

# Holding platforms accountable: Issues and strategies for tracking structural indicators under the Code of Practice on Disinformation

by Teresa Weikmann, Marina Tulin, Michael Hameleers & Claes de Vreese

The [2022 Strengthened Code of Practice on Disinformation](#) is a self-regulatory framework developed collaboratively by industry stakeholders with support and guidance from the European Commission. Alongside fact-checking services, several Very Large Online Platforms and Search Engines (VLOPSEs) – such as Meta, Google, LinkedIn and TikTok – have signed onto the Code, committing, among others, “to support good faith research into disinformation that involves their services” (European Commission, 2022, p.29). Notably, X (formerly Twitter) withdrew from the Code in 2023.

Since the Code’s implementation, a major challenge has been finding reliable and valid ways to measure its effectiveness. A central aim of **BENEDMO 2.0** is to help evaluate whether platform signatories are meeting their responsibilities under the Code, particularly in effectively curbing the spread of disinformation. Central to this is **Commitment 41**, which requires signatories to provide data on specific [structural indicators](#) – comparable and standardized metrics designed to evaluate the tangible effectiveness of platforms’ anti-disinformation measures. These indicators were proposed by specialized [expert groups of the European Digital Media Observatory \(EDMO\)](#) and include (Nenadic et al., 2023, 2024):

1. **Prevalence of Disinformation:** Tracking the extent of disinformation on each platform
2. **Sources of Disinformation:** Identifying the origins and spread of disinformation
3. **Audiences of Disinformation:** Measuring the reach and demographics of disinformation
4. **Demonetization of Disinformation:** Evaluating actions taken to reduce financial incentives for disinformation
5. **Collaboration and Investments in Fact-Checking:** Assessing platforms’ investments in fact-checking and partnerships with fact-checking organizations

*Structural indicators* aim at measuring the absolute numbers and sheer quantitative impact of disinformation and the signatories’ efforts to combat it. However, monitoring



*This project has received co-funding from the European Union  
under Grant Agreement number  
101158277-BENEDMO-DIGITAL-2023-DEPLOY-04*

efforts have, at the time of writing this post (December 2024), mainly looked into *service-level indicators*; these assess whether promised actions were indeed taken, and if they are qualitatively valuable, thus focusing on processes, compliance, and effort by VLOPSEs. In the [Transparency Center](#) of the Code, platform signatories submit bi-annual *transparency reports* detailing their progress in fulfilling the commitments they make to serving the research community, fact-checkers and media literacy experts with their actions. Accordingly, transparency reports are a key resource for EDMO working groups monitoring platform accountability. For example, the *Monitoring Exercise for Assessing VLOPSEs' Transparency Reports* engages experts from local EDMO hubs to assess these reports on a national level, focusing on service-level indicators such as enhancing media literacy, equipping users to identify disinformation and empowering the research community. This exercise offers a preliminary assessment of the quality and completeness of the reports, serving as a first step in evaluating usability. Similarly, EDMO's *Working Group on Transparency of Political Advertising* evaluates platforms' labelling of political ads and the prevalence of disinformation within them. While transparency reports provide some data, e.g., the number of ads labelled or removed, they do not provide details regarding the kind of false or harmful content removed ads contain or examples thereof, limiting insight into disinformation's scope in political advertising. Therefore, this analysis requires supplementation with direct platform ad library reviews. Like the *Monitoring Exercise*, this initiative does not directly assess structural indicators, which is the primary interest for BENEDMO.

### **Key challenges with tracking structural indicators**

While EDMO's list of structural indicators comprehensively addresses the effectiveness of platform actions, it also encompasses broad categories of complex, hard-to-measure variables. Measurement of these indicators remains challenging due to two key issues: (1) the (efficient and comparable) usability of VLOPs self-reports and (2) access to platform data through APIs. First, platforms must not only collect but also share detailed information about the content they identify as disinformation and any actions taken to address it, such as removal or demotion. Second, for these measurements to be reliable, researchers need independent access to platform data, ideally through APIs or similar tools, to verify the accuracy of reported metrics. Currently, both aspects – platform data-sharing and independent access – fall short, as self-reports are incomprehensive and API access is often difficult to obtain (Ginsborg et al., 2024; Nenadic et al., 2023). As a result, the effectiveness of the Code's accountability framework is compromised, highlighting the need for stronger mechanisms for data transparency and third-party verification.



### *Usability of transparency reports*

Transparency reports in their current form are extensive but lack clarity and potentially suffer from reliability issues due to the reliance on self-reported assessments of risks and threats by platforms themselves, making it difficult to extract key details or compare efforts across platforms. For instance, TikTok's most recent [transparency report](#) from September 2024 spans 301 pages, listing efforts to comply with commitments. However, some fields, such as those related to regulating political advertising, are left blank, as the platform "does not allow political advertising" (TikTok, 2024). Notably, there is evidence, including from EDMO research, that political ads can in fact be found on TikTok (Murphy, 2024). In contrast, Google's 299-page [report](#) provides more details on actions taken to regulate political ads, but lacks comparable metrics, such as the proportion of ads removed and the specific reasons for their removal. Furthermore, the reports primarily report on *service-level indicators* – essentially confirming whether a commitment has been met, without providing insights into the impact of these efforts at the EU-wide or member state level. Specifically, the transparency reports do not give numeric information on the prevalence, sources, audience or demonetization of disinformation on platforms, especially *in proportion* to other information (i.e., information that is not disinformation) that can be found on the platforms. Moreover, platforms may indicate in their transparency report that they invest in fact-checking services, but we lack information on how much money is invested per member state and whether this investment has an actual impact. In conclusion, transparency reports are currently not usable to assess VLOPSEs' compliance along the structural indicators, as they are mainly geared towards assessing service-level indicators. This, in turn, makes it infeasible for BENEDMO to assess structural indicators based on transparency reports.

### *Limited API-access*

Access to data from social media platforms through APIs is an ongoing obstacle for the academic community, an issue that has seen little improvement under the Code of Practice on Disinformation or the EU's Digital Services Act (Ginsborg et al., 2024). This limitation seriously impedes the assessment of structural indicators, as it restricts independent, third-party verification of platform actions. However, the latest EDMO report on structural indicators (Nenadic et al., 2024) proposes a methodology that would only be feasible when API access is granted. In addition, even if API-access is granted, effectively utilizing such data requires expertise in computational methods and robust technical infrastructure – skills and resources that are not yet fully established across all EDMO hubs. Emphasizing the importance of structural indicators across EDMO will be crucial, as it can enable hubs to expand their capabilities, including hiring specialized expertise.



As an initial step, the European Commission has contracted independent researchers to do research on structural indicators, who have turned to alternative methods to analyse platform data. For instance, [TrustLab](#) – an independent organization contracted by the European Commission – conducted a pilot study on structural indicators using third-party data brokers to analyse disinformation prevalence and sources across VLOPs in Poland, Slovakia, Spain, and France (TrustLab, 2023). However, it was only meant to serve as an initial assessment, as it neither covers all structural indicators nor all EU member states. Another more recent initiative is undertaken by the non-profit [Science Feedback](#), who are funded by the European Media and Information Fund. Their aim is to measure prevalence, sources and monetisation of disinformation on major VLOPs in France, Poland, Spain and Slovakia. This project was only launched in October 2024, and their methodological approach and results are not known yet.

### **Moving forward**

At the time of writing this at the end of 2024, research on structural indicators is in its very early stages across EDMO hubs. However, this is likely to change in the upcoming year.

**BENEDMO** has committed to increasing efforts to address these current gaps, and other EDMO hubs have plans to follow. For now, we suggest the following steps to improve research on platform accountability, which is in line with previous EDMO-research (Nenadic et al., 2023, 2024):

1. **Revise transparency reports with detailed structural metrics:** Platform transparency reports should include a standardized appendix with specific metrics on structural indicators, allowing for comparability across platforms. Rather than merely indicating if actions were taken (service-level indicators), reports should include relative figures. For instance, as suggested by Nenadic et al. (2024), platforms could provide estimates on disinformation prevalence in a random sample of public content, weighted by views (e.g., 10,000 views, adjusted by member state population size), along with reach and engagement metrics. This would make assessments of compliance with the Code of Practice based on transparency reports easier and more efficient.
2. **Radically improve access to platform data:** Enhanced API access is essential for researchers to analyze the origins of disinformation on platforms. Crucial in achieving this will be the currently discussed access to online platform data for vetted researchers under the [Digital Services Act \(DSA\)](#). Expertise in processing such data has to be enhanced across EDMO hubs, for instance by bringing on board more computational researchers that have experience in doing such analysis. Recognizing the complexity of this issue, we also recommend using alternative methods for data access, such as using data donations or other tools. For example,



BENEDMO partner *Textgain* recently launched [Eerste Hulp bij Twijfel](#) (English: “First Aid when in Doubt”) a tool that collects social media discussions in a Flemish/Dutch context potentially containing disinformation. Such tools can provide valuable insights into structural indicators and help hold platforms accountable at member state level.

In sum, monitoring platforms based on structural indicators, rather than just service-level indicators, is crucial for holding platforms accountable. This will become increasingly important not only for BENEDMO, but for all EDMO hubs. To achieve this, we need a strong collaboration among hubs to develop and refine effective assessment methodologies, conduct joint monitoring exercises, and address challenges such as hard-to-measure categories and data gaps. By taking these steps, we can enhance accountability, gain deeper insights into the effectiveness of platform interventions, and significantly advance efforts to reduce disinformation across the EU.

## References

- European Commission. (2022). *The Strengthened Code of Practice on Disinformation 2022*.
- Ginsborg, L., Bontcheva, K., Châtelet, V., Darius, P., Motyl, M., & Neumeier, A. (2024). *Report on EDMO Workshop on Platform Data Access for Researchers*.
- Murphy, S. (2024). *TikTok's Struggle with Political Advertising*.  
<https://edmohub.ie/index.php/tiktoks-struggle-with-political-advertising/>
- Nenadic, I., Brogi, E., & Bleyer-Simon, K. (2023). *Structural indicators to assess effectiveness of the EU's Code of Practice on Disinformation*.
- Nenadic, I., Brogi, E., Bleyer-Simon, K., & Reviglio, U. (2024). *Structural Indicators of the Code of Practice on Disinformation: The 2nd EDMO report*.
- TrustLab. (2023). *A Comparative Analysis of the Prevalence of Misinformation and Sources of Disinformation across Major Social Media Platforms in Poland, Slovakia, Spain, and France*.

