Adria Digital Media Observatory

MONITORING THE IMPLEMENTATION OF THE CODE OF PRACTICE ON DISINFORMATION IN EMPOWERING THE RESEARCH COMMUNITY IN CROATIA AND SLOVENIA

Mato Brautović

Author
Mato Brautović
University of Dubrovnik

Graphic Designer
Katarina Banović

Dubrovnik, January 2024

Coordinator of the ADMO
University of Dubrovnik

Partners
Faculty of Political Science (University of Zagreb)
Faculty of Electrical Engineering and Computing (University of Zagreb)
Faculty of Social Sciences and Humanities (University of Zagreb)
Gong
Oštro
Adria news
Xwiki
AFP

SUMMARY

This report is part of the research activities carried out within the project "Adria Digital Media Observatory" funded by the Digital Europe Program. The aim was to monitor the implementation of the Code of Practice on Disinformation in empowering the research community in Croatia and Slovenia.

Key findings:

- Access to platform data for researchers in the EU is still in its early stages.

- Researchers used data from Meta (N=9), X (N=8), Google (N=5), TikTok (N=2), Reddit (N=2) and Microsoft (N=1).

- Four researchers (N=4) requested "access to Meta data "4 or more" times, on average, the approval process took "more than 8 weeks" for five researchers (N=5) and six researchers (N=6) reported that access to data was denied without explanation.

- Five researchers (N=5) from Croatia and Slovenia reported that they were treated unequally or unfairly when applying for data access compared to other regions/countries.

- Seven researchers (N=7) reported that accessing the data from X and Meta was "difficult" or "very difficult" for them.

- Five researchers (N=5) had their access to platform data withdrawn after initially being granted access.

INTRODUCTION

There are two main policy documents that regulate access to platforms' data for researchers in the EU: Code of Practice on Disinformation (CoP) and Digital Service Act (DSA). The strengthened Code of Practice on Disinformation from 2022 that was signed by 34 signatories regulates access to platforms data for disinformation researcher in commitments 26 to 29. The CoP requires researchers to go through a third-party review process to ensure that researchers are not misusing access for malicious or commercial purposes. Even if the independent third party body is not established, the platform should grant access (Regulation 2065:2022).

"Relevant Signatories commit to provide access, wherever safe and practicable, to continuous, real-time or near real-time, searchable stable access to non-personal data and anonymised, aggregated, or manifestly-made public data for research purposes on Disinformation through automated means such as APIs or other open and accessible technical solutions allowing the analysis of said data." (CoP, 2022)

Similarly, DSA (Art. 40.4 & 40.12) provides regulation for access to data that are necessary to monitor and assess compliance with DSA for very large online platforms or very large online search engines. Based on researcher's request, platforms have right within 15 days to not give access to the data requested because "they do not have access to the data" or "access to the data will lead to significant vulnerabilities in the security of their service or the protection of confidential information, in particular trade secrets". (Regulation 2065:2022).

Researchers who wish to obtain the status of 'vetted researchers' and gain access to the required data must demonstrate that they belong to a research organization, that the data will not be used for commercial purposes, that they have disclosed funding, that they are able to comply with security and confidentiality requirements, that the planned research activities are aimed at detecting, identifying and understanding systemic risks in the EU and will be conducted in a reasonable timeframe, and that they will make their research results available to the public free of charge (Regulation 2065:2022).

The following research questions on monitoring the way platforms provide access to data for research purposes arise from the provisions of the CoP and the DSA:

RQ1: How do platforms report on access to data for research purposes?

RQ2: Do researchers have access to public data for research purposes and what limitations they experienced?

RQ3: How do researchers view the process for requesting access to data and how do they assess its transparency?

RQ4: Are researchers from CRO and SLO treated unequally or unfairly when applying for access to data?

RQ5: Have researchers experienced access to data being withdrawn after it was originally granted?

METHODOLOGY

A combination of methods was used to produce this report: document analysis and survey. The document analysis was conducted for the second round of reports submitted by the Meta, Microsoft, Google and TikTok in July 2023 and X in January 2023 concerning access to signatories' data for research defined in CoP (Commitments 26 – 28). The aim was to find out how many researchers from Croatia and Slovenia had access to the platform data.

To verify the data provided by the platforms and their localization to Slovenia and Croatia, an online questionnaire was used and the sample was collected by snowballing. The initial sample was compiled by searching Slovenian and Croatian research information systems (CRORIS & COBISS)[1]. A total of 36 researchers (N=36) were contacted who had previously used data from platforms for different research and asked to fill the form and share it with their colleagues. In the end, 12 researchers (N=12) completed the questionnaire, all of whom are academics from Slovenia (N=7) and Croatia (N=5).

As the number of researchers working on disinformation in the SEE region is relatively small, a broader approach in selecting respondents was taken to measure general access to public data for research purposes (not just for disinformation research). Furthermore, the results are given in relative and raw figures.

RESULTS

ACCESS IN REPORTS

Meta reports that access to data is provided via the Crowd Tangle and Meta Content Library and API for Facebook and Instagram, but without quantitative data on the users of these tools (Meta, 2023).

"We cannot provide metrics regarding the use of our new Content Library and Content Library API for this reporting period as the tools are being progressively rolled out since the end of June 2023." (Meta, 2023)

---

[1] Prof. Blaž Rodič helped compile the initial list for Slovenia and Ivana Grkeš collected the emails used to send the surveys to the researchers.

However, they reported that more than 70 researchers worldwide have had access to the Ads Targeting API since it was released in September 2022, and that more than 250 researchers worldwide have had access to the URL Shares dataset since it was released in February 2020. (Meta, 2023)

Microsoft provides access to data from LinkedIn through API access and Bing search results through previously compiled collections such as MS MARCO (Microsoft, 2023). No figures were published on the number of researchers using these tools.

Google reported that it provides researchers with access to data through Google Trends, Google Fact Check Explorer and the YouTube Researcher Program (Google, 2023). Instead of data on the number of researchers, the number of regular users of the above tools was given.

Table 1: The number of regular users of Google tools (Google, 2023)

Country	Number of Fact Check Explorer tool users	Number of Google Trends users researching Google Search	Number of Google Trends users researching YouTube
Croatia	167	22617	948
Slovenia	70	12009	602

Google reports that 25 researchers from the EU have been approved to use the YouTube Research API, while no researchers from CRO and SLO have applied for access to the YouTube Researcher Program in the first half of 2023 (Google, 2023).

TikTok is working on expanding the Research API to the EU and granting access to the Commercial Content API without providing a metric for access (TikTok, 2023).

As X only submitted a report in January 2023, stating that researchers could request access for different levels of API access at that time (Twitter, 2023). Unfortunately, access was shut down in February 2023 and a monthly subscription was introduced.

EXPERIENCES OF RESEARCHERS

The researchers stated that they used data from Meta (N=9), X (N=8), Google (N=5), TikTok (N=2), Reddit (N=2) and Microsoft (N=1) for their research activities. They requested access to data for research purposes from the following platforms: X (N=7), Meta (N=7), Reddit (N=2), Google (N=1), and Microsoft (N=1). The discrepancy between using the data and formally requesting access could be explained by the use of commercial services to access the platforms' data or/and scraping data without using the tools/APIs provided by the platforms and formal approval. Researchers from Croatia and Slovenia have never applied for access to TikTok data because they have never had the opportunity to do so, as TikTok has only just started to provide access for EU researchers (TikTok, 2023).

Four researchers (N=4) requested "access to Meta data "4 or more" times, but only one (N=1) was granted access, while the other three (N=3) were not approved. Three researchers (N=3) requested access to Meta's data twice and were granted access, while one (N=1) requested access once and was not granted access. On the other hand, Google, Microsoft and Reddit granted access to all researchers who requested access, while Twitter granted access to five researchers (N=5).

On average, the approval process took "more than 8 weeks" for five researchers (N=5), while six researchers waited "2-4 weeks" (N=3) and "less than a week" (N=3). One researcher (N=1) reported that it took "1-2 weeks" to gain access.

Six researchers (N=6) reported that access to data was denied without explanation, including three (N=3) who reported that they never received a response from the platforms.

Seven researchers (N=7) reported that accessing the data from X and Meta was "difficult" or "very difficult" for them.

Table 2: Experience with accessing (using) data from platforms for research purposes[2]

Platform	Very easy	Easy	Neutral	Difficult	Very difficult
X	0	1	4	4	3
Meta	0	1	4	3	4
TikTok	0	0	6	1	1
Google	0	1	5	1	1
Microsoft	0	1	5	1	1

[2] The total for all platforms is not 12, as the question was not compulsory.

Some of the researcher's explanations were:

> *We cannot collect data from Twitter any more. It was very hard to collect datasets and now it is impossible.*

> *Limits in amount of data per month and speed of transfer. Many errors in accessing the data source (API) resulting in failed transfers.*

> *Unclear API and data terms of use. Difficulty getting approved. Restricted access to APIs.*

The transparency of the vetting process in terms of the information required and the criteria for approval was rated by the researchers as "transparent" (N=2), "neutral" (N=5), "not transparent" (N=1) and "not transparent at all" (N=4). Five researchers (N=5) felt that researchers in Croatia and Slovenia were treated unequally or unfairly when requesting data access compared to other regions/countries. Some of the researchers' explanations were:

> *We did not received answers when we tried to get permissions to collect the data.*

> *Not receiving any answer from Meta seems as a disparity to me.*

> *Much harder to obtain Slovenian data than e.g. English with simple, popular, robust approaches.*

> *Some data is only accessible to top universities, mostly from the UK and the USA. Usually, they work together with researchers working for platforms on some research.*

Five researchers (N=5) had their access to platform data withdrawn after initially being granted access, which negatively impacted the progress or completion of their research project.

CONCLUSION

Access to data on platforms is crucial for the scientific community, regardless of the topic they are researching, because platforms play an important role in contemporary life. However, as the reports show, access to platform data for researchers in the EU is still in its early stages and there are discrepancies between what the platforms claim and what researchers experience. TikTok, for example, reported that

it is working on extending the Research API to the EU, which means that researchers from EU countries do not have access to their data.

Researchers in Croatia and Slovenia mainly used data from Meta, X and Google, and in some cases this access was not formal. Researchers had to apply several times before they were granted access, later their access was withdrawn and they had difficulties using it. Many of them find the explanation for this in the unequal or unfair relationship between the platforms and the researchers in Croatia and Slovenia.

The collaboration between platforms and the EDMO Working Group on Platform to Researchers will ensure the establishment of an independent intermediary body to support research on digital platforms, facilitating the process of granting access to data from platforms by researchers in a transparent and fair way and developing a code of conduct for both platforms and researchers.

REFERENCES

CoP (2022). *The Strengthened Code of Practice on Disinformation 2022.* https://disinfocode.eu/wp-content/uploads/2023/01/The-Strengthened-Code-of-Practice-on-Disinformation-2022.pdf

Google (July 2023). *Code of Practice on Disinformation – Report of Google for the period of 1 January 2023 - 30 June 2023*. https://disinfocode.eu/reports-archive/?years=2023

Meta (July 2023). *Code of Practice on Disinformation – Report of Meta for the period 01 January 2023 to 30 June 2023*. https://disinfocode.eu/reports-archive/?years=2023

Microsoft (july 2023). *Code of Practice on Disinformation Report of Microsoft for the period 1 January – 30 June 2023*. https://disinfocode.eu/reports-archive/?years=2023

Regulation 2065:2022. *Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act)* https://eur-lex.europa.eu/eli/reg/2022/2065/oj

TikTok (July 2023). *Code of Practice on Disinformation – Report of TikTok for the period 1 January 2023 - 30 June 2023*. https://disinfocode.eu/reports-archive/?years=2023

Twitter (January 2023). *Code of Practice on Disinformation – Report of Twitter for the period H2 2022.* https://disinfocode.eu/reports-archive/?years=2023

ABOUT AUTHOR

Mato Brautović is a professor at the Department of Mass Communication at the University of Dubrovnik. He is an author and co-author of many books, book chapters, and scientific papers in professional journals in the field of his expertise – media and war, journalism, internet history, online disinformation, digital methods, and social media. Professor Brautović has a long and rich experience with scraping, cleaning and analyzing big data, going back to 2006, when he conducted his first social network analysis (SNA).

ADMO

Project Adria Digital Media Observatory (ADMO) ensures the establishment and continuation of the Croatian and Slovenian hub operation for tackling disinformation and cooperation with EDMO on multiple levels by a multidisciplinary and intersectoral approach. As an effective response to the challenges for democracy and society, the project sees a combination of different expertise and methods - attacking sources/actors, networks, reasoning, and disinformation narratives from different spheres and combining techniques. The implemented activities will contribute to the general objective - creating and ensuring the long-term development of a broader Croatian/ Slovenian hub on digital media composed of a multidisciplinary community of academic researchers, fact-checkers and media practitioners, journalists and civil society advocators capable of detecting and analyzing disinformation campaigns; and supporting the government and regulator in strengthening the collective response to disinformation threat to democracy.

The project comprehensively approaches the social, political, and health problem of spreading disinformation campaigns through research, awareness-raising, and capacity-building methods. The key projects' outcomes/deliverables are: the project's website and its integration with the EDMO platform, production of the continuous flow of fact-checks, publication of analytic reports on disinformation content, disinformation sources, origins of disinformation narratives, the role of international actors in spreading disinformation, disinformation acceptance and debunking, on the reasoning behind conspiracist and anti-conspiracist behavior, training for journalists and student of journalism, media literacy education activities for citizens, reports on Code of Practice on Disinformation in Croatia and on the financial viability of the Croatian and Slovenian news media sector.

SVEUČILIŠTE
U DUBROVNIKU
UNIVERSITY
OF DUBROVNIK

fpzg
Sveučilište u Zagrebu
Fakultet političkih
znanosti

FER
Faculty of Electrical
Engineering and
Computing

Filozofski fakultet
Sveučilišta u Zagrebu
FF

GONG

oštro

N1
EXCLUSIVE NEWS
CHANNEL AFFILIATE

WIKI

AFP

EDMO
European Digital Media Observatory