin contingut es tracta ni de fer un monitoratge a gran escala. Alguns governs han treballat, de moment sense èxit, per saltar la mesura. Així doncs, el Regne Unit amb la regulació de *Online Safety Bill*²⁵ o Espanya recolzant la regulació en l'àmbit europeu amb *Chat Control Bill*²⁶. Això fa que ningú tingui accés

²⁴ <https://www.bloomberg.com/news/articles/2023-02-09/twitter-twtr-draws-eu-ire-for-sending-incomplete-disinformation-report>

²⁵ <https://www.theverge.com/2023/10/26/23922397/uk-online-safety-bill-law-passed-royal-assent-moderation-regulation>

²⁶ <https://www.techradar.com/news/spain-seeks-to-ban-encryption-leaked-document-reveals>

al contingut i calguin altres enfocaments per monitorar aquestes plataformes. Cal continuar buscant l'equilibri entre la seguretat i la privacitat de la ciutadania.

A partir de 2022 l'aparició amb força de noves eines d'intel·ligència artificial (IA) generativa ha facilitat la creació de contingut no veraç de tota mena: vídeos, imatges, texts, mems, clonació de webs, *deep fakes*... Un conjunt de contingut no veraç que per la seva qualitat és impossible de distingir de l'autèntic o generat per fonts oficials i que es pot generar de manera senzilla i amb pocs recursos. Eines que abans només estaven a l'abast de grans organitzacions ara són accessibles a qualsevol persona. El que provoca un augment del nombre de continguts falsos amb motivacions nocives per les persones. Existeix la tecnologia per crear contingut el qual sembla que una autoritat, per exemple el president del país, amb la seva imatge, digui el que els creadors vulguin. Quan ens trobem que el contingut generat per IA és indistingible de l'original trobem que el canal que difon la informació és l'únic mitjà de verificació directe de la informació.

Tot i aquestes amenaces som optimistes i pensem que tot i l'auge inicial fomentat per les novetats tecnològiques dels darrers anys: internet, plataformes socials digitals i IA generativa, els investigadors continuen fent una gran tasca en la definició i anàlisi del fenomen. En tractar-se d'un fenomen amb tantes implicacions cal continuar amb un enfocament multidimensional per combatre la desinformació. És més necessari que mai que les institucions públiques legislin sobre aquests nous entorns i que obliguin a facilitar l'accés a les dades i algorismes. Els investigadors externs han de ser els auditors externs d'aquests entorns. Alhora cal que aquests no treballin en departaments individuals, sinó que cal establir mecanismes per fomentar l'obertura de dades, la col·laboració entre governs, plataformes socials, organitzacions de fact-checking, mitjans de comunicació i acadèmia. En definitiva, només estem començant a veure la dimensió real del fenomen i els seus efectes negatius respecte a la seguretat i la salut de les persones, i cal que les diferents organitzacions públiques i les plataformes socials treballin per abordar el problema per cada àmbit territorial alhora que de forma conjunta també es treballi de manera global.

Bibliografia

- Almansa-Martínez, Ana, María Jesús Fernández-Torres, i Leticia Rodríguez-Fernández. 2022. «Desinformación en España un año después de la COVID-19. Análisis de las verificaciones de Newtral y Maldita». *Revista Latina de Comunicación Social* (80):183-200. doi: 10.4185/RLCS-2022-1538.
- Altay, Sacha, Manon Berriche, Hendrik Heuer, Johan Farkas, i Steven Rathje. 2023. «A Survey of Expert Views on Misinformation: Definitions, Determinants, Solutions, and Future of the Field». *Harvard Kennedy School Misinformation Review*. doi: 10.37016/mr-2020-119.
- Avaaz. 2020. *How Facebook can Flatten the Curve of the Coronavirus Infodemic*.
- Bakshy, E., S. Messing, i L. A. Adamic. 2015. «Exposure to Ideologically Diverse News and Opinion on Facebook». *Science* 348(6239):1130-32. doi: 10.1126/science.aaa1160.
- Bakshy, Eytan, Jake M. Hofman, Winter A. Mason, i Duncan J. Watts. 2011. «Everyone's an Influencer: Quantifying Influence on Twitter». P. 65 en *Proceedings of the fourth ACM international conference on Web search and data mining - WSDM '11*. Hong Kong, China: ACM Press.
- Bakshy, Eytan, Itamar Rosenn, Cameron Marlow, i Lada Adamic. 2012. «The Role of Social Networks in Information Diffusion». P. 519-28 en *Proceedings of the 21st international conference on World Wide Web - WWW '12*. Lyon, France: ACM Press.
- Ben Nimmo i Eric Hutchins. 2023. «Phase-based Tactical Analysis of Online Operations». *Carnegie Endowment for International Peace*.
- Ben Thompson. 2023. «Threads and the Social Communications Map». *Stratechery*. Recuperat 11 juliol 2023 (<https://stratechery.com/2023/threads-and-the-social-communications-map/>).
- Brennen, J. Scott, Felix M. Simon, i Rasmus Kleis Nielsen. 2021. «Beyond (Mis)Representation: Visuals in COVID-19 Misinformation». *The International Journal of Press/Politics* 26(1):277-99. doi: 10.1177/1940161220964780.
- Briand, Sylvie C., Matteo Cinelli, Tim Nguyen, Rosamund Lewis, Dimitri Prybylski, Carlo M. Valensise, Vittoria Colizza, Alberto Eugenio Tozzi, Nicola Perra, Andrea Baronchelli, Michele Tizzoni, Fabiana Zollo, Antonio Scala, Tina Purnat, Christine Czerniak, Adam J. Kucharski, Akhona Tshangela, Lei Zhou, i Walter Quattrocchi. 2021. «Infodemics: A New Challenge for Public Health». *Cell* 184(25):6010-14. doi: 10.1016/j.cell.2021.10.031.
- Broniatowski, David A., Amelia M. Jamison, SiHua Qi, Lulwah AlKulaib, Tao Chen, Adrian Benton, Sandra C. Quinn, i Mark Dredze. 2018. «Weaponized

- Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate». *American Journal of Public Health* 108(10):1378-84. doi: 10.2105/AJPH.2018.304567.
- Bruns, Axel. 2019. «After the ‘APIcalypse’: Social Media Platforms and Their Fight against Critical Scholarly Research». *Information, Communication & Society* 22(11):1544-66. doi: 10.1080/1369118X.2019.1637447.
- Buntain, Cody, Jennifer Golbeck, Brooke Liu, i Gary LaFree. 2016. «Evaluating Public Response to the Boston Marathon Bombing and Other Acts of Terrorism through Twitter». P. 4 en *Proceedings of the Tenth International AAAI Conference on Web and Social Media (ICWSM 2016)*.
- Calo, Ryan, Chris Coward, Emma S. Spiro, Kate Starbird, i Jevin D. West. 2021. «How Do You Solve a Problem like Misinformation?» *Science Advances* 7(50):eabn0481. doi: 10.1126/sciadv.abn0481.
- Castillo, Carlos. 2016. *Big Crisis Data: Social Media in Disasters and Time-Critical Situations*. Cambridge University Press.
- Cotter, Kelley, Julia R. DeCook, i Shaheen Kanthawala. 2022. «Fact-Checking the Crisis: COVID-19, Infodemics, and the Platformization of Truth». *Social Media + Society* 8(1):205630512110690. doi: 10.1177/20563051211069048.
- DataReportal. 2022. *Digital 2022 Global Digital Overview*.
- DiResta, Renee, Kris Shaffer, Becky Ruppel, David Sullivan, Robert Matney, Ryan Fox, Jonathan Albright, i Ben Johnson. 2018. «The Tactics & Tropes of the Internet Research Agency».
- Disinfo Lab. 2022. *Towards an Impact-Risk Index Of Disinformation: Measuring The Virality and Engagement of Single Hoaxes*.
- Eriksson, Mats. 2018. «Lessons for Crisis Communication on Social Media: A Systematic Review of What Research Tells the Practice». *INTERNATIONAL JOURNAL OF STRATEGIC COMMUNICATION* 27. doi: 10.1080/1553118x.2018.1510405.
- Espinet, Eva Ortoll, Montserrat Garcia-Alsina, i Agustí Canals. 2015. «Gestionar dades en situacions d’emergència. Estudi del cas dels focs de l’Alt Empordà a Catalunya». 8.
- European Commission. 2023. «Digital Services Act: Commission designates first set of Very Large Online Platforms and Search Engines». *Digital Services Act: Commission designates first set of Very Large Online Platforms and Search Engines*. Recuperat (<https://digital-strategy.ec.europa.eu/en/news/digital-services-act-commission-designates-first-set-very-large-online-platforms-and-search-engines>).
- Eysenbach, Gunther. 2002. «Infodemiology: The Epidemiology of (Mis)Information». *The American Journal of Medicine* 113(9):763-65. doi: 10.1016/S0002-9343(02)01473-0.

- Eysenbach, Gunther. 2020. «How to Fight an Infodemic: The Four Pillars of Infodemic Management». *Journal of Medical Internet Research* 22(6):e21820. doi: 10.2196/21820.
- Francois, Camille, i Herb Lin. 2021. «The Strategic Surprise of Russian Information Operations on Social Media in 2016 in the United States: Mapping a Blind Spot». *Journal of Cyber Policy* 6(1):9-30. doi: 10.1080/23738871.2021.1950196.
- Gobierno de España. 2022. *Lucha contra las campañas de desinformación en el ámbito de la seguridad nacional. Propuestas de la sociedad civil.*
- González-Bailón, Sandra. 2017. *Decoding the Social World. Data Science and the Unintended Consequences of Communication.* Massachusetts Institute of Technology.
- Gorman, Jack M., i David A. Scales. 2022. «Leveraging Infodemiologists to Counteract Online Misinformation: Experience with COVID-19 Vaccines». *Harvard Kennedy School Misinformation Review.* doi: 10.37016/mr-2020-92.
- Gruzd, Anatoliy, Philip Mai, i Felipe Bonow Soares. 2022. «How Coordinated Link Sharing Behavior and Partisans' Narrative Framing Fan the Spread of COVID-19 Misinformation and Conspiracy Theories». *Social Network Analysis and Mining* 12(1):118. doi: 10.1007/s13278-022-00948-y.
- Guess, Andrew M., Michael Lerner, Benjamin Lyons, Jacob M. Montgomery, Brendan Nyhan, Jason Reifler, i Neelanjan Sircar. 2020. «A Digital Media Literacy Intervention Increases Discernment between Mainstream and False News in the United States and India». *Proceedings of the National Academy of Sciences* 117(27):15536-45. doi: 10.1073/pnas.1920498117.
- Guess, Andrew M., Neil Malhotra, Jennifer Pan, Pablo Barberá, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio, Drew Dimmery, Deen Freelon, Matthew Gentzkow, Sandra González-Bailón, Edward Kennedy, Young Mie Kim, David Lazer, Devra Moehler, Brendan Nyhan, Carlos Velasco Rivera, Jaime Settle, Daniel Robert Thomas, Emily Thorson, Rebekah Tromble, Arjun Wilkins, Magdalena Wojcieszak, Beixian Xiong, Chad Kiewiet De Jonge, Annie Franco, Winter Mason, Natalie Jomini Stroud, i Joshua A. Tucker. 2023a. «How Do Social Media Feed Algorithms Affect Attitudes and Behavior in an Election Campaign?» *Science* 381(6656):398-404. doi: 10.1126/science.abp9364.
- Guess, Andrew M., Neil Malhotra, Jennifer Pan, Pablo Barberá, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio, Drew Dimmery, Deen Freelon, Matthew Gentzkow, Sandra González-Bailón, Edward Kennedy, Young Mie Kim, David Lazer, Devra Moehler, Brendan Nyhan, Carlos Velasco Rivera, Jaime Settle, Daniel Robert Thomas, Emily Thorson, Rebekah Tromble, Arjun Wilkins, Magdalena Wojcieszak, Beixian Xiong, Chad Kiewiet De Jonge, Annie Franco, Winter Mason, Natalie Jomini Stroud, i Joshua A. Tucker. 2023b. «Reshares on Social Media Amplify Political News but Do Not

- Detectably Affect Beliefs or Opinions». *Science* 381(6656):404-8. doi: 10.1126/science.add8424.
- Guess, Andrew, Jonathan Nagler, i Joshua Tucker. 2016. «Less than you think : Prevalence and predictors of fake news dissemination on Facebook». (January):1-9. doi: 10.1126/sciadv.aau4586.
- Hagen, Loni, Ashley Fox, Heather O’Leary, DeAndre Dyson, Kimberly Walker, Cecile A. Lengacher, i Raquel Hernandez. 2022. «The Role of Influential Actors in Fostering the Polarized COVID-19 Vaccine Discourse on Twitter: Mixed Methods of Machine Learning and Inductive Coding». *JMIR Infodemiology* 2(1):e34231. doi: 10.2196/34231.
- Hagen, Loni, Thomas Keller, Stephen Neely, Nic DePaula, i Claudia Robert-Cooperman. 2018. «Crisis Communications in the Age of Social Media: A Network Analysis of Zika-Related Tweets». *Social Science Computer Review* 36(5):523-41. doi: 10.1177/0894439317721985.
- Hagen, Loni, Ryan Scharf, Stephen Neely, i Thomas Keller. 2018. «Government Social Media Communications during Zika Health Crisis». 10. doi: 10.1145/3209281.3209364.
- Houston, J. Brian, Joshua Hawthorne, Mildred F. Perreault, Eun Hae Park, Marlo Goldstein Hode, Michael R. Halliwell, Sarah E. Turner McGowen, Rachel Davis, Shivani Vaid, Jonathan A. McElderry, i Stanford A. Griffith. 2015. «Social Media and Disasters: A Functional Framework for Social Media Use in Disaster Planning, Response, and Research». *Disasters* 39(1):1-22. doi: 10.1111/disa.12092.
- Innes, Martin. 2018. *From Minutes to Months. A Rapid Evidence Assessment of the Impact of Media and Social Media during and after Terror Events*. Crime & Security Research Institute.
- Knuutila, Aleks, Lisa-Maria Neudert, i Philip N. Howard. 2022. «Who Is Afraid of Fake News? Modeling Risk Perceptions of Misinformation in 142 Countries». *Harvard Kennedy School Misinformation Review*. doi: 10.37016/mr-2020-97.
- López-García, Xosé, Carmen Costa-Sánchez, i Ángel Vizoso. 2021. «Journalistic Fact-Checking of Information in Pandemic: Stakeholders, Hoaxes, and Strategies to Fight Disinformation during the COVID-19 Crisis in Spain». *International Journal of Environmental Research and Public Health* 18(3):1227. doi: 10.3390/ijerph18031227.
- Malhotra, Pranav. 2020. «A Relationship-Centered and Culturally Informed Approach to Studying Misinformation on COVID-19». *Social Media + Society* 6(3):205630512094822. doi: 10.1177/2056305120948224.
- Marx, Julian, Milad Mirbabaie, i Christian Ehnis. 2018. «Sense-Giving Strategies of Media Organisations in Social Media Disaster Communication : Findings from Hurricane Harvey». *Australasian Conference on Information Systems* 1-12.

- Newman, Nic, Richard Fletcher, Kirsten Eddy, Craig T. Robertson, i Rasmus Kleis Nielsen. 2023. «Reuters Institute Digital News Report 2023».
- Nguyễn, Sarah, Rachel Kuo, Madhavi Reddi, Lan Li, i Rachel E. Moran. 2022. «Introduction: Mis- and Disinformation as Global». *Harvard Kennedy School Misinformation Review* 12. doi: 10.37016/mr-2020-95.
- Nielsen, Rasmus Kleis, Richard Fletcher, Nic Newman, J. Scott Brennan, i Philip N. Howard. 2020. «Navigating the ‘Infodemic’: How People in Six Countries Access and Rate News and Information about Coronavirus». 36.
- Nimmo, Ben. 2020. *The Breakout Scale: Measuring the Impact of Influence Operations*.
- North Atlantic Treaty Organization. 2022. *Strategic Concept 2022*.
- Nyhan, Brendan, Jaime Settle, Emily Thorson, Magdalena Wojcieszak, Pablo Barberá, Annie Y. Chen, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio, Drew Dimmery, Deen Freelon, Matthew Gentzkow, Sandra González-Bailón, Andrew M. Guess, Edward Kennedy, Young Mie Kim, David Lazer, Neil Malhotra, Devra Moehler, Jennifer Pan, Daniel Robert Thomas, Rebekah Tromble, Carlos Velasco Rivera, Arjun Wilkins, Beixian Xiong, Chad Kiewiet De Jonge, Annie Franco, Winter Mason, Natalie Jomini Stroud, i Joshua A. Tucker. 2023. «Like-Minded Sources on Facebook Are Prevalent but Not Polarizing». *Nature* 620(7972):137-44. doi: 10.1038/s41586-023-06297-w.
- Palen, Leysia, Jennings Anderson, Melissa Bica, Carlos Castillos, John Crowley, Paloma Díaz, Megan Finn, Rob Grace, Amanda Hughes, Muhammad Imran, Marina Kogan, Nicolas Lalone, Prasenjit Mitra, Wendy Norris, Katie Pine, Hemant Purohit, Christian Reuter, Caroline Rizza, Lise St Denis, Bryan Semaan, Valerie Shalin, Lea Shanley, Patrick Shih, Robert Soden, Kate Starbird, Keri Stephen, Z. Toups, i Tom Wilson. 2020. «Crisis Informatics: Human-Centered Research on Tech & Crises». 15.
- Palen, Leysia, i Kenneth M. Anderson. 2016. «Crisis Informatics—New Data for Extraordinary Times». *Science* 353(6296):224-25. doi: 10.1126/science.aag2579.
- Pennycook, Gordon, Tyrone D. Cannon, i David G. Rand. 2018. «Prior Exposure Increases Perceived Accuracy of Fake News». *Journal of Experimental Psychology* 147(12):1865-80. doi: 10.1037/xge0000465.
- Pennycook, Gordon, Jonathon McPhetres, Yunhao Zhang, Jackson G. Lu, i David G. Rand. 2020. «Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention». *Psychological Science* 31(7):11. doi: https://doi.org/10.1177/0956797620939054.
- Recuero, Raquel, i Felipe Soares. 2022. «#VACHINA: How Politicians Help to Spread Disinformation About COVID-19 Vaccines». *Journal of Digital Social Research* 73-97. doi: 10.33621/jdsr.v4i1.112.

- Reuter, Christian, Amanda Lee Hughes, i Marc-André Kaufhold. 2018. «Social Media in Crisis Management: An Evaluation and Analysis of Crisis Informatics Research». *International Journal of Human-Computer Interaction* 34(4):280-94. doi: 10.1080/10447318.2018.1427832.
- Rosenberg, Matthew, Nicholas Confessore, i Carole Cadwalladr. 2018. «How Trump Consultants Exploited the Facebook Data of Millions». *The New York Times*, març 17.
- Roth, Yoel. 2020. «Bot or Not? The Facts about Platform Manipulation on Twitter». *Twitter Bog*. Recuperat 13 febrer 2024 (https://blog.twitter.com/en_us/topics/company/2020/bot-or-not).
- Roth, Yoel. 2023. «Opinion | Trump Attacked Me. Then Musk Did. It Wasn't an Accident.» *The New York Times*, setembre 18.
- Royan, Regina, Tricia Rae Pendergrast, Marina Del Rios, Shannon M. Rotolo, N. Seth Trueger, Eve Bloomgarden, Deanna Behrens, Shikha Jain, i Vineet M. Arora. 2022. «Use of Twitter Amplifiers by Medical Professionals to Combat Misinformation During the COVID-19 Pandemic». *Journal of Medical Internet Research* 24(7):e38324. doi: 10.2196/38324.
- Salaverria, Ramón, Nataly Buslón, Fernando López-Pan, Bienvenido León, Ignacio López-Goñi, i María-Carmen Erviti. 2020. «Desinformación en tiempos de pandemia: tipología de los bulos sobre la Covid-19». *El Profesional de la Información* 29(3). doi: 10.3145/epi.2020.may.15.
- Shane, Tommy. 2020. «The Psychology of Misinformation: Why We're Vulnerable». *First Draft*. Recuperat 13 febrer 2024 (<https://firstdraftnews.org/articles/the-psychology-of-misinformation-why-were-vulnerable/>).
- Starbird, Kate. 2017. «Examining the Alternative Media Ecosystem through the Production of Alternative Narratives of Mass Shooting Events on Twitter». P. 10 en.
- Starbird, Kate, i Tom Wilson. 2020. «Cross-Platform Disinformation Campaigns: Lessons Learned and Next Steps». *Harvard Kennedy School Misinformation Review*. doi: 10.37016/mr-2020-002.
- Theocharis, Yannis, Ana Cardenal, Soyeon Jin, Toril Aalberg, David Nicolas Hopmann, Jesper Strömbäck, Laia Castro, Frank Esser, Peter Van Aelst, Claes de Vreese, Nicoleta Corbu, Karolina Koc-Michalska, Joerg Matthes, Christian Schemer, Tamir Sheafer, Sergio Splendore, James Stanyer, Agnieszka Stepińska, i Václav Štětka. 2021. «Does the Platform Matter? Social Media and COVID-19 Conspiracy Theory Beliefs in 17 Countries». *New Media & Society* 146144482110456. doi: 10.1177/14614448211045666.
- Tucker, Joshua, Andrew Guess, Pablo Barbera, Cristian Vaccari, Alexandra Siegel, Sergey Sanovich, Denis Stukal, i Brendan Nyhan. 2018. «Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature». *SSRN Electronic Journal*. doi: 10.2139/ssrn.3144139.

- Unión de Televisiones Comerciales en Abierto. 2022. *1r Estudio desinformación en España*.
- Uscinski, Joseph E., Adam M. Enders, Casey Klofstad, Michelle Seelig, John Funchion, Caleb Everett, Stephan Wuchty, Kamal Premaratne, i Manohar Murthi. 2020. «Why Do People Believe COVID-19 Conspiracy Theories?» *Harvard Kennedy School Misinformation Review*. doi: 10.37016/mr-2020-015.
- Valencia, Stephanie. 2021. «Misinformation Online Is Bad in English. But It's Far Worse in Spanish.» *Washington Post*, octubre 28.
- Valero, Myriam Vidal. 2023. «Thousands of Scientists Are Cutting Back on Twitter». *Nautre*.
- Vosoughi, Soroush, Deb Roy, i Sinan Aral. 2018. «The Spread of True and False News Online». *Science* 359(6380):1146-51. doi: 10.1126/science.aap9559.
- Vraga, Emily K., i Leticia Bode. 2017. «Using Expert Sources to Correct Health Misinformation in Social Media». *Science Communication* 39(5):621-45. doi: 10.1177/1075547017731776.
- Vraga, Emily K., i Leticia Bode. 2021. «Addressing COVID-19 Misinformation on Social Media Preemptively and Responsively». *Emerging Infectious Diseases* 27(2):396-403. doi: 10.3201/eid2702.203139.
- Vraga, Emily K., Ullrich K. H. Ecker, Iris Žeželj, Aleksandra Lazić, i Arina A. Azlan. 2023. «To Debunk or Not to Debunk? Correcting (Mis)Information». P. 85-98 en *Managing Infodemics in the 21st Century*, editat per T. D. Purnat, T. Nguyen, i S. Briand. Cham: Springer International Publishing.
- Wardle, Claire, i Hossein Derakhsan. 2017. *Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making*. 162317GBR. Council of Europe.
- Watts, Duncan J., David M. Rothschild, i Markus Mobius. 2021. «Measuring the News and Its Impact on Democracy». *Proceedings of the National Academy of Sciences* 118(15):e1912443118. doi: 10.1073/pnas.1912443118.
- World Health Organization. 2021. *WHO Competency Framework: Building a Response Workforce to Manage Infodemics*. Geneva: World Health Organization.
- World Health Organization. 2022. *WHO policy brief: COVID-19 infodemic management*.
- Xaudiera, Sergi. 2018. «Crisis Management through Twitter: The Case of the Barcelona Attacks». 2017(August 2017):1-12.
- Xaudiera, Sergi, i Ana S. Cardenal. 2020. «Ibuprofen Narratives in Five European Countries During the COVID-19 Pandemic». *Harvard Kennedy School Misinformation Review*. doi: 10.37016/mr-2020-029.

Annexos

Annex A

Authoritative Agency

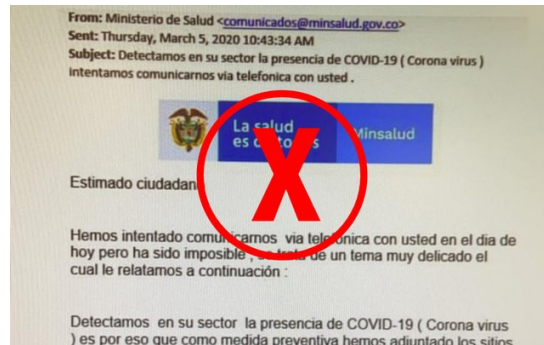


Figure 5. Email simulating an official communication from the Colombian Government



Figure 6. False statement attributed to the Minister of Health of Ecuador.

Virulence



Figure 7. Photo in which the author tries to convince us that these are people falling from the coronavirus in China.



Figure 8. Image allegedly showing a doctor in a Madrid hospital who died of coronavirus.

Medical Efficacy

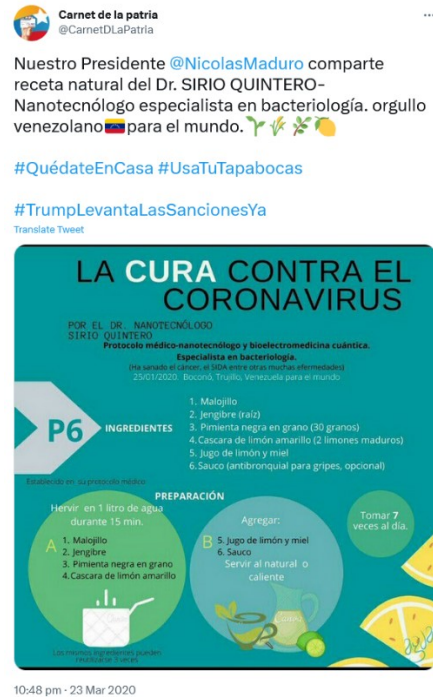


Figure 9. Infographic supposedly showing how to cure the coronavirus.

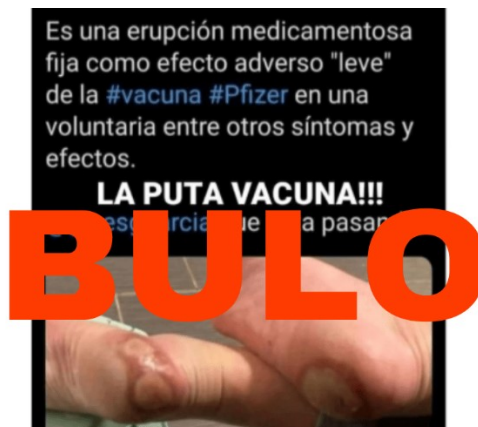


Figure 10. Alleged adverse reactions to Pfizer's vaccine in a female volunteer.

Satire



Figure 11. Bottle of wine sent by drone to quarantined cruise passengers.



Figure 12. Image showing a lion in Russia to keep people from leaving their homes.

Prophecy

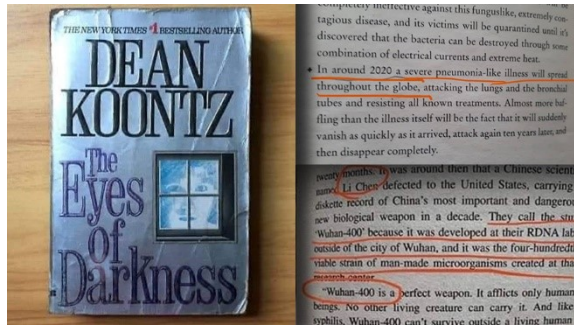


Figure 13. Alleged prophecy in a book by Dean Koontz.



Figure 14. Alleged prophecy on a DVD movie.

Intolerance



Figure 15. Image showing Asian people supposedly eating a baby. The author criticises them for allegedly being linked to the coronavirus.



Figure 16. Image showing a group of young people of Arab origin who have allegedly taken advantage of the pandemic confusion to steal from shops.



Figure 17. Alleged nurses criticising government party voters who applauded them at 8pm on the balcony (a practice of thanking the medical staff).



Figure 18. False information about an alleged economic measure of the Spanish Government.