European Digital Media Observatory

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

M6

This project has received funding from the European Union under Contract number: LC-01464044

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
Period covered by the periodic description:	from 01/06/2020 to 05/11/2020
Periodic description:	1st
Date of submission of the periodic description:	30/11/2020
Project website address:	https://edmo.eu/
The deliverable has been elaborated by:	Aarhus University (AU)

This project has received funding from the European Union under Contract number: LC-01464044

Introduction

This short document is the first *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 first of 5 deliverables that will be made every 6 months through the duration of the project. 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 organisations (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 and finalised and as researchers and organisations are further engaged. These reports will start with a presentation of the support and coordination of academic research in Europe followed by an overview of key research topics and findings within the field of information disorder, and finally an evaluation of the research repository (not applicable in this first iteration). This update covers the first six months of EDMO.

Support and coordination of academic research in Europe

In the following section we shortly summarize undertaken actions which aimed at supporting and coordinating academic research in Europe.

Truly Media Platform

A part of the actions taken to support academic research in Europe has been additional collaboration with ATC in order to make sure that the Truly Media platform addresses even more the needs of researchers. As an initial step, these needs were defined in discussions

This project has received funding from the European Union under Contract number: LC-01464044

among the researchers active in the EDMO project (from DATALAB at Aarhus University) and will be further expanded with the establishment of a research network. The discussions among the researchers at Aarhus University together with invited workshop with world-leading researchers in the area at Queensland University of Brisbane (Centre led by Axel Bruns), resulted in three primary points of attention: 1) making extracting lists in dynamic formats possible, 2) providing clear documentation of where the fact-checked stories derive from, and 3) providing additional information such as engagement scores, that is, how often and by whom are debunked stories of fact-checking services retrieved.

Among the researchers, it was made clear that for making the Truly Media platform even more attractive for researchers a high priority should be to establish the possibility of extracting the information about disinformation collected at the platform. This especially implies the possibility to extract a list of debunked stories/detected disinformation in dynamic formats such as json or as CSV-files in order to facilitate further analysis on external sources of various kinds. In computational social science research, researchers depend on using lists or similar datafiles in their research and will less likely use the user interface itself because researchers need to document every step they take and make sure to be in control of the process in order to make the research reliable and replicable at any given time together with the need to make the data and full code available for review. This is because the lists of debunked stories often need to be joined with other datasets and media lists, which can only be done if the lists are downloadable in dynamic formats. What the research community is missing at the moment, is clear lists of debunked stories that also include additional information such as country codes on the story level and source level.

In addition, a clear documentation of where the fact-checked stories derive from (e.g. Facebook, twitter) is needed preferably as a column in itself together with a URL, this also goes for pictures and videos that need a filename attached to them. The documentation should also take the merging of different channels such as the East Stratcom Task Force list or the list provided by FactCheck.org into consideration by harmonizing which and how information is documented. This clear documentation is key in order to make the platform

This project has received funding from the European Union under Contract number: LC-01464044

usable for research as researchers have a strong obligation towards transparency. Transparency in turn depends on a good documentation of methods to increase reliability and replicability and includes the outlining of all shortcomings and their potential implications.

Finally, the provision of engagement scores (number of likes, shares etc e.g. from CrowdTangle), country statistics and additional information are a potential additional contribution to research. For example, information about how often specific searches were conducted or who searches for specific terms/actors can contribute to a better understanding of the implications and spread of disinformation. Of course, data protection issues need to be addressed and data can only be made accessible if personal information is protected according to legal regulations such as the GDPR. In addition, the provision and merging of additional statistical data was addressed in the discussion. Examples of statistical data, which is potentially relevant for fact-checking for researchers, is provided in the following list:

a) Number of journalists/employees in media in a country/region;
b) Number of media in a country/characteristics of media in a country;
c) Information about the political system, party system;
d) Information about internet penetration: who uses the internet at which times, to what extent and in which ways;
e) Survey data such as trust in political institutions/ the media; perception of social media - maybe existing survey data, e.g. the Eurobarometer, includes interesting indicators;
f) Basic characteristics of the population for example the average educational level or age distribution;
g) Information about public media and financial expenses related to public media.

However, there is no exhaustive way to identify all the researchers' needs, since they depend on the research questions which are manifold. Thus, a close collaboration between ATC and Aarhus University – as well as researchers in general – is needed to ensure the

This project has received funding from the European Union under Contract number: LC-01464044

continued research relevance of the Truly Media platform. It was, however, noted by the researchers that additional statistical data was of less importance than facilitated data extraction, a unifying data model across fact checkers and a good documentation.

Interest shown in EDMO from researchers

A great interest in EDMO has been shown by European researchers and research organisations. Especially in connection with the 2020 CEF Telecom European Digital Media Observatory call[1] for national/multinational hubs for research on digital media, which is a part of the next phase of EDMO (European Commission, p. 3), both Aarhus University and the coordinating partner EUI received a number of enquiries from interested research organisations. As this is an ongoing call, the partners of EDMO could not assist these research organisations in their application process or answer specific questions, but had to refer to the call-text and the EU helpdesk. These research organisations will however be assessed and potentially be included in the list of relevant academic institutions and organisations (IV.D.B) and then be contacted later in the ongoing project.

Upcoming engagement with researchers

Aarhus University will assist EDMO in constructing the requirements for researchers to be members of EDMO in 2021. In this way, researchers can be an integrated part of EDMO and the resources, activities and data made available by the observatory. Here, we initially emphasize requirements for including academic research institutions that are part of the EU and thus comply with GDPR and that inclusion is done at the university level. Secondly, that European university teams can collaborate with academic universities outside the EU but are then held responsible for these research teams at host university level. Any breach of contract then falls on the host university. We argue for a strong need to separate membership for academic and non-academics in order to potentially utilize special legal and

[1]https://ec.europa.eu/inea/sites/inea/files/cefpub/2020-2_european_digital_media_observatory_call_text_final.pdf

This project has received funding from the European Union under Contract number: LC-01464044

ethical requirements for independent researchers that might provide them with other digital media access solutions going forward.

Relevant academic research

Disinformation and digital media is a vast field of research, and this will therefore not be an exhaustive list of the findings within the field, but rather a look into some relevant results. A more systematic description will be provided regarding IV.D.A. In this specific report, we have identified six relevant topics in the field of research, which are described next in more detail. This list of research topics is based on a literature review performed as part of the latest SOMA-report (D.2.3).

Terminology and definitions

A fundamental focus in the reviewed research is the definition of the research objects as well as the suggestion of adjustments and the challenging of the existing terminology. This goes for the field of information disorder as well, which is often split in different categories and subcategories that are either used interchangeably or carefully separated. Some of these terms include mis-/dis- and malinformation to describe respectively false information spread without intention of harm, false information spread with the intention to harm, and finally information based in reality but used to inflict harm (Carmi et al, 2020). Another term that is often used is “fake-news” (e.g. Graves, 2018) , which was initially used as a means of describing both dis- and misinformation in news reporting, but the term has now been adopted by political actors in an attempt to discredit news reporting or other political actors that they disagree with (ibid.). In some papers the terms “disinformation” and “fake-news” are, for example, used interchangeably (e.g. Shu et al., 2020), while others argue for distinguishing between the two (e.g. Gelfert, 2018; Vargo et al., 2018). Regarding the term “fake-news”, an analysis performed on 34 academic articles between 2003 and 2017 using the term has identified the use of the term to include: news satire, news parody, fabrication, manipulation, advertising, and propaganda (Tandoc et al. 2017).

This project has received funding from the European Union under Contract number: LC-01464044

Intentions behind the spread of disinformation

In extension of the previous point, some of the research done within the field of disinformation is focussed on the intentions behind spreading it. Misinformation differs from disinformation, as misinformation is not created and spread with the intention of causing harm or influencing the receiver(s), while disinformation is (e.g. Carmi et al., 2020). Therefore, the separation between mis- and disinformation is something that can be challenging, as it depends on the analysis of intentions, which are difficult to measure explicitly. There is no doubt that the spread of disinformation poses a threat to society, and therefore the intentions behind this intentional spread of false information are of interest to a number of researchers within the field. Newer research points to the fact that the intentions behind the spread are manifold and can vary from intentions of disrupting or influencing (foreign) societies to having profit-oriented ideological, normative or financial aims (see e.g. EU DisinfoLab, 2020; Shu et al., 2020; Tandoc et al., 2018).

Covid-19

As for more case-specific research, one topic is, not surprisingly, dominant. The Covid-19 pandemic has shown to be a relevant example of how false information can spread on digital media. For example, mis- and disinformation has been circulated stating that 5g radiation is the cause of the disease, which has resulted in engineers at telecoms having faced threats that are both physical and verbal in nature (Waterson, 2020). This sudden emergence of mis- and disinformation following the pandemic was with its new blending of medical and political misinformation and disinformation named *the first infodemic* (Alam et al., 2020). The nature of this infodemic is very complicated, not least due to the fact that factuality alone cannot explain it - much less solve it, as the infodemic not only includes fake news in the form of rumours and conspiracy theories, but also “promotion of fake cures, panic, racism, xenophobia, and mistrust in the authorities, among others” (ibid.). It is suggested that the way to solve this is by fostering multidisciplinary collaborations (ibid.), which will be expanded upon in section 3.5 and 3.6 of this report. An example of such interdisciplinary collaborations can be seen in Sharma et al.’s paper (2020), where they use

This project has received funding from the European Union under Contract number: LC-01464044

fact-checking sources to analyse sentiment, topics and emerging trends related to the Covid-19 pandemic.

Fact-checking & fact-checkers

A lot of research is focusing on fact-checking, fact-checkers and fact-checking sites (eg. Amazeen, 2020, Dias & Sippitt, 2020, Robertson et al., 2020). Fact-checking is something that has expanded both domestically and globally over the past decade (Amazeen, 2020, Dias & Sippitt, 2020). Some research studies fact-checking as a tool to strengthen democracy in areas where the democratic institutions "are perceived to be weak or are under threat" (Amazeen, 2020), while other research focuses on who uses fact-checking sites and what seems to be a divide in interest depending on political belief (Robertson et al., 2020). Focusing on American politics specifically, findings show that fact-checking sites have a particular appeal to liberal audiences while conservative users see these sites as less useful (ibid.). There are however still some key gaps in the academic literature within this field. One key gap is, as presented by Dias and Sippitt (2020), that fact-checking has only been viewed as a means of persuasion rather than a way of providing people with the information needed to make up their mind. This frames fact-checking as a practice with a specific political agenda rather than a means of providing needed information. Furthermore, Dias and Sippitt (2020) argue that a closer collaboration between researchers and fact-checkers is needed, where researchers should seek to reflect what is published by fact-checkers more closely by relying on fact-checkers' archives. This is, however, something that researchers such as Sharma et al. (2020) already started to do. There should be a continuation of this approach with high priority, as much of current research has been made under artificial, experimental conditions, thus, challenging the diversity and possibility of external validation of the research results (Dias & Sippitt, 2020). This need for closer collaborations between researchers and fact-checkers is a point that is both echoed in case-research on the Covid-19 infordemic (Alam et al., 2020) as well as in the newest SOMA-report, where it is argued that especially European organizations, which provide fact-checking services and index sites, need to increase their attraction for researchers (Walter et al., 2020). Thus, adapting the Truly Media platform to researchers' needs would address

This project has received funding from the European Union under Contract number: LC-01464044

this gap in research and help to make the results even more relevant to stakeholders such as journalists, fact-checkers and legislators. Furthermore, the fostering of closer collaborations between journalists, fact-checkers, policymakers, government entities, social media platforms, and society as a whole would be a way of adopting a more holistic approach, which potentially leads to a more efficient and faster way of overcoming challenges connected to the ongoing Covid-19 infodemic (Alam et al., 2020).

Disinformation detection tools and methods

In the latest SOMA-report (D.2.3), which includes a systematic literature-review, it is concluded that two different approaches within disinformation detection can be distinguished - namely a practice-oriented approach and a research-oriented approach. The practice-oriented approach is typically characterized by a journalistic perspective (Graves, 2018) and often relies on experts in order to debunk disinformation. One of the primary challenges of this approach is at the moment the high velocity and vast spread of disinformation in peaks of digital media usage, which an expert-centered approach can hardly meet (Shu et al., 2019). Besides, a research-oriented approach can be described, which is often based on automation of disinformation detection for example by identifying abnormalities in a dataset. An example is the use of natural language processing methods in order to detect abnormalities within the text body (Zhang & Ghorbani, 2020). There are, however, also potential improvements in this approach. In the SOMA-report (D.2.3) it is suggested that there should be a focus on, for example, the identification of intentions and the impact of disinformation, as well as on an intensified interdisciplinary approach between researchers within the fields related to disinformation research for example the field of computer science, neuroscience, psychology or social sciences (Walter et al., 2020). A general point that applies to fact-checkers, journalists and researchers is that it can be challenging to identify trustworthy tools or to find the appropriate tool for checking a specific information since existing tools apply different methods and evaluate different statements (Lim, 2018).

This project has received funding from the European Union under Contract number: LC-01464044

Finally, research is focused on potential policy changes to help combatting disinformation. In recent years the EU has been using soft law and self-regulation policies and strategies to accelerate the combat against misinformation and disinformation and it is suggested that this might not be the most beneficial approach (Bechmann, 2020). Disinformation has had an increased political focus in the wake of Brexit, Trump political campaigns, COVID-19 and general information warfare from Russia (Bechmann 2020; Buning et al. 2018). This can be seen in the backdrop of researchers pointing to the possible danger of disinformationon decreased trust in organizations and democratic systems (Bradshaw & Howard, 2018), but also that disinformation may be used to manipulate voters by targeting core values in democratic society (Guess et al., 2019; Lazer et al., 2018), and that such disinformation strategies may potentially increase polarization (Benkler et al., 2018; Osmundsen et al., 2020). Even though little empirical research exists on the area, the legislative part of the combat against disinformation is of course a point of attention. One suggestion on how to combat the spread of disinformation, from a policy point-of-view, is to focus more on creating a responsible media infrastructure rather than just focussing on content-level problem solving (Bechmann, 2020). It is argued that this media infrastructure needs to take into consideration "the global nature of media industries, the collapse of media and tech industries, and the extensive use of data and machine learning algorithms as driving engines on modern media platforms" (ibid.). Furthermore, Helberger even adds "current approaches quite literally establish platform governance, by making platforms the primary governors of online communication." (Helberger, 2020). In other words, she highlights that soft law and self-governance cannot in any way provide the media pluralism needed in order to solve the information disorder we see in critical democratic events (e.g. elections) and health crises as seen at the moment with COVID-19. Instead, she turns to the need to revisit EU law in order to fit media concentration law to a "digital media reality with online platforms" (ibid).

This project has received funding from the European Union under Contract number: LC-01464044

Assessment of the scientific repository

As part of the EDMO-project a repository of relevant scientific articles will be established by Aarhus University in collaboration with ATC (IV.D.C & IV.D.D). These two tasks are initiated, but their progress depends on other tasks in Task 4, which have to be completed first. Our strategy is to use the literature search results gathered in IV.D.A. as the first input into the repository and progressively update the repository using Google Scholar searches with carefully selected keywords.

This project has received funding from the European Union under Contract number: LC-01464044

References

Alam, F., Dalvi, F., Shaar, S., Durrani, N., Mubarak, H., Nikolov, A., Martino, G. D. S., Abdelali, A., Sajjad, H., Darwish, K., Nakov, P. (2020). Fighting the COVID-19 Infodemic in Social Media: A Holistic Perspective and a Call to Arms. arXiv.org. https://arxiv.org/abs/2007.07996

Bechmann, A. (2020). Tackling Disinformation and Infodemics Demands Media Policy Changes, Digital Journalism, 8(6): 855-863. https://doi.org/10.1080/21670811.2020.1773887

Benkler, Y ., Faris, R., & Roberts, H. (2018). Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics, Oxford University Press, New York (NY).

Buning et al., M. d. C. (2018). A Multi-dimensional Approach to Disinformation: Report of the independent high level group on fake news and online disinformation. EU Commission. 1-40. https://ec.europa.eu/digital- single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation, ISBN 978-92-79-80420-5

Bradshaw, S., & Howard, P. N. (2018). The Global Organization of Social Media Disinformation Campaigns. Journal of International Affairs, 71(1.5): 23–32. https://www.jstor.org/stable/26508115

Carmi, E., Yates, S. J., Lockley, E., & Pawluczuk, A. (2020). Data citizenship: Rethinking data literacy in the age of disinformation, misinformation, and malinformation. Internet Policy Review, 9(2). https://doi.org/10.14763/2020.2.1481

Dias, N., & Sippitt, A. (2020). Researching Fact Checking: Present Limitations and Future Opportunities. The Political Quarterly, 1467-923X.12892. https://doi.org/10.1111/1467- 923X.12892

EU DisinfoLab. (2020). The Few Faces of Disinformation (p. 7). https://www.disinfo.eu/wp-content/uploads/2020/05/20200512_The-Few-Faces-of-Disinformation.pdf

European Commission. (2020). Call for Proposals Concerning Projects of Common Interest under the Connecting Europe Facility in the Field of Trans-European Telecommunication Networks.

This project has received funding from the European Union under Contract number: LC-01464044

https://ec.europa.eu/inea/sites/inea/files/cefpub/2020-2_european_digital_media_observatory_call_text_final.pdf

Gelfert, A. (2018). Fake News: A Definition. Informal Logic, 38(1), 84–117. https://doi.org/10.22329/il.v38i1.5068

Graves, L. (2018). Boundaries Not Drawn: Mapping the institutional roots of the global fact-checking movement. Journalism Studies, 19(5), 613–631. https://doi.org/10.1080/1461670X.2016.1196602

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. Science Advances, 5(1): eaau4586. https://doi.org/10.1126/sciadv.aau4586

Helberger, N. (2020). The Political Power of Platforms: How Current Attempts to Regulate Misinformation Amplify Opinion Power. Digital Journalism 8(6): https://www.tandfonline.com/doi/full/10.1080/21670811.2020.1773888

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. Science, 359(6380): 1094–1096. https://doi.org/10.1126/science.aao2998

Lim, C. (2018). Checking how fact-checkers check. Research & Politics, 5(3), 205316801878684. https://doi.org/10.1177/2053168018786848

Osmundsen, M., Bor, A., Vahlstrup, P. B., Bechmann, A., & Petersen, M. B. (2020). Partisan polarization is the primary psychological motivation behind "fake news" sharing on Twitter. https://doi.org/10.31234/osf.io/v45bk

Robertson, C. T., Mourão, R. R., & Thorson, E. (2020). Who Uses Fact-Checking Sites? The Impact of Demographics, Political Antecedents, and Media Use on Fact-Checking Site Awareness, Attitudes, and Behavior. The International Journal of Press/Politics, 25(2), 217–237. https://doi.org/10.1177/1940161219898055

Sharma, K., Seo, S., Meng, C., Rambhatla, S., Liu, Y. (2020). COVID-19 on Social Media: Analyzing Misinformation in Twitter Conversations. arXiv.org. https://arxiv.org/abs/2007.07996

This project has received funding from the European Union under Contract number: LC-01464044

Shu, K., Bhattacharjee, A., Alatawi, F., Nazer, T. H., Ding, K., Karami, M., & Liu, H. (2020). Combating disinformation in a social media age. WIREs Data Mining and Knowledge Discovery. https://doi.org/10.1002/widm.1385

Shu, K., Mahudeswaran, D., & Liu, H. (2019). FakeNewsTracker: A tool for fake news collection, detection, and visualization. Computational and Mathematical Organization Theory, 25(1), 60–71. https://doi.org/10.1007/s10588-018-09280-3

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

Vargo, C. J., Guo, L., & Amazeen, M. A. (2018). The agenda-setting power of fake news: A big data analysis of the online media landscape from 2014 to 2016. New Media & Society, 20(5), 2028–2049. https://doi.org/10.1177/1461444817712086

Walter, J. G., Sørensen, M. H., Bechmann, A. (2020). Detecting mis- and disinformation in digital media; SOMA report D2.3: Outlier (disinformation) detection solution; H2020-ICT-2018-2020; p.:49; https://www.disinfobservatory.org/resources/

Zhang, X., & Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and discussion. Information Processing & Management, 57(2), 102025. https://doi.org/10.1016/j.ipm.2019.03.004

This project has received funding from the European Union under Contract number: LC-01464044