

D.10 Periodically updated description of the DSI efforts to support and facilitate the coordination of academic research

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 months
Period covered by the interim report:	from 06/11/2020 to 05/05/2021
Date of submission of the periodic description:	31/05/2021
Project website address:	https://edmo.eu/
The deliverable has been elaborated by:	Mathias Holm Sørensen, DATALAB, Aarhus University Petra de Place Bak, DATALAB, Aarhus University Jessica Walter, DATALAB, Aarhus University Anja Bechmann, DATALAB, Aarhus University

Table of Contents

<i>1.0 Introduction</i>	4
<i>2.0 Engagement with researchers</i>	4
<i>3.0 Relevant academic research</i>	5
3.1 Detection.....	5
3.2 COVID-19	6
3.4 Misinformation and human behavior.....	8
<i>4.0 Assessment of the scientific repository</i>	8
4.1 Quantitative evaluation	9
4.2 Qualitative evaluation	10
<i>5.0 References</i>	11

1.0 Introduction

This report is the second periodically updated description of the DSI efforts to support and coordinate academic research in Europe (D.10) related to Task IV in EDMO. This is the second of five deliverables that will be made during the duration of the project – one every six months. The reports will provide an overview of the current and recent achievements in order to support and coordinate academic research in Europe within the field of digital media and information disorder. They will also provide an overview of the most relevant academic research results in the field of disinformation and, when relevant, a quantitative and qualitative assessment of the scientific repository. These reports will naturally expand as the project goes on, as academic activities studying disinformation in the EU (IV.D.A) and relevant academic institutions and organizations (IV.MS.2) are mapped, listed and reached out to, and as repositories of relevant scientific articles (IV.MS.1 & IV.D.C) and of relevant policy papers and other content (IV.D.D) are established. These reports will start with a presentation of the support and coordination of academic research in the European Union followed by an overview of key research topics and findings within the field of information disorder, and finally conclude with an evaluation of the research repository. This report covers M7-M12 of the project-period.

2.0 Engagement with researchers

Engagement and support of researchers studying disinformation in the European Union will be an ongoing activity in task IV. The first two steps in doing this are the establishing of the preliminary repository of academic literature studying disinformation at scale in the European Union and the integration of researchers in EDMO.

As it will be described in section 4.0 of this report on the assessment of the scientific repository, the literature search that will act as the foundation for the repository has been finalized. This initial search resulted in 2.021 results, which was filtered thoroughly for relevance, which resulted in the inclusion of 117 pieces of literature to be included in the repository. The process of setting up and providing access to the repository from the EDMO website has been initiated in collaboration with Athens Technology Center.

Furthermore, this literature search will act as the foundation for task *IV.D.B List of relevant academic institutions and organizations*, where the plan is 1. to identify institutions within IV.D.A., 2. to expand the list using the snowballing method, and 3. to reach out to all organizations on the list. This process has thereby been initiated through the literature search and following filtering-process.

As the final ongoing action to engage with and support academic research studying disinformation in the European Union, Aarhus University will assist EDMO in constructing the requirements for researchers to be members of EDMO in 2021. Later in the project researchers will be able to become an integrated part of EDMO and the resources, activities and data made available by the observatory. In this process a registration document has been drafted by Aarhus University, which has been through a round of internal reviews within the EDMO executive board. Following this, Aarhus University will make a second draft, taking the comments and suggestions into account.

3.0 Relevant academic research

The following description is not an exhaustive list of the findings within the field of disinformation, but rather a look into some relevant results as identified by the EDMO-affiliated researchers at Aarhus University. The section is based on a search performed using the keywords described in *IV.D.A Academic research on disinformation at scale in the EU* within the period of M7-M12 of the project. In this periodic report, we have chosen three relevant topics in the field of research, which are described next in more detail. The three topics are 1) detection, 2) COVID-19 and 3) psychological effects of disinformation on humans.

3.1 Detection

As with the first periodically updated description of new research within the field of information disorder, methods for detecting false information is a prominent theme within research in this reporting period.

One way to identify false information is through deep-learning techniques. In a lot of research, deep-learning models have been used on the body of text of news articles, which is typically what is included in publicly available datasets. However, new experimental research presents the possibility of including visual features in the detection of fake news. This is done using a hybrid multimodal framework that evaluates data based both on textual and visual features (images). This framework provides a method for higher accuracy in the detection of false information. However, the lack of images in the scraped publicly available fake news datasets poses both an academic and practical challenge (Meel & Vishwakarma, 2021).

One of the other big challenges that have been identified in the current development of accurate methods for detecting misinformation is the lack of understanding of language alterations (Zhou, et. al., 2021). Especially for speakers with English as their second language, switching between languages has been identified as a very common practice when communicating misinformation (ibid.), and it therefore poses a big challenge to the accuracy of the detection, if they are developed to be only unilingual. The same goes for comparative or regional research that crosses linguistic borders.

A third challenge related to the current methods for detecting false information highlighted in new academic literature is that a lot of the current detection is based on fully supervised methods, which require human annotations. This is a time-consuming process, which makes it hard to follow the velocity at which misinformation is shared. Therefore, researchers are working on developing semi-supervised detection. A way to do this is by limiting the amount of annotation having to be made by humans. Preliminary research shows a potential for both higher quality and velocity of the detection by switching over to a semi-supervised method, where machine learning is used to make some of the annotations that would otherwise have to be made by humans (Abdali, Shah & Papalexakis, 2021).

3.2 COVID-19

COVID-19 is of course a big topic of discussion within the academic field studying disinformation. Furthermore, it is an unusual topic compared to the likes of research

studying detection methods, fact-checking or legislation, as COVID-19 has been used as a case to study all these different research topics as “people's everyday lives during this pandemic mirror a misinformation experiment” (Greenspan & Loftus, 2020).

Some research published within the last six months focus on the relationship between believing misinformation regarding COVID-19 and failure to comply with for example social distancing (Reyes et. al., 2021 & Chan, et. al., 2021), some arguing that this exposure to disinformation can affect people's memories of given events, if they are continually informed by false information (Greenspan & Loftus, 2020). This can be problematic as false memories of the severity of such an event can potentially affect attitudes towards similar future events. Related to this, research has also been conducted that indicates a relationship between the spread of misinformation and vaccine hesitancy (Lockyer, et. al., 2021).

Other research related to the COVID-19 pandemic has focused on governments' means of addressing corona-related misinformation. In one paper, researchers identified five categories of government activities to address the misinformation: “(1) disseminating and increasing access to accurate information; (2) restricting access to accurate information; (3) disseminating disinformation, false information, and misinformation; (4) addressing commercial fraud; and (5) criminalizing expression” (Pomeranz & Schwid, 2021). This same article problematizes some of these means, and states that governments should, instead of violating human rights by criminalizing expression, protect expression, disseminate factual information, protect whistleblowers and support the independent media environment (ibid.).

Some researchers also focus on the responsibility of the platforms regarding the sharing of misinformation related to COVID-19 (e.g. Greenspan & Loftus, 2020). The possible steps suggested to minimize this spread are: including warnings to posts, featuring related content and providing additional messaging before making it possible to share (ibid.).

3.4 Misinformation and human behavior

Another topic that was identified in multiple papers was the topic of misinformation and how it exists in interaction between people. As already presented in the previous section on COVID-19 related disinformation research, some relationship between believing misinformation regarding COVID-19 and failure to comply with for example social distancing (Reyes et. al., 2021 & Chan, et. al., 2021), as well as willingness to be vaccinated have been identified. (Lockyer, et. al., 2021).

In other research regarding the relationship between exposure and human behavior, it has been hypothesized that even short-term exposure to conspiracy theories as a general phenomenon can affect human behavior in heightening the degree of strategic sophistication (Balafoutas, et. al. 2021) – this meaning a heightened reflection on the incentives of others in decision making. Based on a laboratory experiment the researchers conclude that even simple contact with conspiracy theories can affect human behavior. The researchers do, however, clarify that, through short-term exposure conspiracy theories can be positive in heightening the strategic sophistication, the harmful effects of over-exposure and believing in such theories is very well documented in other academic literature (ibid.).

Finally, another facet of the relationship between information disorder and human behavior is the concept of self-corrections. One of the identified academic articles focuses on what motivates self-corrections, and they argue that self-correction should, instead of being seen as a unidirectional action, be seen as two-way communication, where citizens monitor and evaluate the spread of falsehood by others, as well as by themselves. In this sense self-correction can both happen based on external reactions to false information, but also internal reflection of accidental circulation (Koo, et. al., 2021).

4.0 Assessment of the scientific repository

As part of the EDMO-project a repository of relevant scientific articles as well as a repository of relevant policy papers will be established by Aarhus University in collaboration with ATC (IV.D.C & IV.D.D). These two subtasks are initiated, but their progress depends on other subtasks in Task 4, which have to be completed first.

IV.D.C and IV.D.D are respectively due to be finished by month 25 and 30 and are progressing as expected. Our strategy is to use the literature search results gathered in IV.D.A. as the first input into the repository of scientific content and progressively update the repository using searches with carefully selected keywords.

In the first *D.10 Periodically updated description of the DSI efforts to support and facilitate the coordination of academic research* it was mentioned that this literature search would be conducted on Google Scholar. However, during the preliminary searches and identification of keywords it became apparent that choosing Google Scholar as the access point would result in the exclusion of relevant literature, as their search engine only allows for a limited number of characters in the search bar. We instead chose to use the Danish Royal Library which gives access to search in 10.113 collections including Scopus and Web of Science.

The full description of the ongoing work with the establishment of the scientific repository can be found in the deliverable *IV.D.A: Academic research on disinformation at scale in the EU*.

4.1 Quantitative evaluation

The search initially resulted in 2.021 results. These have been filtered for relevance, defined as the relevance to research in digital media in relation to mis- and disinformation at scale. After the careful manual filtering, 117 results remain which is lower than the set goal of 200 entries in the final repository. This search is however reflecting the scope of EDMO (see method in IV.DA) and only preliminary, and it will be updated during the project as new research is published. The repository will furthermore be expanded by using input from the identified organizations in IV.D.B to identify other sources of relevant content. Finally, we consider expanding the repository to not only include literature in English through the assistance of the national and multinational hubs established during the second phase of EDMO. We will continually evaluate the criteria for the repository to secure meeting the goal of 200 relevant entries.

4.2 Qualitative evaluation

As for the qualitative evaluation of the literature identified in the initial literature search, all 2.021 were manually coded to secure included works that fulfill the relevance criteria (the main exclusion criteria can be found in deliverable IV.D.A). This manual coding was done to secure a high relevance of every piece of literature included in the repository. The final 117 results ended up including research from the fields of research: Computer Science and Information Science, Behavioral Science, Social Sciences, Media Law and Economics, Communication and Media Studies, Neuroscience and Psychology, Health and Medical SciencesCare Studies, and Others. This is seen as a positive fact, as a primary goal of the repository is to highlight research done within different fields of research to potentially foster interdisciplinary collaborations.

5.0 References

- Greenspan, R.L. & Loftus, E. F. (2021). "Pandemics and infodemics: Research on the effects of misinformation on memory". In: *Hum Behav & Emerg Tech*, 3, p. 8– 12. <https://doi.org/10.1002/hbe2.228>
- Reyes L. M., Ortiz L., Abedi M., Luciano Y., Ramos W. & Reyes P. (2021). "Misinformation on COVID-19 origin and its relationship with perception and knowledge about social distancing: A cross-sectional study". In: *PLoS ONE*, 16(3). <https://doi.org/10.1371/journal.pone.0248160>
- Chan, H.W., Chiu, C.P.Y., Zuo, S., Wang, X., Liu, L. & Hong, Y. (2021) "Not-so-straightforward links between believing in COVID-19-related conspiracy theories and engaging in disease-preventive behaviours". In: *Humanit Soc Sci Commun* 8, 104. <https://doi.org/10.1057/s41599-021-00781-2>
- Lockyer, B., Islam, S., Rahman, A., Dickerson, J., Pickett, K., Sheldon, T., Wright, J., McEachan, R., Sheard, L. (2021). "Understanding COVID-19 misinformation and vaccine hesitancy in context: Findings from a qualitative study involving citizens in Bradford, UK". In: *Health Expect.* <https://doi.org/10.1111/hex.13240>
- Pomeranz, J. L., Schwid, A. R. (2021). "Governmental actions to address COVID-19 misinformation". In: *J Public Health Pol.* <https://doi.org/10.1057/s41271-020-00270-x>
- Meel, P., Vishwakarma, D. K. (2021). "HAN, image captioning, and forensics ensemble multimodal fake news detection". In: *Information Sciences*, 567, pp. 23-41. <https://doi.org/10.1016/j.ins.2021.03.037>
- Abdali, S., Shah, N., Papalexakis, E. E. (2021). "Semi-supervised Multi-aspect Detection of Misinformation Using Hierarchical Joint Decomposition". In: Dong Y., Ifrim G., Mladenović D., Saunders C., Van Hoecke S. (eds) *Machine Learning and Knowledge Discovery in Databases. Applied Data Science and*

Demo Track. ECML PKDD 2020. Lecture Notes in Computer Science, 12461.
https://doi.org/10.1007/978-3-030-67670-4_25

Zhou, L., Lim, J., Alsaleh, H., Wang, J. & Zhang, D. (2020). "Language Alternation in Online Communication with Misinformation". In: Lang K.R. et al. (eds) *Smart Business: Technology and Data Enabled Innovative Business Models and Practices. WeB 2019. Lecture Notes in Business Information Processing, 403.*
https://doi.org/10.1007/978-3-030-67781-7_15

Koo, A. Z., Su, M., Lee, S., Ahn, S. & Rojas, H. (2021). "What Motivates People to Correct Misinformation? Examining the Effects of Third-person Perceptions and Perceived Norms", In: *Journal of Broadcasting & Electronic Media*, DOI: [10.1080/08838151.2021.1903896](https://doi.org/10.1080/08838151.2021.1903896)

Balafoutas, L., Libman, A., Selamis, V. & Vollan, B. (2021). "Exposure to conspiracy theories in the lab". In: *Economic and Political Studies*, 9(1), pp: 90-112, DOI: [10.1080/20954816.2020.1818930](https://doi.org/10.1080/20954816.2020.1818930)