



# Exploring global challenges of regulating researcher access to platform data

June 2024



# Contents

<b>Regulating researcher access to platform data: challenges beyond the EU</b>	<b>2</b>
Introduction	3
Background	4
Approach	5
Insights	5
Global efforts to mandate researcher access: nuances of the 'Brussels Effect'	5
Researchers are not a regulatory priority	6
The challenges of regulatory implementation	7
Regulation is not a panacea	7
The DSA may both help and hinder other regions, exacerbating existing structural inequalities	8
Conclusion	9
Acknowledgements	9
Appendix A: Interview participants	10

## About

The Open Data Institute (ODI) was founded in 2012 by Sir Tim Berners-Lee and Sir Nigel Shadbolt, placing us at the heart of the data economy. Our primary mission is to foster transparency, accountability and innovation through data. We are independent of government, non-partisan and non-profit, making us the trusted delivery partner for corporate, public and civil society organisations. We are funded through our market-leading commercial activities and grants from philanthropic and government bodies worldwide.

The report is based on document research and supported by insight and perspective from focused conversations with regional platform data access experts from around the world. We would like to thank them for their generous help and insightful perspective.

This report was researched and produced by the ODI. It was written by Sophia Worth and Claudine Tinsman, with support from Jared Keller, Sasha Moriniere and Gefion Thuermer. If you want to share feedback by email or to get in touch, contact Jared Keller at [jared.keller@theodi.org](mailto:jared.keller@theodi.org).

# Regulating researcher access to platform data: challenges beyond the EU

*Researchers face a complex and shifting landscape in accessing social media data. Despite the vast social importance of large social media platforms, data access has traditionally been through APIs, web-scraping, and informal methods, due to a lack of legal mandates. In the EU, the recently implemented Digital Services Act (DSA) requires platforms to provide data access for researchers under specific conditions, aiming to safeguard users' rights and address societal risks.<sup>1 2</sup> Globally, however, researchers continue to navigate challenges in accessing such data. To better understand these difficulties, we conducted 11 exploratory interviews with platform data access experts worldwide about the challenges of platform data access in their respective regions. We present these high-level findings in this exploratory report.*

---

<sup>1</sup> The Digital Services Act (DSA) (2022), [Article 34, Risk Assessment](#).

<sup>2</sup> The Digital Services Act (DSA) (2022), [Article 40, Data access and scrutiny](#).

# Introduction

Researchers who need social media data operate in a very uncertain and shifting landscape.<sup>3</sup> The vast social significance of these platforms is evident: Meta had 2.9 billion monthly active users in 2023 and more than half of the global population now uses social media<sup>4</sup>. This impacts everything from political movements and elections to public health, wellbeing and far beyond. Researchers all over the world have expressed their enthusiasm for studying these online systems, but since data is primarily in the hands of the private companies that own them, there are complicated dynamics for access, and very different contexts for this access across the world.

In the vast majority of the world, researcher access to platform data is not mandated by law, and researchers therefore must resort to a range of methods to access such data. Therefore, researchers have done so via Application Programming Interfaces (APIs) and portals, or other informal approaches such as web-scraping or data donation.<sup>5</sup> However, the EU Digital Services Act (DSA) regulation is breaking ground in requiring that in specific circumstances, researchers must be provided access to certain data held by social media companies.

The DSA introduced a framework that governs the duties of online intermediaries, including platforms such as social media, search engines and online marketplaces. Its purpose is to foster a more secure digital environment that safeguards the fundamental rights of all users. The DSA defines specific due diligence requirements for platforms based on their function, size and influence in the digital realm. Under Article 40 of the DSA, Very Large Online Platforms (VLOPs) such as Meta must give researchers access to the data they hold to study, identify and understand any systemic risks to the EU.<sup>6</sup>

In the rest of the world, access to social media data remains largely unregulated. So, in the absence of specific legislation mandating access to data, how do researchers access platform data? To address this question, we conducted 11 exploratory interviews with experts around the globe about researcher access to platform data and explored how researchers collect data and the challenges they face.

---

<sup>3</sup> Davidson et al. (2023), '[Platform-controlled social media APIs threaten open science](#)'.

<sup>4</sup> Wong, B. and Bottorff, C. (2023), '[Top Social Media Statistics And Trends Of 2024](#)'.

<sup>5</sup> Davidson et al. (2023), '[Platform-controlled social media APIs threaten open science](#)'.

<sup>6</sup> The Digital Services Act (DSA) (2022), [Article 40, Data access and scrutiny](#).

# Background

Until 2023, APIs were particularly useful to researchers looking into a range of socially important topics, such as disinformation or online harm.<sup>7</sup> However, over the course of 2023, several social media platforms, such as X (formerly Twitter), TikTok and Reddit, severely restricted data availability and/or made access more costly by increasing users' fees.<sup>8</sup> As a result, collecting such data has become considerably more challenging for researchers, forcing them to rely on self-reported platform data such as transparency reports, or risk being legally pursued for breach of contract by collecting data through means that may violate platform terms and conditions, such as scraping.<sup>9 10</sup>

This wave of changes has forced researchers to abandon their projects or obtain data through unofficial channels, thereby restricting their ability to examine crucial questions requiring platform data.<sup>11</sup> To address the instability of this access, Article 40 of the DSA mandates

that 'providers of very large online platforms...provide access to data to vetted researchers...for the sole purpose of conducting research that contributes to the detection, identification and understanding of systemic risks in the Union'.<sup>12</sup>

However, the act is still unclear about what the vetting process entails in practical terms and the conditions for using the data, which leaves a great deal of room for interpretation by platforms.<sup>13</sup>

Globally, researcher access to data access is ad hoc, requiring researchers either to request access from platforms directly, access data using third-party tools, or purchase data from expensive data brokers.

---

<sup>7</sup> Davidson et al. (2023), '[Platform-controlled social media APIs threaten open science](#)'.

<sup>8</sup> Moriniere, S. (2023), '[We must fix researcher access to data held by social media platforms](#)'.

<sup>9</sup> Ibid.

<sup>10</sup> Gerken, T. (2024), '[Elon Musk's X anti-hate group case thrown out](#)'.

<sup>11</sup> Davidson et al. (2023), '[Platform-controlled social media APIs threaten open science](#)'.

<sup>12</sup> The Digital Services Act (DSA) (2022), [Article 40, Data access and scrutiny](#).

<sup>13</sup> Brown, M.A., Lukito, J. and Yang, KC. (2024),

# Approach

We do not yet know the implications of the DSA's implementation. To understand the specific challenges for data access around the world, we undertook 11 interviews with academics and civil society stakeholders from the UK, the US, China, India, Latin America, Africa and the EU (see Appendix) to identify shared and different challenges across several different national/regional contexts. To better understand the wide range of contexts and provide additional context to our interviews, we gathered relevant resources, including blogs, articles and information about legislative proposals in these regions.

## Insights

The EU is not the only region making efforts to improve researcher access. Legislative proposals have been introduced at various points in the US, the UK and Brazil (to name but a few) to support researcher access to platform data, though the DSA is the sole piece of ratified and operational legislation to address the problem. In this next section, we share three key insights from our conversations with regional experts and a worldwide literature review on this topic.

### **Global efforts to mandate researcher access: nuances of the 'Brussels Effect'**

During our interviews with experts, we found that many countries are tracking the DSA's implementation to observe its benefits and disadvantages. Several experts interviewed pointed out that it benefits from a 'Brussels Effect' of Article 40.

The term 'Brussels Effect' refers to the phenomenon whereby the EU influences global markets and sets global standards.<sup>14</sup> Because the EU is such a large market, multinational companies often find it more practical to implement these standards globally rather than tailor their practices to different regions when it sets high regulatory standards. Consequently, the EU's regulations are becoming the de facto global standard.

---

<sup>14</sup> Bradford, A. (2012), ['The Brussels Effect'](#).

Experts pointed out that under the DSA, platforms will have to create data access channels for vetted EU researchers, and could extend those processes to researchers from the rest of the world. However, there was consensus from all of our experts that platforms were unlikely to do so, due to their concern that extensive analysis of platform data may damage their public image and adversely affect their market value.<sup>15</sup>

Though several interviewees viewed Article 40 as a source of inspiration for platform regulation proposals worldwide, it cannot be transposed wholesale to other countries and regions with different social contexts, regulatory bodies and infrastructure. For instance, the DSA's implementation relies on independent regulators and a parliamentary commission presiding over the implementation of the law – which is mostly not present in Latin America<sup>16</sup> or Africa.<sup>17</sup>

## Researchers are not a regulatory priority

In recent years, several regulatory efforts have been underway across the globe to enable researcher access. For instance, the Platform Accountability and Transparency Act (PATA) and Kids Online Safety Act (KOSA) have both been introduced to the US Senate several times, but neither has yet been debated in the wider legislature. In Brazil, the latest iteration of the proposed 'Fake News Bill' includes provisions for researcher access to platform data for misinformation and disinformation research.<sup>18</sup> In the UK, amendments mandating researcher access to platform data were proposed during deliberations on the Online Safety Act (OSA) but were ultimately tabled.<sup>19</sup> Instead, communications regulator Ofcom has been tasked with producing a report determining how much access researchers currently have, and if it should be mandated in future legislation.<sup>20</sup>

---

<sup>15</sup> Research ICT Africa (2023), '[Towards an African alliance for meaningful access to Intermediaries' data holdings](#)'.

<sup>16</sup> Brito Cruz et al. (2023), '[In the first interview of the series, Agustina Del Campo talks about access to data from digital platforms for research](#)'.

<sup>17</sup> Research ICT Africa (2023), '[Towards an African alliance for meaningful access to Intermediaries' data holdings](#)'.

<sup>18</sup> Brito Cruz et al. (2023), '[In the first interview of the series, Agustina Del Campo talks about access to data from digital platforms for research](#)'.

<sup>19</sup> Havilicek, S. (2023), '[Why the Online Safety Bill must include provisions for data access](#)'.

<sup>20</sup> Online Safety Act (2023), [Section 162](#).

These efforts share a common attribute of not being passed into law, at least not for now. While there are undoubtedly specific political factors in each context, interviewees indicated that, in the landscape of social media concerns, mandating researcher access has been considered less urgent than other matters, such as implementing online safety measures or instituting regulatory oversight for platforms.

## The challenges of regulatory implementation

Questions remain under the DSA, and Article 40 in particular, that have yet to be resolved. The Act does not clearly explain how platforms should decide if a research question pertains to ‘systemic risks’ under Article 40; nor does it provide guidance for sharing data with researchers. The DSA does not provide any further guidance on whether the data shared to address ‘systemic risks’ may include information about users outside the EU, enabling researchers in wealthy Global North countries to access population data that is currently out of bounds to them.

The lack of clarity on these details will likely lead to different interpretations about what data should be given to whom, and how it should be made available.<sup>21</sup> Furthermore, experts expressed concern about the ambiguity over whether the DSA would be enforced, pointing to the GDPR as an example of the difficulties EU regulators face when implementing sweeping changes to the union’s regulatory regime.

## Regulation is not a panacea

One point that came across strongly from our interviews is that specific local political, economic and cultural contexts shape the challenges that researchers face in accessing platform data in specific countries or regions. Meanwhile, experts cited that researchers in parts of Africa and Latin America face significant challenges in accessing data that regulation alone cannot fix. As large platforms do not provide API access for researchers in these locations, researchers resort to purchasing data from expensive brokers or through limited scraping of public platform data.<sup>22</sup>

---

<sup>21</sup> Miller, G. (2024), ‘[The Digital Services Act Is Fully In Effect, But Many Questions Remain](#)’.

<sup>22</sup> Research ICT Africa. (2023), ‘[Towards an African alliance for meaningful access to Intermediaries’ data holdings](#)’.

Furthermore, experts pointed out that some researchers in African and Latin American contexts (Kenya, for example), lack regional points of contact to communicate with regarding access to data, and lack the legal authority over platforms that is present largely in the US, where the social media companies are based. Researchers find themselves in a position of having to contact platform decision makers who do not have the requisite local legal knowledge of data protection to provide access. To circumvent this problem, researchers are sometimes forced to rely on research partnerships with organisations based in the Global North that have access to points of contact and greater financial resources to obtain data, which creates overdependence dynamics.

Finally, our experts indicated that when researchers are able to access platform data directly, they face a skills gap problem. Data collected directly from platforms is seldom provided in a ‘user-friendly’ format and requires substantial wrangling and reformatting to be usable, which involves a significant level of technical skill. However, researchers in the social sciences and from civil society do not necessarily have the technical expertise to do this and can find it difficult to partner with researchers in other fields who do.<sup>23</sup> In such cases, researchers find they have little choice but to fall back on existing datasets that other organisations have put together with greater resources, even if they do not contain all the necessary information.

## **The DSA may both help and hinder other regions, exacerbating existing structural inequalities**

As indicated previously, there are already significant differences between regions in terms of informal mechanisms of access to online platforms. Beyond acting as an inspiration for legislation elsewhere, the DSA may offer a pathway for insights into global challenges, given that the ‘systemic risks’ it investigates may be global. It remains an open question how far researchers from regions beyond Europe will be able to become ‘vetted researchers’ under the DSA.<sup>24</sup> Still, such dynamics perpetuate similar concerns, whereby researchers globally may be dependent on the ‘first mover’ nations with regard to regulation for access to data, replicating the current unbalanced dynamics.

---

<sup>23</sup> Ibid.

<sup>24</sup> Helberger & Samuelson. (2024). [‘The Digital Services Act as a Global Transparency Regime’](#).; Brito Cruz et al. (2023). [‘In the first interview of the series, Agustina Del Campo talks about access to data from digital platforms for research’](#).

## Conclusion

While legislation may be an important source of leverage to ensure access to platform data sources, it is also clear that the impacts of legislation on a global scale need to be considered and that such impacts are not felt equally, replicating the dynamics of inequality.<sup>25</sup> Furthermore, researcher access is clearly not a simple challenge and will require a suite of approaches across different environments. There was significant interest in promoting collaboration between researchers across different fields and regions to leverage the data access of vetted EU researchers for the global good. These findings indicate that further investigation into the legal, technical and practical means of such knowledge exchange may be valuable to researchers across different regional and global contexts.

## Acknowledgements

The report would not have been possible without the insight and perspective from interviews with platform data access experts worldwide, and we would like to thank them for their time and regional expertise. A complete list of experts and their affiliations is included in this report's appendix.

---

<sup>25</sup> Inverarity et al. (2023). [‘Power, ecology and diplomacy in critical data infrastructures’](#).

# Appendix A: Interview participants

Table 1: Organisations and regions of expertise

Country/Region of Expertise	Expert	Affiliation
Africa	Liz Orembo	ICT Research Africa
China	Jessica Batke	Chinafile
China	Dr Fen Gao	Open Data China
EU	Dr Jef Ausoloos	University of Amsterdam
EU	Noemie Krack	KU Leuven
India	Divij Joshi	University College London
India	Shashank Mohan, Shobhit Shukla, Tavishi Ahluwalia, Jhalak Kakkar	National Law University Delhi
Kenya	Odanga Madung	Mozilla
Latin America	Iná Jost	InternetLab
UK	Dylan Sparks	Reset
United States	Dr Brent Eastwood	R Street Institute