Coimisiún na Meán

Researcher Survey on
Article 40 Data Access

Publication Date: September 2025

Table of Contents

List of Figures

Figure 1: The number of planned data access request across VLOPSEs

Figure 2: Types of data researchers expect to request

Figure 3: Challenges for researcher data access

1.0 Introduction

Art. 40 of the Digital Services Act (DSA) establishes a framework that compels access to data from Very Large Online Platforms and Very Large Online Search Engines (VLOPSEs) to vetted researchers. Data accessed in this way is used for the purposes of investigating and evaluating systemic risks within the online environment in the EU. Under Art. 40(4) DSA, researchers who are approved by the relevant Digital Services Coordinator (DSC) of establishment are awarded 'vetted researcher' status and can obtain access to information not ordinarily available through standard platform access routes.

The DSC of establishment is responsible for assessing applications and requesting data access from VLOPSEs on behalf of vetted researchers who meet the criteria of Art. 40 DSA. As the DSC of establishment for 15 out of the 25 designated VLOPSEs under the DSA, Coimisiún na Meán will play a significant role in the vetted researcher data access process.

At Coimisiún na Meán, we have been working to develop our operational system for implementing Art. 40(4) DSA. Our focus has been on developing clear, effective processes for the rigorous assessment of applications, which respect rights and obligations of individuals, organisations and the data providers associated with the requested data. This work is led by a dedicated team within the Research & Strategy unit, staffed with academics with a strong understanding of the dynamic and complex challenges of data access, protection and safeguarding.

From March to May 2025, we conducted an online survey to explore researchers' needs, readiness, and barriers to data access under Art. 40 DSA. Using a snowball sampling method, the survey was shared through established research networks and direct outreach. Recipients were encouraged to circulate it further, resulting in 116 responses.

The survey was designed to gather insights from researchers about their needs and preparedness for the process of requesting data access under Art. 40(4) DSA, as well as understanding the current and potential challenges associated with this process. The survey also explored researcher demand, though this should be seen as indicative rather than definitive given the early stage of implementation. While not a formal scientific study, this research provided valuable context to help inform our internal processes and external guidance. Our goal was to listen directly to researchers to understand and better align our approach to application and processing of data access requests with their real-world experiences and needs.

We are sharing the survey results publicly to inform other stakeholders such as DSCs, VLOPSEs and the research community. By making this information accessible, we aim to support broader efforts to address researchers' needs and challenges during the operationalisation of the vetted researcher and data access process under Art. 40 DSA.

2.0 Survey Results

2.1 Profile of researchers

Most researchers who responded to the survey came from universities or Academic Institutions (82%), while 12% were from non-governmental organisations (NGOs). The remaining 6% were either independent researchers, private sector, or governmental agency employees.

EU-based research organisations accounted for 73% of the total respondents, while 27% were based outside the EU. The EU Member States most represented were Ireland (28%) and Germany (11%). Outside of the EU, the USA (13%) and UK (6%) saw the highest levels of representation. A further 8% of researchers came from other non-EU countries including India, Uganda, Brazil, New Zealand, and Yemen.

At 43%, social sciences was the most represented field of study among researchers, followed by computer and data science at 28%, and law at 10%. The remaining 19% comprised other fields including international relations, economics and natural sciences.

2.2 Choice of DSC for submission of application

Researchers applying for vetted researcher data access, who are based within the EU, have two routes to application. They can either apply directly to the DSC of establishment for the VLOPSE from which they are requesting data access or they can apply to the DSC where their research organisation is based, in which case this DSC will perform an initial assessment and forward this to the DSC of establishment.
In either case, the DSC of establishment is responsible for the final vetting decision.

We asked EU-based researchers which DSC they would prefer to submit their application to. Almost one third (28%) indicated that they would need more information to make a decision. 23% of respondents said they would submit to the DSC where their research organisation is based and 18% said they would submit directly to the DSC of establishment where the VLOPSE is located. Only 3% responded that they did not know.

We encourage EU-based researchers to apply to the DSC in the country where their research organisation is based for an initial assessment of their application. Some of the criteria applicants must meet, such as demonstrating commercial independence, disclosure of funding, or confirming whether the organisation qualifies as a research organisation under Art. 2(1) of Directive (EU) 2019/790, often require national-level assessment, which may be carried out more effectively by DSCs familiar with the local context.

Coimisiún na Meán are working closely with other DSCs to develop a harmonised assessment process. Given the high volume of applications expected for VLOPSEs established in Ireland, we believe that starting with a national-level initial assessment, where possible, will support a more efficient and timely review.

2.3 Organisational support and GDPR awareness

We asked researchers about the types of support services available at an organisational level to support development of their applications. Access to ICT/tech specialists and access to an ethics committee or review board were the most-selected answers (both at 85%,) while 74% of the researchers stated that legal/contractual services are available to them.

We asked about these support services as they represent areas which may prove difficult to navigate for more complex data access requests. For example, if a researcher is proposing to use their own organisation's IT infrastructure for data storage, they may wish to consult with their organisation's IT

services for support in ensuring that data security or confidentiality requirements, technical and organisational measures are met in accordance with Art. 40(8) DSA.

A data sharing agreement may be required to formalise the terms under which data is shared. In cases where the data contains commercially sensitive information, additional safeguards such as a Non-Disclosure Agreement (NDA) or other type of contractual agreements may also be necessary. In these instances, researchers may require support from their organisation's legal services. Research ethics is a cornerstone of good research practices, and it is expected that research which may pose an ethical risk is assessed and approved by an ethics committee.

Some researchers have expressed concerns related to limited access to support services in academic settings. We encourage both researchers and research organisations alike, to review their support services for researchers who wish to apply for Vetted Researcher status and to access data under Art. 40 DSA.

Technical Infrastructures

Part of the DSC of establishment's responsibility is to ensure that proposed methods for accessing data from VLOPSE's are appropriate to fulfil the requirements of data security, data confidentiality and protection of personal data. For more sensitive or high-risk data requests, this may require the use of an SPE (secure processing environment) rather than more accessible methods such as an API (Application Programming Interface) or direct file transfer. 39% of our researchers stated that they have access to an SPE.

Data Protection Compliance

Regarding data protection, 79% of researchers indicated that they have access to a Data Protection Officer (DPO) or similar, followed by 63% for data protection training. When asked about their level of familiarity with General Data Protection Regulation (GDPR) requirements, 57% said they were very familiar, 38% said they were somewhat familiar and only 5% stated they were not familiar with GDPR.

While not all applications will request access to personal data, applications that do, will require adherence to GDPR and clear demonstration of compliance. We strongly encourage researchers who are considering requesting access to personal data to undertake GDPR training, and to consult early with their research organisation's DPO to support the development of their application. This is of particular importance to researchers outside the EEA who may be less familiar with GDPR. More information about data protection law requirements can be found on the websites of the various Data Protection Commissions in the EU.

2.4 Experience with platform data

The majority (72%) of researchers who responded to this survey have accessed platform data before and 64% have previously published research based on social media or search data. When asked how they accessed this data, most (54%) stated API access, with other means including scraping (47%) and downloading datasets (15%) also mentioned.

There are some existing researcher data access programmes made available by VLOPSEs, particularly as part of their obligations under Art. 40(12) DSA. We asked researchers if they had any prior experience applying for data directly from a VLOPSE as part of an existing data access programme. The majority of researchers (72%, n=84) reported having accessed platform data before. Within this group, 68% (n=57) had not used an existing access programme, while 32% (n=27) had. Among those 27 researchers, only 7% (n=2) received all the data they requested, 33% (n=9) did not receive the data they requested, and 59% (n=16) reported receiving only part of it.

These findings highlight a key difference between the data access provisions of Art. 40(12) DSA and Art. 40(4) DSA. Under Art. 40(12) DSA, VLOPSEs have significantly more discretion over the application process including setting requirements and approval standards—which influences what data they ultimately

Coimisiún
na Meán

provide. In contrast, under Art. 40(4) DSA, the DSC assesses each application, and for each approved application issues specific reasoned requests that dictate precisely what data must be shared and the method of sharing.

We asked those who had experience with platform data for any insights or feedback they think would be useful for DSCs to consider in implementing this new data access process. Researchers emphasised the need for the process and application guidance to be user friendly, simplified and organised in a sequential, step-by-step format. They referenced poor data access application experiences, that were scattered with unclear or inconsistent criteria and documentation. Researchers also emphasised the need for an application system that is easy to use and provides clear communication about application progress.

2.5 Number of planned applications

Overall, 89% of respondents indicated that they expect to apply for data access under Art. 40(4) DSA.

In terms of timeline, 54% of researchers expected to submit an application within 3 months of applications opening, 19% within 3-6 months and 17% within 6-12 months.

TikTok, Facebook, Instagram, YouTube and X emerged as the VLOPSEs of greatest interest to researchers planning data access requests. It should be noted that the seven most requested VLOPSEs are all established in Ireland (see Fig. 1).

Under Art. 40(4) DSA, researchers must submit a separate data access application for each service they wish to request data from. An individual researcher may submit multiple applications—either for different platforms or for different research projects. It is notable that the 103 researchers who plan to apply within the first 12 months expect to submit a total of 602 individual applications. Of these, 448 are expected within the first six months, with 317 relating to VLOPSEs established in Ireland. (see Fig. 1).

The number of planned data access
requests across VLOPSE

Platform	Requests
TikTok	77
Facebook	76
Instagram	73
Youtube	71
X	69
LinkedIn	38
Google Search	30
Snapchat	26
Wikipedia	15
Amazon Store	13
Pornhub	12
Google Play	12
Apple Store	12
Pinterest	12
Bing	10
Google Maps	10
AliExpress	7
Zalando	6
Shein	6
Booking.com	6
Stripchat	5
XVideos	5
Temu	4
XNXX	4
Google Shopping	3
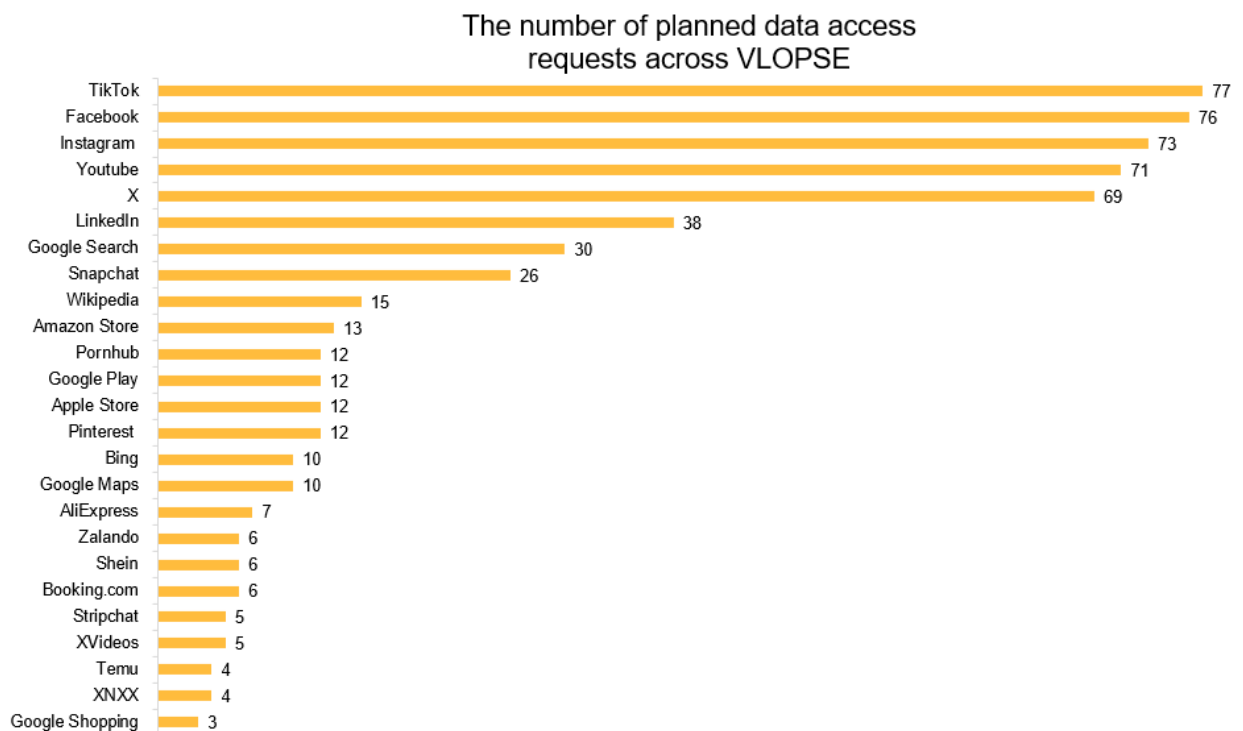
Fig. 1 The number of planned data access request across VLOPSEs

2.6 Planned data types and access modalities

Researchers were asked what type of data they would expect to request access to for their research (see Fig. 2).

The most requested data type, at 57%, was anonymous individual-level data such as user engagement histories or anonymised profiles of relationship networks. Regarding personal data, pseudonymised individual-level data is expected to be requested in 53% of cases, while directly identifiable personal data was lower at 34%.

These findings suggest positive awareness of data protection considerations among researchers, with over half planning to request access to anonymous data to mitigate privacy risks. This approach is significant because the GDPR should not apply to anonymised data and if the data is anonymised in this way, it would not be subject to the same processing restrictions. However, researchers will need a clear understanding of what constitutes "anonymous data" and should be prepared to describe their proposed anonymisation techniques. This ensures they do not inadvertently request access to data subject to the GDPR without implementing the necessary technical, organisational and legal safeguards.

Researchers may receive data which is pseudonymised data or may have to pseudonymise data which they will use for specific research purposes. Pseudonymisation as a safeguard may protect personally identifiable information, however it is important for researchers to remember pseudonymised data is considered personal data under GDPR and the processing of any personal data, even if pseudonymised, must be compliant and justified in every case. Researchers should be prepared to describe any pseudonymisation techniques they may employ as a safeguard in their data access application. Other important data protection considerations that researchers may be aware of include the principle of data

Coimisiún
na Meán

minimization which may underpin a vetted researcher data access request. Researchers should consider their needs and ensure their requests specify the least identifiable data possible, the minimum volume required, and the shortest retention period necessary to accomplish their research goals. Further details on data protection law requirements are available on the websites of the Data Protection Commissions across the EU.

The second most common type of data researchers expect to request is related to platform systems or governance at 55% and aggregated or summary statistical data at 45%.

Under the vetted researcher data access process different types of data may be requested under Art. 40(4) DSA and includes previously unavailable data types such as "data related to the testing of new features prior to their deployment, including the results of A/B tests; data related to content moderation and governance, such as data on algorithmic or other content moderation systems and processes, including changelogs, archives or repositories documenting moderated content, including accounts as well as data related to prices, quantities and characteristics of goods or services provided or intermediated by the data provider". These present new opportunities for advancing platform research. We expect researchers to consider measures like non-disclosure agreements (NDAs) and data sharing agreements, as proposed safeguards, particularly if the requested data contains commercially sensitive or confidential information.

Data relating to advertising and monetisation is expected to be requested in 36% of cases.

Types of data researchers expect to request

Category	%
Anonymised individual-level data, e.g. user engagement histories (likes, shares, or views), anonymised profiles of relationship networks, or interaction patterns that do not directly identify individuals.	57%
Data related to platform systems or governance, e.g., datasets on algorithmic moderation processes, A/B test results from new features, or archives documenting decisions about flagged content.	55%
Pseudonymised individual-level data, e.g. datasets with user activity linked by pseudonyms, enabling behaviour tracking without directly revealing identity; personalisation mechanisms used in content recommendations; or pseudonymised ad targeting data.	53%
Aggregated or summary data, e.g. aggregated statistics on flagged posts by region or high-level trends in content recommendations.	45%
Data related to advertising and monetisation, e.g. data on ad targeting and delivery mechanisms or platform-generated insights on advertiser segmentation.	36%
Data containing personal data, identifiable information or potentially identifiable, e.g. user-generated content such as posts, comments, or other interactions that could include details revealing identity indirectly, such as workplace, family connections	34%
I am unsure/have not fully considered what data I would require	10%

Fig. 2 Types of data researchers expect to request (multi-code)

Coimisiún
na Meán

Data Access Modalities

The term data access modality refers to the technical means by which a researcher would receive access to the data they request.

When asked which types of modalities researchers expect would be suitable for accessing data for their research purposes, the most common answer was API access at 67%, followed by a secure data transfer to their organisation's servers at 61% and access to file sharing via a shared folder at 43%. Interestingly, more secure access methods were anticipated to be required the least. 33% of researchers anticipated the use of a first party platform SPE and 33% expected to use a third party SPE.

With respect to both requested data and proposed modalities, researchers are expected to consider the level of risk and sensitivity associated with their request, including, but not limited to factors such as whether the data includes personal data or commercially sensitive information. For example, requests involving anonymous aggregated data could be considered lower risk than those involving personal data or commercially sensitive material. Researchers must demonstrate that they can meet the relevant data security and confidentiality requirements of their request. This should include a description and evidence of the technical, organisational, and legal measures they have in place to meet those requirements.

2.7 Research Purpose: Systemic risk

Under Art. 40(4) DSA, access to data by vetted researchers must serve the sole purpose of conducting research that contributes to:

- the detection, identification, and understanding of systemic risks in the Union (as outlined in Art. 34(1) DSA), and

- the assessment of the adequacy, efficiency, and impacts of systemic risk mitigation measures (under Art. 35 DSA).

While this survey did not ask for information outlining the systemic risk specific to each planned application, researchers were asked to choose the systemic risk as the primary focus of their planned research. The most common answer chosen was *negative effects of civic discourse, electoral processes and public security* (42%), followed by the *negative effects on the exercise of fundamental rights* (19%) and the *protection of public health and minors* at (8%). *Dissemination of illegal content* (7%), *gender-based violence* (6%) and the *serious negative consequences to personal physical or mental well-being* (3%) were other focus for research.

Art. 34(1) DSA provides more information about systemic risks as those arising from the design or functioning of a VLOPSE's service and related systems (including algorithmic systems), or from the use made of those services.

Vetted Research data access applications must demonstrate that the planned research is being carried out solely for the purpose of studying a systemic risk that is linked to the design, functioning, or use of the VLOPSE's service or its systems.

2.8 Funding of planned applications

Art. 40(8) DSA requires researchers to disclose the source of their research funding.

When asked about the funding in place for their planned data access applications, 16% of respondents had full funding in place, a further 16% had partial funding, while 20% said their funding was either in progress or was dependent upon the access to data. Twenty-one percent stated that they had no funding in place currently. Sources of funding mentioned by researchers included the EU, local government, philanthropic and academic funds.

Coimisiún na Meán recognises that academic funding can be complex and may include non-monetary contributions such as staff time, equipment, materials, or access to facilities. It is important that applicants are transparent about all forms of funding so that the DSC can assess whether the source has been fully disclosed.

2.9 Non-EU data and planned applications

Art. 40 DSA and the Delegated Act do not place geographic restrictions on the data that may be requested. When asked whether they expected to request access to non-EU data, 24% of researchers said they expect to do so, 36% said they would consider it, and only 27% said they did not plan to do so.

One reason given for requesting access to non-EU data was the need for comparative research on non-EU contexts, particularly to study issues such as disinformation or civic discourse related to EU politics. The strong interest in non-EU data reflects researchers' recognition of the cross-border nature of systemic risks.

While researchers may request access to non-EU data, their applications must also provide DSCs with a clear rationale explaining why such data is necessary to study systemic risk within the EU and any proposed safeguards necessary if international transfers of personal data are envisioned.
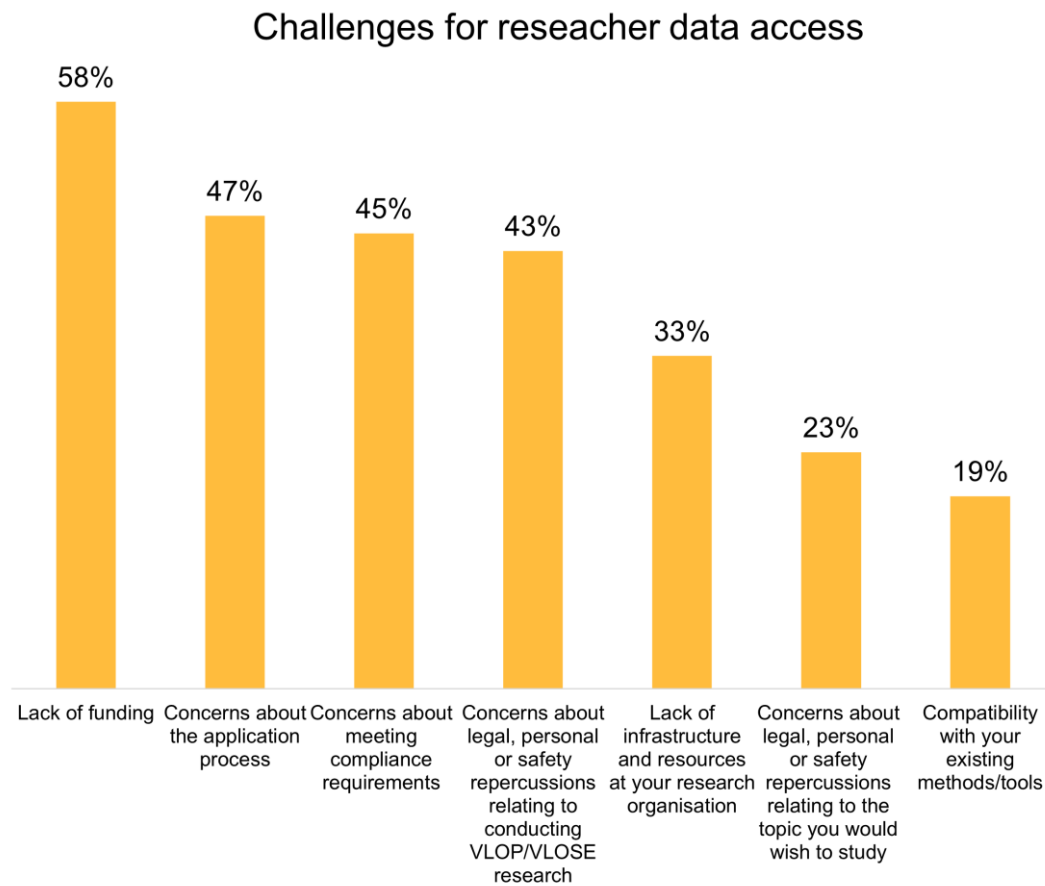
2.10 Challenges and barriers for researchers

Researchers were asked about any foreseen challenges or barriers which could affect their ability to apply for vetted researcher data access. Unsurprisingly, secure and stable funding was the most significant barrier identified by researchers at 58%, followed by concerns about the application process (47%) and meeting compliance requirements at (45%) (see Fig. 3).

Regarding the application process and compliance requirements, we understand that vetted researcher data access under Art. 40(4) DSA is complex and will require significant effort from researchers and their research organisations. An Coimisiún na Meán is committed to providing guidance that will help researchers understand both the requirements and evaluation criteria for assessment.

Concerns about legal, personal or safety repercussions relating to VLOPSE research was identified by 43% of respondents, while the same concerns relating to the topic of the research stands at 23%. This reflects the sentiment that conducting platform or online research has become more difficult in recent years. Lack of infrastructure and legal resources at an organisational level was highlighted as a barrier by 33% of researchers.

Challenges for reseacher data access

Fig. 3 Challenges for researcher data access

3.0 Conclusion

This survey was conducted to better understand the needs, challenges and expectations of researchers regarding vetted researcher data access under Art. 40 DSA. The information gathered has helped inform the development of processes and guidance with the objective of creating frameworks that are practical and responsive to real-world experiences. We sincerely thank all researchers who took the time to contribute their invaluable perspectives.

As the vetted researcher data access process is still at an early stage, and neither the Delegated Act nor researcher guidance was available during the survey period, researchers stated intentions may not fully reflect how they will make applications in practice. Therefore, while the survey provides valuable initial insights into demand and expectations, these findings should be interpreted with caution and revisited as the process develops.

We look forward to assessing applications in October 2025 within a vetted research data access framework that is both novel and unprecedented. As with all new undertakings, challenges may occur in the initial phase, however we are confident in our planned procedures, which will ensure consistency and reliability and allow us to respond effectively as the process unfolds. Coimisiún na Meán, along with our colleagues in other DSCs, are committed to implementing a vetted researcher data access process, that balances the rights and interests of all stakeholders, including researchers, VLOPSEs, and their users, while ensuring the system remains effective, fair, and robust as this important initiative evolves.