European Digital Media Observatory

IV.D.D: Research on disinformation about the war in Ukraine and outlook on challenges in times of crisis

Date: 30-11-2022

Final version

European Digital Media Observatory

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 Tveen, DATALAB, Aarhus University Jessica Walter, DATALAB, Aarhus University Anja Bechmann, DATALAB, Aarhus University

European Digital Media Observatory

Executive summary

In this report we present the final part of Task IV.D, IV.D.D, that involved the active participation in the EDMO Task Force on Disinformation and the War in Ukraine, as well as preparing an academic paper on a research question raised in the Task Force. The previous three IV.D. reports have described the: mapping of relevant academic activities in the EU (IV.D.A), identification of relevant academic institutions and organizations in the EU (IV.D.B), and creation of a repository of relevant scientific articles (IV.D.C). Originally, IV.D.D entailed the creation of a repository of relevant policy papers; however, the task was changed in response to the need to address disinformation circulating on digital platforms in the wake of the Russian/Ukrainian war.

Further, the report presents remaining challenges academics face when researching digital disinformation in times of crisis. The shared purpose of all subtasks in Task IV has been to support and coordinate research on digital disinformation in the EU context. Through this work, the following five main challenges have been identified:

1) Restrictions on data access for researchers on most social media platforms;

2) Restrictions in answering research questions based on the provided data. Twitter, for example, does not share users' geolocation or information about deleted mis-/ and disinformation, and is not demographically representative;

3) Scarcity of computational tools for analysis of low- and middle-resource languages;

4) Inherent biases in fact-checking databases;

5) Difficulty sharing data across academic institutions due to GDPR.

European Digital Media Observatory

Table of content

1. Introduction

This report focuses on the problems academics face when studying digital disinformation in times of crisis. During the first project period of EDMO, especially the COVID-19 pandemic and the current war in Ukraine are examples of events that have stimulated a spread of false and misleading information on digital platforms. From the beginning of EDMO, the long-term aim of Task IV has been to provide support and coordination for academic research activities on digital disinformation in the European Union. In response to the outbreak of the Russian/Ukrainian war, the content of the last deliverable (IV.D.D) changed from the establishment of a repository of policy papers and other content to instead address the Russian-Ukrainian war. This was done through participation in the EDMO Task Force on Disinformation and the War in Ukraine and preparation of a scientific article addressing a research question raised in the Task Force.

In this report, Buning's (2018) definition of disinformation as: “false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit” is expanded on by also including unintentionally false information, often described as misinformation (for further conceptualizations see e.g., Kapantai et al., 2021; Tandoc et al., 2018; Wardle & Derekhshan, 2017). Hereby, disinformation is used as a cover term for various types of false information in consistency with previous periodical reports.

This is the last of four reports included in work-package IV.D. of EDMO. The previous reports concerned academic research on disinformation at scale in the EU (IV.D.A), a list of relevant academic institutions and organizations (IV.D.B), and the final repository of relevant scientific articles (IV.D.C). This report proceeds in two main sections: the first describes the work undertaken in IV.D.D, and the second challenges researchers face when investigating digital disinformation in times of crisis.

2. Work on IV.D.D: Studying disinformation in war times

The outbreak of the Russian/Ukrainian war has yet again underlined the importance of addressing the spread of digital disinformation with both sides spreading misleading and outright false information. In this sense, disinformation plays a role in the escalation of the conflict, both in terms of assessing the development of the war, compromising citizens' right to access trustworthy information, and swaying public opinion. From the infodemic surrounding the COVID-19 pandemic, we know that digital disinformation cross national borders (Stephens, 2020). This means that disinformation circulating within and outside member states of the European Union presents a threat to democratic processes in all member states by causing confusion, altering the public democratic debate and creating divides. For both COVID-19 and the war in Ukraine studies provide solid documentation that disinformation spreads fast and broadly (e.g., Stephens, 2020; Vosoughi et al., 2018) on digital platforms.

2.1. EDMO Task Force on Disinformation and the War in Ukraine

The EDMO Task Force on Disinformation and the War in Ukraine was established on the 3rd of March 2022. Anja Bechmann, professor, and director of DATALAB at Aarhus University, has joined the Task Force along with 17 other experts with backgrounds, for example, in academia, journalism, and news media. The Task Force has worked with a strong academic research focus to provide policy- and decision-makers, public and private stakeholders, and the public at large relevant insights into the problematics of studying digital disinformation in war times. One of the main points stressed by the Task Force is that for disinformation to stop circulating, a global response is needed. Based on the work carried out by the Task Force since its establishment, it has delivered ten recommendations related to countering digital disinformation during war times. The recommendations are described in full length on the following webpage: https://edmo.eu/2022/06/29/10-recommendations-by-the-taskforce-on-disinformation-and-the-war-in-ukraine/.

2.2. Preparing an academic paper on the war

The research team at DATALAB is preparing an academic paper on a research question raised in the Task Force related to the war in Ukraine for submission to a high ranked scientific journal. The paper in process is an empirical investigation of Twitter users' emotional responses to pro-Russian and Pro-Ukrainian debunks, respectively, across a broad range of European countries, representing different cultural and historical contexts with respect to the conflict and war between Ukraine and Russia. The data is going to be collected in two steps: first, debunked stories will be collected via different European fact-checking organizations' websites, by scraping the Google Fact Check Explorer applying different keywords, e.g., Ukraine, Zelensky, Russia, Putin, and Kijev, in national languages, and by using the EDMO fact-checking repository. Following manual coding of whether Ukraine or Russia benefitted from the debunk, the fact-check URLs will be used to filter data from Twitter. Ultimately, the paper will identify potential systematic differences in fact-checking in different European countries and variances in Twitter users' emotional reactions to pro-Russian and pro-Ukrainian fact-checks shared on this platform.

2.3. ECREA Post-Conference on Digital Media and Information Disorders

In October 2022 the 9th European Communication Research and Education Association (ECREA) conference took place in Aarhus. Following the main conference, DATALAB hosted the hybrid post-conference "Digital media and information disorders: Theorizing and investigating influence operations, impact and contextual factors" on 24th of October 2022. The post-conference was co-funded by EDMO and advertised in the EDMO calendar. The post-conference contributed to the overall aim of Task IV by bringing together young and senior researchers from around the world that are active in different areas of digital disinformation. The post-conference was an opportunity to present and discuss recent research, and hereby, establish new collaborations and/or inspiring new research projects. The participating

speakers presented research focused on various EU countries, such as Finland, Lithuania, Belgium, Denmark, Norway, Sweden, and Italy, and different EDMO Hubs were represented at the event.

3. Additional academic challenges

Besides from accounting for the fulfillment of Task IV.D.D, an objective of this report is to give an overview of challenges to research on digital disinformation in times of crisis. With this aim, the following subsections summarize the obstacles for academic research on digital disinformation in times of crisis that have been identified through the efforts to support and coordinate research through different subtasks (IV.D.A-IV.D.D) in Task IV. In addition, the presented challenges have also been identified through DATALAB's research on different aspects of digital disinformation and fact-checking: for instance, on digital infrastructures of COVID-19 misinformation fact-checking (Nissen et al., 2022), types of COVID-19 misinformation and their emotional valence on Twitter (Charquero-Ballester et al., 2021), and a review of academic research on digital false information at scale in Europe (Bak et al., 2022). The challenges can be divided into five main themes: 1) restrictions on access to social media platforms, 2) limitations of Twitter data, 3) scarcity of computational methods for low- and middle-resource languages, 4) fact-checkers as a data-source, and finally, 5) cross-country collaborations. These challenges are most pressing for quantitative studies with large datasets that aim at inferring conclusions to larger populations and, hereby, provide general insights into certain aspects of the disinformation surrounding the crisis.

3.1. Restricted access to social media platforms

Social media platforms play an increasingly prevalent role in Europeans' lives (Eurostat, 2021), and are today used for many different purposes, from information sharing and seeking to news consumption, entertainment, and socialization. From an empirical research perspective, the increased use of these platforms presents new

opportunities to study a wide range of open questions, also beyond disinformation research, as human behavior on digital platforms is traced. Further, data from digital social media platforms are unique data sources as they connect people across borders, while users often are not aware of how their data is used. This means researchers can get insights into actual exposure, engagement patterns, networks, and so forth.

Despite these motivating possibilities for research on social media data, the platforms restrict data access to an extent that prevents the full use of these extraordinary data archives. For instance, Meta, allows researchers access to so-called *privacy-safe* data from Facebook and Instagram via Crowd-tangle,[1] but only to publicly available data (Meta data, n.d.) on their platforms, including Facebook, Instagram, and WhatsApp. Similarly, the video-blogging platform YouTube provides scaled access to video metadata from public content only. Finally, while today Twitter has a comparably liberal data sharing policy that allows for extraction of large scale datasets through the platform's Application Programming Interface (API) (Stephens, 2020), protected tweets are not included (Twitter Community, 2021). Still, the structural properties of Twitter with only basic functions, such as retweet, mention, hashtags, and a character limit, make Twitter data much easier to analyze compared to, for instance, Facebook's varying formats. This may be one of the reasons why Twitter is the most studied social media platform in large scale studies on digital disinformation in the EU (Bak et al., 2022), and in social media data research in general (Tufekci, 2014). In times of crisis, private pages constitute blind spots in the datasets that can systematically skew the findings on different platforms, for instance, caused by demographically skewed user groups.

[1] At the time of writing, it is uncertain whether CrowdTangle or a similar resource will be available in the future. Meta has announced a new solution for research access to data and paused new users from gaining access to CrowdTangle (Reuters, 2022).

Finally, a general challenge to digital disinformation studies on social media data, is the platforms' hesitancy to answer researchers' requests. For instance, during the current war in Ukraine, the Task Force chaired by Claire Walter contacted social media platforms to ask for data related to the war, but none of the contacted provided data (EDMO, 2022). However, Facebook did not deliver timely data and does not facilitate research access to all data created on the platform. To gain a more comprehensive understanding of how disinformation spreads across platforms during a crisis, it is crucial to get access to data from all social media platforms.

3.2. Problems with Twitter data

As Twitter is the most studied social media platform in the field of digital disinformation studies in the EU context (Bak et al., 2022) and serves as "best case" as a platform cooperating with academic researchers (as discussed in the previous section), we will use this section to zoom in on Twitter's two main limitations researchers still face, based on our own academic work: that is, Twitter does not share 1) complete geolocation information of tweets, nor 2) information of the extent and systematics of content deletion. Together, these limitations are preventing research with a country specific focus or a focus on cross-country analysis. Further, the limitations disrupt the conclusions as it is not transparent how much of the data is missing from the datasets. In times of crisis, this means that it is difficult to determine from which countries claims originate and prevail, as well as how generally known certain claims are.

With regards to geolocation, Twitter does offer researchers access to two types of geographical metadata: tweet location, if shared by the user when tweeting, and account location, if provided by the user (Twitter, n.d.). As far from all users adjust their geolocation, scraping based on these settings return fewer tweets. If the extracted datasets are incomplete, it presents an issue for the reliability of the research. If instead Twitter provided this information, it could be the basis of research that could help guide policymakers, fact-checkers, and other stakeholders in allocating

resources to counter disinformation, especially in times of crisis when rapid reactions to disinformation are essential.

A way researchers can work around the lack of tweets searched on geolocation, is to instead extract data based on language settings, for instance, by using language specific stop words, i.e., commonly used words such as: 'to', 'a', 'and' in English. However, the notable deficit of this method is that users can tweet in non-native languages, be bilingual, or share something in a native language while living abroad. Finally, another approach is to use the time zone as an indicator of geolocation, however, this approach is not effective in an EU context where most member states are in the same time zone. Further, these methods could also return results from countries outside the EU, for instance, French speaking African countries.

Finally, in response to the problem of digital disinformation on their platform, Twitter removes content that violates the platform's policy, such as disinformation. This enforcement of the policy restricts researchers' data access, especially to disinformation, as Twitter is not transparent with the extent of deletion nor provides researchers insights into the deleted content. What is more, Twitter does not account for the exact reasons for deletion or for how systematic deletion is carried out and based on which rules or regulations. In consequence, it is difficult to draw generalizable observations from the findings of studies based on Twitter data as data deletion can cause measurement errors and constitutes an undocumented blind spot compromising the validity of the research conducted on Twitter data. Researchers to some extent could overcome the problem by scraping the same criteria constantly and to identify deleted content themselves. However, this requires substantial data storage, computational capacities, and personnel.

3.3. Computational methods for low-resource languages

One challenge for large scale studies in the EU context is the lack of natural language processing (NLP) tools to analyze the 24 official languages. At present, it is more

convenient to use English language data for scientific papers, as there are more tools available (Bazzaz Abkenar et al., 2021). The creation of new tools for low- and middle-resourced languages in the EU is mostly attractive to linguists interested in each language and researchers dedicated to developing these tools. It is possible that the preference for English data is further strengthened by the industry in English-speaking countries that invests both money and time in the development of English NLP tools, e.g., to monitor customer satisfaction, brand related discussions, and trends on digital media. Another contributing factor could be a more general Anglocentric bias in research (Levisen, 2019).

A further challenge is to develop models that work in several languages comparably well to enable analysis across countries. This could, for instance, allow for comparative analysis of how users in different countries respond emotionally to information related to a given crisis. For this to happen it is necessary to establish wide-covering and available corpora in all European languages to build efficient automated analysis tools for these low- and middle resource languages. Even though this problem is prevalent in all contexts in which low-resource language texts need to be analyzed, it is especially critical in times of crisis as tool development is a time intense process that hinders a quick response.

Some corpora projects already exists for Danish, French, Spanish, and Icelandic (Strømberg-Derczynski et al., 2020). A source of inspiration for future projects could be the Danish Gigaword Project at Aarhus University. Here, a collective of researchers have created a Danish corpus that is representative, accessible, and general-purpose, i.e., covering different times, sociolects, dialects, and so forth (Strømberg-Derczynski et al., 2020). Establishing a corpus is time consuming and it is difficult to automize the process, for instance, with Danish the automatized sampling blends Danish with Norwegian Bokmål and is biased toward web-speak and modern use of language.

3.4. Fact-checking as a data-source

Researchers and fact-checkers could benefit mutually from a stronger collaboration (Walter et al., 2020), as fact-checking organizations, on the one hand, have large archives of debunked claims that can be used by researchers as baselines for identifying and characterizing false content. Further, researchers can support fact-checkers by analyzing audiences, spread, effect of exposure and influence. From the beginning of EDMO, the benefit of uniting fact-checkers and researchers in the fight against disinformation has been an apparent and central focus. Both tasks are of high importance in times of crisis, as shown during the Covid-19 infodemic, for instance to estimate anti-vaxx narratives influence of vaccine compliance, and to use fact-checkers resources on the most damaging narratives. At this point, EDMO has taken important steps to integrate fact-checkers and researchers' workflows, for instance by establishing the regional Hubs.

An obstacle for researchers' use of fact-checkers' databases is, among other things, variations in the source transparency between fact-checkers (Humprecht, 2020); in a study of eight fact-checkers in the US, the UK, Germany, and Austria, Humprecht (2020) concluded that these variate in professionalism and organizational structure. This lowers the comparability of data across fact-checking organizations. Similarly, Nissen et al. (2022) found great variance between the two fact-checking infrastructures Poynter and Google Fact Check Explorer in terms of who is conducting the fact-checks, which platform they focus on, the people they focus on, which themes are fact-checked, and, finally, the number of categories assigned to the claims in the context of the COVID-19 pandemic. Taken together, the differences and inherent biases need to be considered if infrastructures are used as data sources for academic investigations, as high standards for transparency of the methods and data are preconditions for sound research.

In addition, a more practical problem is that the fact-checking websites are not optimized for scraping, i.e., automating the data extraction process. This presents

another obstacle for potential use of fact-checking databases for large-scale academic research. Finally, fact-checking as an approach that still strongly relies on experts and human work cannot provide a comprehensive account of all circulating disinformation due to limited personnel and financial resources (Konstantinovskiy et al., 2021), the fast velocity in which false content is produced, and limited data access also for fact-checking organizations. Thus, in the future, the collaboration between fact-checkers and academics can be further improved by increasing transparency and streamlining of fact-checking databases, as well as by optimizing the fact-checking sites for scraping the content.

3.5. Collaboration between universities

As previously stated, the overall objective with Task IV has been to support and coordinate academic research. Research suggest that fostering collaborations on digital disinformation across disciplines is beneficial for the research community, for instance to improve detection methods (Walter et al., 2020), and across national borders to identify similarities and differences on cross-national, regional and international levels (Bak et al., 2022). However, data protection rules make it difficult to share information with external partners, and hereby, to collaborate across research institutions. This is a challenge that requires new structures for how data can be shared or at least analyzed across universities in a GDPR compliant manner, which now is often difficult especially for large scale datasets that require storage facilities. Further, research collaborations and financing take time to establish, which can make it challenging to respond quickly to new crises.

4. Conclusion and outlook

During the first project period of EDMO, important efforts have been made to support and coordinate academic research on digital disinformation. The completion of the subtask IV.D.D has contributed to this focus by active participation in the EDMO Task Force on Disinformation and the War in Ukraine, and by preparing an academic paper

on a related topic. In addition, the preparation of a post-conference on information disorder has brought together researchers and hereby created an opportunity to establish new collaborations and exchange of ideas.

While enhancing the visibility of academic research on digital disinformation in the EU context and establishing a network of academic scholars active in the field, the work carried out in Task IV.D over the course of the project period has led to the identification of five focal points for future endeavors to combat disinformation, namely to:

1. Communicate to platforms the need to provide academic data access that enables the analysis of various research questions across platforms, especially in times of crisis.
2. Encourage Twitter and other platforms to provide access to geolocations, and information about data removed from Twitter for violating the platform's policies or, in best case, an archive of deleted content.
3. Develop better analytical methods for low- and middle-resource languages to be able to study the disinformation phenomena at scale on national, regional, and international levels and hereby, find synergies and discrepancies in disinformation characteristics.
4. Enhance standardization and transparency across fact-checking organizations and infrastructures to expose and minimize inherent biases.
5. Identify GDPR compliant ways to share data between universities to allow for more and stronger collaborations between research organizations.

5. References

Bak, P. de P., Walter, J. G., & Bechmann, A. (2022). Digital false information at scale in the European Union: Current state of research in various disciplines, and

future directions. *New Media & Society*, 146144482211221. https://doi.org/10.1177/14614448221122146

Bazzaz Abkenar, S., Haghi Kashani, M., Mahdipour, E., & Jameii, S. M. (2021). Big data analytics meets social media: A systematic review of techniques, open issues, and future directions. *Telematics and Informatics*, *57*, 101517. https://doi.org/10.1016/j.tele.2020.101517

Buning, N. de C. (2018). *A multi-dimensional approach to disinformation—Publications Office of the EU*. https://op.europa.eu/en/publication-detail/-/publication/6ef4df8b-4cea-11e8-be1d-01aa75ed71a1/language-en

Charquero-Ballester, M., Walter, J. G., Nissen, I. A., & Bechmann, A. (2021). Different types of COVID-19 misinformation have different emotional valence on Twitter. *Big Data & Society*, *8*(2), 205395172110412. https://doi.org/10.1177/20539517211041279

EDMO. (2022). *10 Recommendations by the Taskforce on Disinformation and the War in Ukraine*. Edmo.Eu. https://edmo.eu/2022/06/29/10-recommendations-by-the-taskforce-on-disinformation-and-the-war-in-ukraine/

Eurostat. (2021, June 30). *Do you participate in social networks?* https://ec.europa.eu/eurostat/web/products-eurostat-news/-/edn-20210630-1

Humprecht, E. (2020). How Do They Debunk “Fake News”? A Cross-National Comparison of Transparency in Fact Checks. *Digital Journalism*, *8*(3), 310–327. https://doi.org/10.1080/21670811.2019.1691031

Kapantai, E., Christopoulou, A., Berberidis, C., & Peristeras, V. (2021). A systematic literature review on disinformation: Toward a unified taxonomical framework. *New Media & Society*, *23*(5), 1301–1326. https://doi.org/10.1177/1461444820959296

Konstantinovskiy, L., Price, O., Babakar, M., & Zubiaga, A. (2021). Toward Automated Factchecking: Developing an Annotation Schema and Benchmark for Consistent Automated Claim Detection. *Digital Threats: Research and Practice*, *2*(2), 1–16. https://doi.org/10.1145/3412869

Levisen, C. (2019). Biases we live by: Anglocentrism in linguistics and cognitive sciences. *Language Sciences*, *76*, 101173. https://doi.org/10.1016/j.langsci.2018.05.010

Meta data. (n.d.). https://research.facebook.com/data/

Nissen, I. A., Walter, J. G., Charquero-Ballester, M., & Bechmann, A. (2022). Digital Infrastructures of COVID-19 Misinformation: A New Conceptual and Analytical Perspective on Fact-Checking. *Digital Journalism*, 1–23. https://doi.org/10.1080/21670811.2022.2026795

Reuters. (2022, January 29). *Meta pauses new users from joining analytics tool CrowdTangle* [News site]. Reuters.Com.

Stephens, M. (2020). A geospatial infodemic: Mapping Twitter conspiracy theories of COVID-19. *Dialogues in Human Geography*, *10*(2), 276–281. https://doi.org/10.1177/2043820620935683

Strømberg-Derczynski, L., Ciosici, M. R., Baglini, R., Christiansen, M. H., Dalsgaard, J. A., Fusaroli, R., Henrichsen, P. J., Hvingelby, R., Kirkedal, A., Kjeldsen, A. S., Ladefoged, C., Nielsen, F. Å., Petersen, M. L., Rystrøm, J. H., & Varab, D. (2020). *The Danish Gigaword Project*. https://doi.org/10.48550/ARXIV.2005.03521

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

Tufekci, Z. (2014). Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls. *Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014*.

Twitter. (n.d.). *Filtering Tweets by location*. Developer Twitter. https://developer.twitter.com/en/docs/tutorials/filtering-tweets-by-location

Twitter Community. (2021). https://twittercommunity.com/t/does-the-api-academic-research-product-track-exclude-protected-accounts/156780

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Wardle, C., & Derekhshan, H. (2017). *INFORMATION DISORDER: Toward an interdisciplinary framework for research and policy making*. https://rm.coe.int/information-disorder-report-november-2017/1680764666?ct=t()