

European Digital Media Observatory

IV.D.A: Academic research on disinformation at scale in the EU

Date: 31-05-2021

Final version

Project number:	SMART 2019/1087
Project Acronym:	EDMO
Project title:	European Digital Media Observatory
Start date of the project:	01/06/2020
Duration of the project:	30
Project website address:	https://edmo.eu/
Authors of the deliverable	Petra de Place Bak, DATALAB, Aarhus University Mathias Holm Sørensen, DATALAB, Aarhus University Jessica Gabriele Walter, DATALAB, Aarhus University Anja Bechmann, DATALAB, Aarhus University



European Digital Media Observatory

Executive summary:

In this report, we present the progress on the initial part of Task IV that includes four main elements: the mapping of relevant academic activities in the EU, the identification of relevant academic institutions and organizations in the EU, the creation of a repository of relevant scientific articles, and the creation of a repository of relevant policy papers and other content. Task IV.D.A consists of mapping academic activities studying disinformation at scale in the EU and has as per 15.05.2021 been completed. The task is based on an extensive literature search that will also form the basis of the preliminary repository (IV.MS.1). The literature search and following manual filtering for relevance resulted in the inclusion of 117 entries of different fields of research studying mis-/disinformation at scale in the EU. In this report, we make the choices made prior to the search for academic research papers transparent, present the filtering process, the annotation scheme and highlight the most important characteristics of the repository so far. Moving forward, IV.D.A. serves as the foundation for the successful completion of all other subtasks in Task IV.



European Digital Media Observatory

List of figures

Figure 1 Base of researchers based on institutional affiliation within the European Union

Figure 2 Base of researches based on institutional affiliation outside the European Union

Figure 3 Regional interests of relevant studies within the European Union

Figure 4 Regional interest of relevant studies outside the European Union

Figure 5 Fields of research of relevant articles

Table of content

List of figures	4
Table of content	5
1.0 Introduction	6
2.0 Method	6
2.1 Search engine	7
2.2 Keywords	7
2.3 Timeframe	8
2.4 Language & Geography	9
2.5 Coding	9
2.6 Filtering process	9
2.7 Risks & Mitigation measures	10
3.0 Results	10
3.1 Regional affiliation of researchers	11
3.2 Regional interest of listed studies	12
3.3 Fields of research	13
4.0 Actions needed	14
5.0 Conclusion	15
6.0 References	16



European Digital Media Observatory

1.0 Introduction

The long-term aim of Task IV is to provide support and coordination for academic research activities on disinformation in the European Union. This entails mapping relevant academic research capabilities in the European Union and the creation of a repository with relevant peer-reviewed scientific literature on disinformation. The research team at DATALAB – Center for Digital Social Research at Aarhus University, Denmark, is responsible for Task IV which includes four main elements: IV.D.A: The mapping of relevant activities in the EU, IV.D.B: The identification of relevant academic institutions and organizations in the EU, IV.D.C: The creation of a repository of relevant scientific articles and IV.D.D: The creation of a repository of relevant policy papers and other content.

The purpose of the first action and this deliverable is to create an overview of European academic activities that study disinformation at scale from the viewpoint of different academic disciplines including Computer Science and Information Studies, Behavioral Science, Social Sciences, Media Law and Economics, Communication and Media Studies, Neuroscience and Psychology, Health Care Studies, and Other. Based on this initial task, the researchers at the DATALAB will reach out to other independent university-based researchers and organizations. In the process, the preliminary repository of academic literature on disinformation will be extended based on the input from the identified partners and by periodic updates of the search to include the newest academic publications. This task constitutes an important step toward supporting the coordination, collaboration and information exchange between research communities within the EU.

In this report, we outline the methods used for conducting the literature search and highlight selected preliminary results based on the manual filtering and coding of all academic work included in the initial search.

2.0 Method

In this section, we first describe the choices made prior to the literature search, i.e. choice of search engine, definition of keywords, timeframe, language, and relevance criteria. Second, we reflect on the following filtering process of all entries in the initial

results from the literature search. By providing these descriptions, we ensure transparency as to how the repository was initiated.

2.1 Search engine

Both Google Scholar and the Danish Royal Library were considered as access points for the literature search. Google Scholar was rejected as an option due to having a character limit of 230 in the search bar, which did not allow for the inclusion of all the selected and relevant keywords. Hence, we chose the Danish Royal Library, which gives access to search in 10,113 collections including for example Scopus and Web of Science (an overview of the collections will be available in the online repository), and therefore provides an extensive access to all relevant literature.

2.2 Keywords

The literature search was based on the careful selection of different keywords. The process of selecting keywords is crucial to obtaining relevant search results and limits the risks of missing out on relevant academic research articles. To best meet the objective of the literature search, the keywords were divided into two categories: category 1 contains words related to disinformation, while category 2 contains country names of current or former members of the EU. Due to the complexity of the topic and in order to find all relevant research articles, we assigned different relevance to these two categories. At least one of the words from the first category had to be included in the title of the publications, whereas it was sufficient if at least one of the words from the second category was included anywhere in the text additionally to the word from the first category.

There are several terms in use for different types or aspects of false information and several studies contribute to the conceptualization of “disinformation”, “misinformation”, “fake news” and related terms (e.g. Bechmann, Anja & O’Loughlin, Ben, 2020; Buning, 2018; Farkas & Schou, 2019; Kalsnes, 2018; Tandoc et al., 2018; Wardle, Claire & Derakhshan, Hossein, 2017). The concepts address differences, for example, regarding intentions behind the fabrication and spreading - with e.g.

disinformation being intentional and misinformation not necessarily so. The literature search is designed with the aim to capture most terms used for false information, as intentions for example are not the focus in all relevant publications and in addition, some terms sometimes are used interchangeably. Thus, category 1 comprises several keywords used related to disinformation, namely: disinformation, misinformation, “fake news”, malinformation, “information disorder”, “false information”, hoax and/or “conspiracy theory”.¹

In order to detect relevant studies and research activities within the European Union, category 2 contains all (former and current) member states and relevant abbreviations: Europe, EU, Germany, Denmark, Sweden, Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Estonia, Finland, France, Greece, Hungary, Italy, Latvia, Luxembourg, Malta, Poland, Ireland, Portugal, Romania, Slovakia, Slovenia, Spain, Lithuania, Netherlands, Scotland, England, Wales, Great Britain, GB and/or Britain.

2.3 Timeframe

We set the timeframe to 2015 onward and will update the search periodically to include the newest academic publications. If necessary, also keywords will be adjusted or extended for the updates. 2015 was chosen as the starting point as it marks the year that the manipulation of information during the Ukraine crisis led the European Council to call for an action plan (Bentzen & Russell, 2015); the plan was published later the same year by the newly formed East StratCom Task Force. Moreover, this timeframe also includes research on the influence of digital mis- and disinformation in relation to the Brexit Campaign and the Donald Trump 2016 Presidential Campaign. Finally, 2015 was the year where the Poynter Institute established The International Fact-Checking Network (IFCN), which is the first international network to bring together fact-checkers worldwide and therefore marks an important event in the fight against mis- and disinformation.

¹Based on Buning et al. (2018) we use the term disinformation and misinformation more generally as “false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit”. The terms misinformation, disinformation and related terms are interchangeably hereupon.

A choice was made to exclusively include academic work published in English for several reasons. On the one hand English is the official language of the EDMO project. On the other hand, articles published in English can be read by a large audience, making the repository more relevant. Furthermore, the purpose of the task is to create an overview of European academic activities that study disinformation at scale from the viewpoint of different academic fields. The priority is therefore not to provide a comprehensive overview of research findings, but rather to identify where research is done and within which academic fields. The search was further narrowed down to research focusing on digital mis- and disinformation within the European Union (EU) and its member states. For that reason, the researchers may be based outside of the EU.

2.5 Coding

Based on the above described search criteria a search was conducted on the 28th of January 2021 which resulted in 2.021 articles, proceedings, books, book chapters and reports were found. These have been filtered for relevance, defined as the relevance to research in digital media in relation to mis- and disinformation at scale. The relevant set from the library was then further filtered based on title and abstract by the DATALAB for 1) Field of research, 2) Affiliation of researcher(s), 3) Base and 4) Geographic area of interest.

2.6 Filtering process

All search results were manually filtered to secure the inclusion of studies that fulfill the relevance criteria as previously described. The main reasons for exclusion of papers and publications were:

1. Papers with none of the keywords in the title. These appeared in the initial search due to one or more keywords present in the abstract.
2. Papers in other languages than English, which appeared in the search results because they include an English version of the abstract.
3. Non-empirical research papers such as theoretical conceptualizations, editorials, commentaries and publications in non-scientific journals.

4. Academic papers that address disinformation not at scale, e.g. smaller lab-based studies or non-representative surveys.
5. Papers that did not focus on disinformation in digital media.
6. Papers by researchers who are affiliated to universities and institutions outside the EU and the research itself did not focus on disinformation within the EU. In the last case, we expect the entries were included in the search results because of the keyword 'EU' which may have matched words such as 'Reuters' and 'Neutral'.
7. Finally, policy papers and media literacy related papers were filtered out as they better fit in the planned repository of relevant policy papers and other content that is part of a later stage of task IV.

2.7 Risks & Mitigation measures

The risk of missing out on relevant literature is inevitable due to the sheer volume of scientific literature on the subject. At DATALAB we have mitigated this risk by carefully defining relevant search terms related to disinformation and names of member states. We have given priority to creating an overview of European academic activities that study disinformation at scale from the viewpoint of different academic fields. As the repository is updated, research published in other languages can be added in collaboration with national/multinational hubs established as part of the second phase of EDMO to create a comprehensive overview of academic research.

3.0 Results

In this section, we describe the preliminary results of the literature search, i.e. member states represented by affiliation, fields of research and member states as areas of interest. After the filtering, 119 results remain, which will be supplemented as the search is periodically updated and supplemented based on inputs from the identified organizations. The final repository aims to list at least 200 entries.

3.1 Regional affiliation of researchers

So far, the literature search has resulted in a broad representation of researchers based both in- and outside the EU. A total of 18 member states are represented, shown here followed by the number of entries in parenthesis: Italy (26), Spain (14), Germany (14), Netherlands (13), Greece (7), Poland (6), Portugal (5), Slovakia (4), Austria (4), France (4), Belgium (3), Romania (3), Sweden (2), Denmark (2), Cyprus (2), Finland (2), Czech Republic (2) and Bulgaria (1). Outside the EU research is included from: UK (24), USA (9), Switzerland (6), Canada (4), Brazil (2), Australia (2), Norway (1), Russia (1), Qatar (1), Turkey (1) and Israel (1) - see also figure 1 and 2. The latter have been included either due to collaborations with researchers affiliated to universities within the EU or because of an explicit research focus on one or more member states. Countries were identified based on affiliations of all authors announced in the publications and in cases where authors were based in different countries, all countries were included in the count. Only in one case was the author affiliation unidentifiable due to the author's independence of any university or organization.

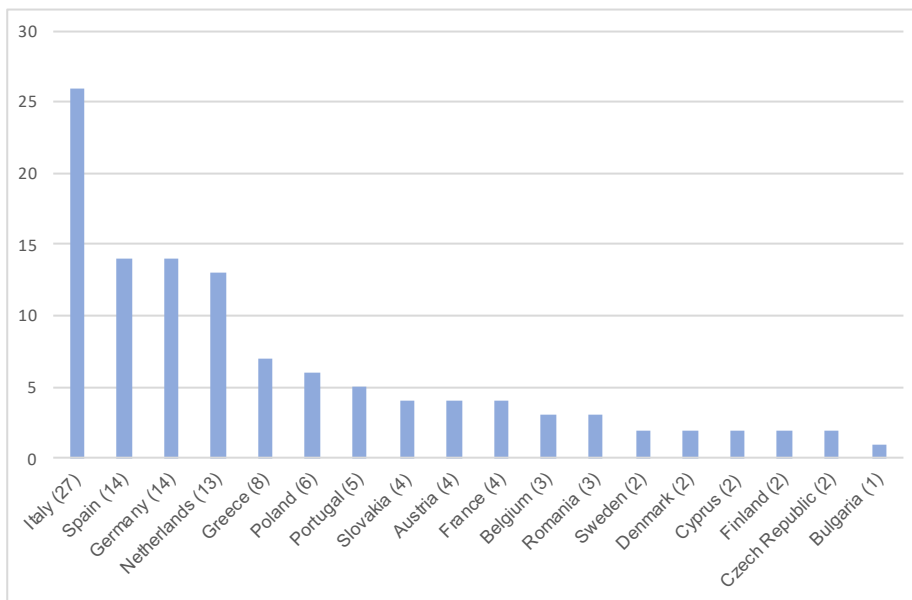


Figure 1 Base of researchers based on institutional affiliation within the European Union

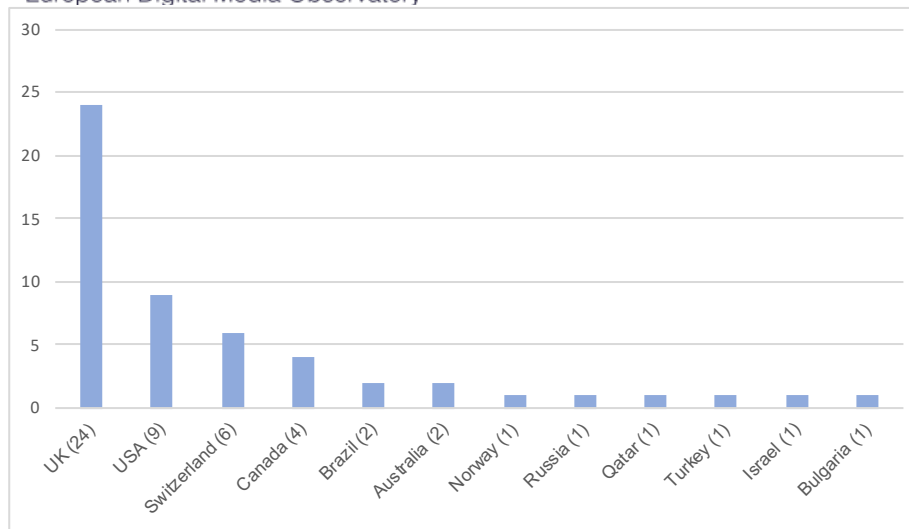


Figure 2 Base of researches based on institutional affiliation outside the European Union

3.2 Regional interest of listed studies

The entries in the preliminary repository were annotated with the geographic area of interest of the study. In 76 cases the study is manually coded as ‘nonspecific’, as many studies use large datasets scraped from social media platforms (e.g. Twitter), which are not country specific. For the remaining articles, the geographic area of interest is identified based on countries mentioned in the titles, abstracts and in some cases the papers themselves. In cases of cross-country studies, all countries are taken into consideration. The member states included followed by the number of studies in parenthesis is listed here (see figure 3): Italy (14), Germany (9), Austria (5), Finland (5), Spain (6), Bulgaria (4), Poland (4), Portugal (4), Netherlands (4), Sweden (4), France (4), Belgium (3), Denmark (3), Greece (3), Romania (3), Ireland (3), Hungary (2), Estonia (2), Cyprus (2), Croatia (2), Czech Republic (2), Latvia (2), Lithuania (2), Luxembourg (2), Malta (2), Slovakia (2) and Slovenia (2). Outside of the EU the countries investigated by authors with an affiliation within the EU are (see figure 4): USA (15), UK (11), China (3), Canada (2), Brazil (2), Turkey (2), Norway (1), Switzerland (1), Israel (1) and Taiwan (1). Note that these are included either because the study is comparative (i.e. compares one or more member states to an area outside the EU) or the study is conducted by researcher(s) affiliated with a

university based in the EU. Out of the remaining 117 entries, 14 entries study more than one country.

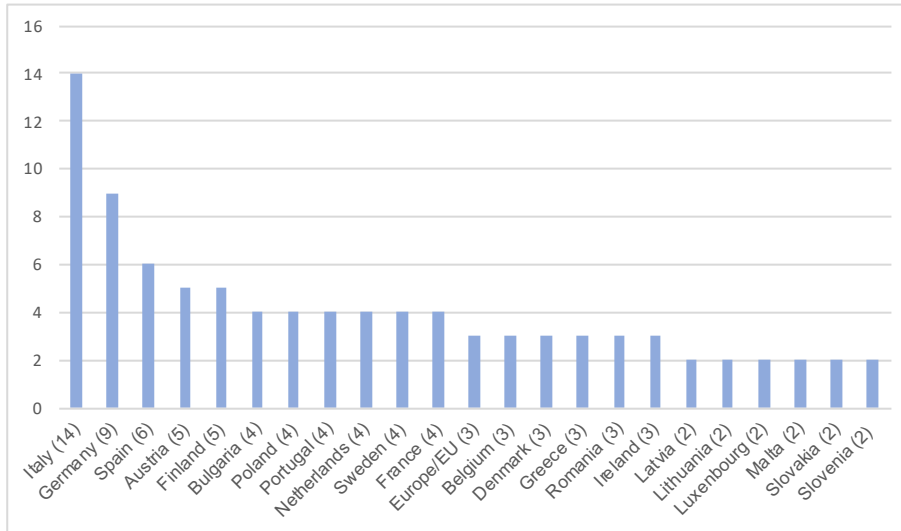


Figure 3 Regional interests of relevant studies within the European Union

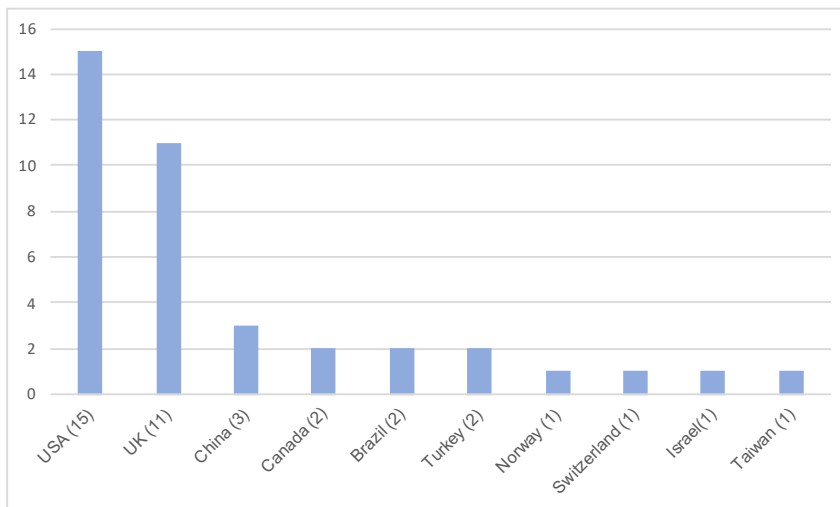


Figure 4 Regional interest of relevant studies outside the European Union

3.3 Fields of research

Based on the included articles we defined the following fields of research categories: Computer Science and Information Studies (73), Behavioral Science (2), Social Sciences (17), Media Law and Economics (4), Communication and Media Studies (18), Neuroscience and Psychology (4), Health Care Studies (8), and Other (3). These differ partly from the fields of research originally listed in the grant agreement (i.e. social and political science (more broadly addressed as “Social Sciences”);

network science (included in “Computer Science and Information studies”); neuroscience and psychology; media law and economics; media literacy (included in “Communication and Media Studies”); behavioral sciences) as these new categories better allow for related disciplines to constitute one category and hereby, increase searchability in the final repository. The fields of research were identified based on the papers’ titles, abstracts and when included the topic descriptions (for some papers this field was left empty). So far, the category ‘Computer Science and Information Studies’ is by far the most represented category, which can be ascribed to the possibilities of gathering large datasets using the methods of these fields of research – see figure 5.

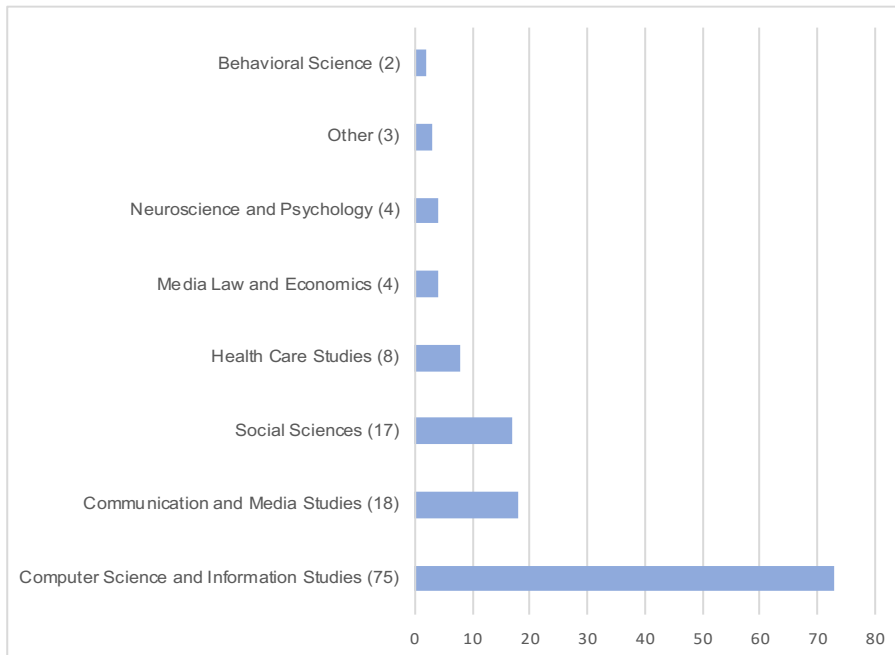


Figure 5 Fields of research of relevant articles

4.0 Actions needed

The literature search forms the foundation for both IV.D.A and the other activities in Task IV (IV.D.B, IV.D.C, IV.D.D). Moving forward, the next steps are to identify and engage with relevant academic institutions and organizations, to build and make accessible repositories of both relevant scientific articles and relevant policy papers.



European Digital Media Observatory

5.0 Conclusion

At this stage, the literature search shows a good diversity in representation of member states both by regional affiliation of researchers and by regional interest of listed studies. The entries in the preliminary repository suggest that Italy is the most active member state when it comes to studying disinformation at scale and perhaps for this reason, the most frequently studied country within the EU. Concerning fields of research, the repository represents a broad range, however with a focus on studies from Computer Science and Information Studies.

6.0 References

- Bechmann, Anja & O'Loughlin, Ben. (2020). *Democracy & disinformation: A turn in the debate* (KVAB Thinkers' Report, p. 37). Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten.
<https://www.kvab.be/sites/default/rest/blobs/2557/Final%20Report%20Dem%20&%20Desinfo.pdf>
- Bentzen, N., & Russell, M. (2015). *Russia's disinformation on Ukraine and the EU's response* [Briefing]. EPRS | European Parliamentary Research Service.
[https://www.europarl.europa.eu/RegData/etudes/BRIE/2015/571339/EPRS_BRI\(2015\)571339_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2015/571339/EPRS_BRI(2015)571339_EN.pdf)
- Buning, M. et al. (2018). *A Multi-dimensional Approach to Disinformation: Report of the independent high level group on fake news and online disinformation*. EU Commission.
- Farkas, J., & Schou, J. (2019). *Post-truth, fake news and democracy: Mapping the politics of falsehood*. Routledge.
- Kalsnes, B. (2018). Deciding what's true: The rise of political fact-checking in American journalism. *Digital Journalism*, 6(5), 670–672.
<https://doi.org/10.1080/21670811.2018.1460211>
- Tandoc, E. C., Lim, Z. W., & Ling, R. (2018). Defining “Fake News”: A typology of scholarly definitions. *Digital Journalism*, 6(2), 137–153.
<https://doi.org/10.1080/21670811.2017.1360143>
- Wardle, Claire & Derakhshan, Hossein. (2017). *INFORMATION DISORDER: Toward an interdisciplinary framework for research and policy making* (Council of Europe Report DGI, p. 109). Council of Europe. <https://rm.coe.int/information-disordertoward-an-interdisciplinary-framework-for-research/168076277c>