A W O

Data Protection and DSA Data Access for Platform Research

Summary Paper

April 2026

A W O

Table of Contents

1 Executive Summary

This summary paper provides practical guidance on data protection compliance for researchers seeking access to data from platforms under the EU's Digital Services Act (DSA).[1] It is particularly relevant for applicants using the Article 40(4) procedure of the DSA, which allows vetted researchers to obtain platform data – via Digital Services Coordinators – to investigate systemic risks stemming from the design or functioning of online platforms available in the EU.

Ironically, while Article 40(4) promises unprecedented opportunities for researcher access – at least in theory – the two years preceding its implementation have been marked by a steep decline in actual access. The data fortress inside platforms has become more sealed than ever – not only because the tools capable of unlocking it have grown more powerful, thereby potentially increasing reputational and regulatory risks for platforms, but also because platforms now have strong commercial incentives to lock down data for AI training purposes. To a certain extent, the DSA already anticipated this paradox: in exchange for greater access to platform data, researchers must demonstrate that they can handle that data securely and in full compliance with the General Data Protection Regulation (GDPR).[2] This is no small task – particularly for organisations that lack the size or resources to employ a dedicated Data Protection Officer, or for those that previously relied on public-interest scraping rather than formal compliance infrastructures.

Amid these changes, a significant cultural shift will be required from researchers. As access requests face increased scrutiny from both regulators and platforms, data protection considerations can no longer be an afterthought; organisations will need to define and scope their research objectives with greater precision and strengthen their data protection posture. This is the new reality researchers must embrace if they want to take full advantage of the opportunities provided by Article 40(4) and improve their chances of a successful application.

To understand the potential improvements required – and the gaps organisations still need to address – Stiftung Mercator engaged AWO to conduct privacy audits and high-level GDPR readiness assessments for a selection of its grantees. The goal was to identify compliance gaps that could hinder organisations from making full use of the Article 40(4) data access procedure.

This paper summarises our high-level findings and outlines the key data protection requirements for any organisation seeking to use Article 40(4). While it cannot replace a project-specific GDPR analysis, integrating these considerations early on should significantly improve an organisation's ability to draft GDPR-compliant access requests.

[1] 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), available at: https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng.

[2] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), available at https://eur-lex.europa.eu/eli/reg/2016/679/2016-05-04.

2 GDPR and DSA Data Access under Article 40(4)

With GDPR compliance now central to accessing platform data in the EU, researchers must understand what this requires in practice. Data protection is not a one-off task that can be satisfied through a simple checklist; it is an ongoing process that must evolve alongside the research activities and the ways in which data are collected, analysed, and stored.

To help researchers with this task, this section of the paper provides some practical guidance that covers the following matters:

* The basics of data protection and its fundamental concepts
* The main requirements under the GDPR
* Key data protection themes that come up in the context of platform research
* Further resources to help with GDPR compliance

The issues covered in these sections are not designed to be exhaustive; the specific measures and processes needed for a given project will depend on the nature of the project, the data requested and how it will be used. Nevertheless, the information provided can form a starting point for researchers to develop an appropriate roadmap for ensuring and demonstrating compliance with the GDPR for their data access applications.

2.1 Primer on data protection

To understand data protection compliance, it is important to have a grasp of some basic concepts, as they form the foundation for identifying the measures required for a given project.

While most researchers are familiar with the existence of the GDPR, its underlying purpose can sometimes be obscured amid day-to-day compliance work.

Data protection is ultimately about making sure that people's information is used in a responsible and ethical manner. This applies across the data lifecycle, including its collection, retention, processing, disclosure and deletion – at each stage, certain principles and rules need to be followed.

The GDPR, which underpins these principles and rules, applies to the processing of personal data by controllers or processors for any purpose. This can be broken down as follows:

Question	*Answer*
What is personal data?	Any information relating to an identified or identifiable individual. This includes information that can be used to directly or indirectly identify the individual, including through the use of names, identification numbers, location data, online identifiers or factors specific to an individual's physical, physiological, genetic, mental, economic, cultural or social identity.
What is sensitive personal data?	This includes personal data revealing: * Racial or ethnic origin * Political opinions * Religious or philosophical beliefs * Trade union membership It also includes: * Genetic data * Biometric data (when processed for the purpose of uniquely identifying an individual)

	* Data concerning health * Data concerning a natural person's sex life or sexual orientation
What is processing?	This is any operation or set of operations performed on personal data, including its collection, recording, organisation, structuring, storing, adaptation or alteration, retrieval, consultation, use, disclosure transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.
What are controllers and processors?	Controllers are those who determine the purposes and means of the processing of personal data. If multiple entities determine the purposes and means of the processing together, then they are all joint controllers. An entity which processes personal data on behalf of the controller is a processor.

2.2 Main GDPR requirements

The table below summarises the main requirements that are imposed on data controllers under the GDPR along with some of the practical compliance measures that can be implemented for each:

Requirement	*Detail*	*Practical Compliance Measures*
Lawfulness, Transparency and Fairness	'Lawfulness' means that the processing of personal data must have a basis in law. 'Transparency' means information about the data being collected and processed is presented to data subjects in an accessible manner and is easy to understand. 'Fairness' means data has not been processed through unfair means, by deception or without the knowledge of the data subject. It also means that the processing cannot result in unjustified negative effects on the individual.	* Identify and document an appropriate legal basis for the processing of personal data * Complete a legitimate interest assessment (if applicable) * Draft and publish a data protection/information notice explaining how personal data is used * When processing special category data, ensure that you have both a legal basis and valid exception under the relevant data protection rules
Purpose Limitation	This requires that data are processed for specified, explicit and legitimate purposes. In addition, data cannot be further processed in a manner that is incompatible with the original purposes for which it is processed. This principle therefore means that data cannot be collected and used for unlimited, unspecified purposes.	* Establish access control policy * Set user roles and permissions for systems/applications/platforms used for processing personal data * Set up logging mechanisms to track access to data (e.g., login attempts, modifications to documents or databases) for the systems/applications/platforms used for processing personal data * Raise awareness in relation to the principle of purpose limitation and establish processes to ensure that personal data is not further processed for incompatible purposes
Data Minimisation	Data that is collected should be adequate, relevant and limited to what is necessary for the processing purposes. Data should therefore not be processed if the purpose can be reasonably fulfilled through other means.	* Document the list of categories of personal data that is required for the project * Document the justification for the use of each category of

		personal data used by explaining its relevance and utility for the project * Develop a process to identify and delete data that is irrelevant for project purposes * Apply pseudonymisation techniques to personal data where appropriate (e.g., removing direct identifiers)
Accuracy	Personal data that are processed must be accurate and, where necessary, kept up to date. Every reasonable step must be taken, without delay, to ensure that this is the case.	* Establish criteria used for labelling data (e.g., how inauthentic accounts are identified) * Implement manual or automated data quality checks (e.g., duplicate detection) * Establish a process for human review of any automated processes used for analysing data (e.g., for AI systems/tools/services)
Storage Limitation	This means that if data are being stored in a format that allows for the identification of data subjects, then it should be kept for any longer than is necessary for the processing purposes. This may entail establishing time limits for erasure or periodic reviews of which stored data should be kept.	* Develop an internal data retention schedule specifying the retention periods for each category of personal data used for the research * Implement manual or automated retention policies/schedules in line with the timeline for the project for the systems/applications/platforms used for processing personal data
Integrity and Confidentiality (Data Security)	Any data processed should be protected by appropriate security. The security measures put in place should address the specific risks identified in the processing operation. In particular, those measures should protect against unauthorised or unlawful processing and against accidental loss, destruction or damage.	* Document the third-party service providers involved in the project that are also processing personal data (e.g., cloud service providers or providers of AI systems/tools/services) * Firewalls * Encryption (at-rest and in-transit) * Backups * Incident response plans * Confidentiality clauses/obligations imposed on employees handling personal data * Staff training
International Data Transfers	Transfers of personal data to locations outside of the EU/EEA can only take place if either the European Commission has made an adequacy decision with respect to the third country, adequate safeguards have been provided (e.g., SCCs) or one of the	* Exercise data residency options with service providers to ensure data storage on servers in the EU or at least a third country granted data adequacy by the

	derogations apply and the transfer is not systematic.	European Commission (where possible)[3] * Document relevant transfer mechanism for data transfers out of the EU/EEA
Data Subject Rights	Data subjects are afforded certain rights with respect to their personal data, including the right to be informed about the processing, to access the personal data processed, to rectify their personal data, to erase their personal data, to restrict the processing of their personal data, to the portability of their personal data and to object to the processing of their personal data.	* Implement standard operating procedures for responding to requests from data subjects
Accountability	This requires that the above principles be complied with, and such compliance should be demonstrated. In other words, it must be possible to show *how* the above principles have been followed.	* Maintain a record of processing activities (ROPA) for the project * Carry out a data protection impact assessment (DPIA) for the project (if required) * Carry out data protection reviews of any third-party providers being onboarded for the project * Enter into data processing agreements with any third-party providers onboarded for the project

Certain of these obligations will broadly look the same across different processing activities that an organisation may undertake. Those organisation-wide obligations would include the following:

* **Data security.** Ensuring the confidentiality, integrity and availability of data will usually involve the same set of measures which may be technical or organisational in nature. This includes the use of encryption (both for data at-rest and in-transit), backups and confidentiality clauses imposed on staff handling personal data. Additionally, incident response plans should be in place that apply across all processing operations and assign the necessary roles and responsibilities for responding to incidents when they occur, including those arising from research projects involving platform data.
* **Data subject rights.** The rights afforded to data subjects under the GDPR will apply regardless of the processing activity carried out. This means that organisations will need to have standard operating procedures or plans in place manage for responding to data subject requests they receive. Those requests could relate to personal data handled for platform research projects, and so measures should be in place for this before undertaking the research.
* **International data transfers.** As part of the research, researchers may use products or services that process personal data outside of the EU or EEA. Researchers will need to be aware of this possibility and check the privacy policies and data processing agreements of the products and services they are using to ensure that data resides in the EU/EEA or that an appropriate transfer mechanism is in place. Additionally, researchers that are based outside of

[3] List of adequate countries available at: https://commission.europa.eu/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en.

the EU/EEA will also need to have an appropriate transfer mechanism in place to access to data of EU users from platforms.

* **Accountability.** Certain measures for fulfilling the accountability principle will need to cover different processing activities. For example, records of processing activities (ROPAs) will need to be developed and maintained for projects using personal data.

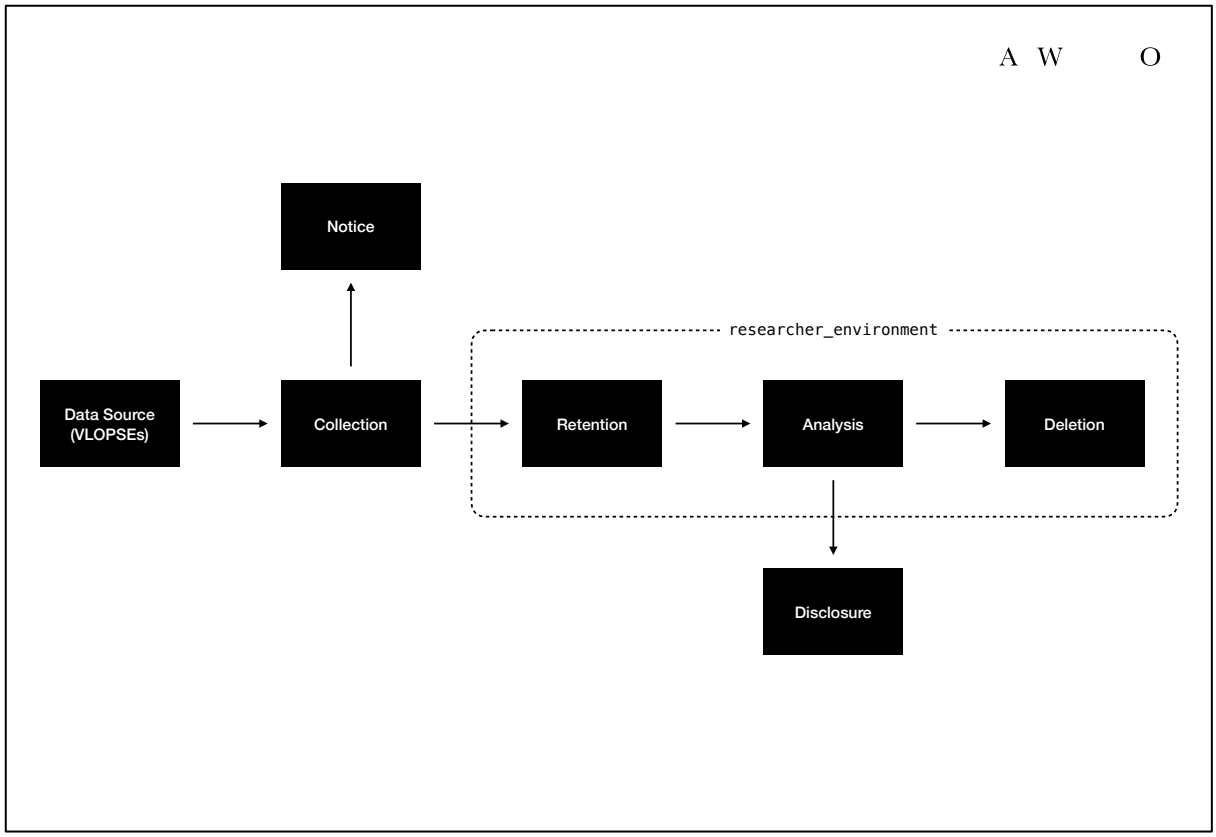
2.3 Key Data Protection Themes for Platform Research

2.3.1 How to get started

Researchers should begin their data protection compliance journey by putting together a detailed project plan that includes the following:

* The purposes of the research project
* The key research questions the project seeks to answer
* The specific DSA systemic risk that the research relates to
* The roles and responsibilities of those involved in the project
* A description of the data to be collected for the project, why it is needed and how it will be used
* A description of the specific data quality measures that will be used for the project
* A description of the technical architecture that will be deployed for the project, including a list of any third-party tools or services that may be used
* A timeline for the project

Putting together the above project plan will make it easier to identify the data lifecycle that the research project will involve. This data lifecycle will capture the data flows from which researchers can identify the controller obligations that apply across that lifecycle, which is shown in the diagram below:



For each obligation, researchers can refer to the table in Section 2.2 above to help identify the potential risks of non-compliance and measures that can be implemented to ensure compliance.

2.3.2 The concept of personal data

Personal data means any information that can be used to identify an individual, even if indirectly. The concept of personal data is very broad, and researchers obtaining data from VLOPSEs will most likely be obtaining personal data.

However, the concept of personal data is relative, meaning not all types of information will be regarded as personal data under the GDPR. This includes, in certain circumstances, pseudonymised data.[4]

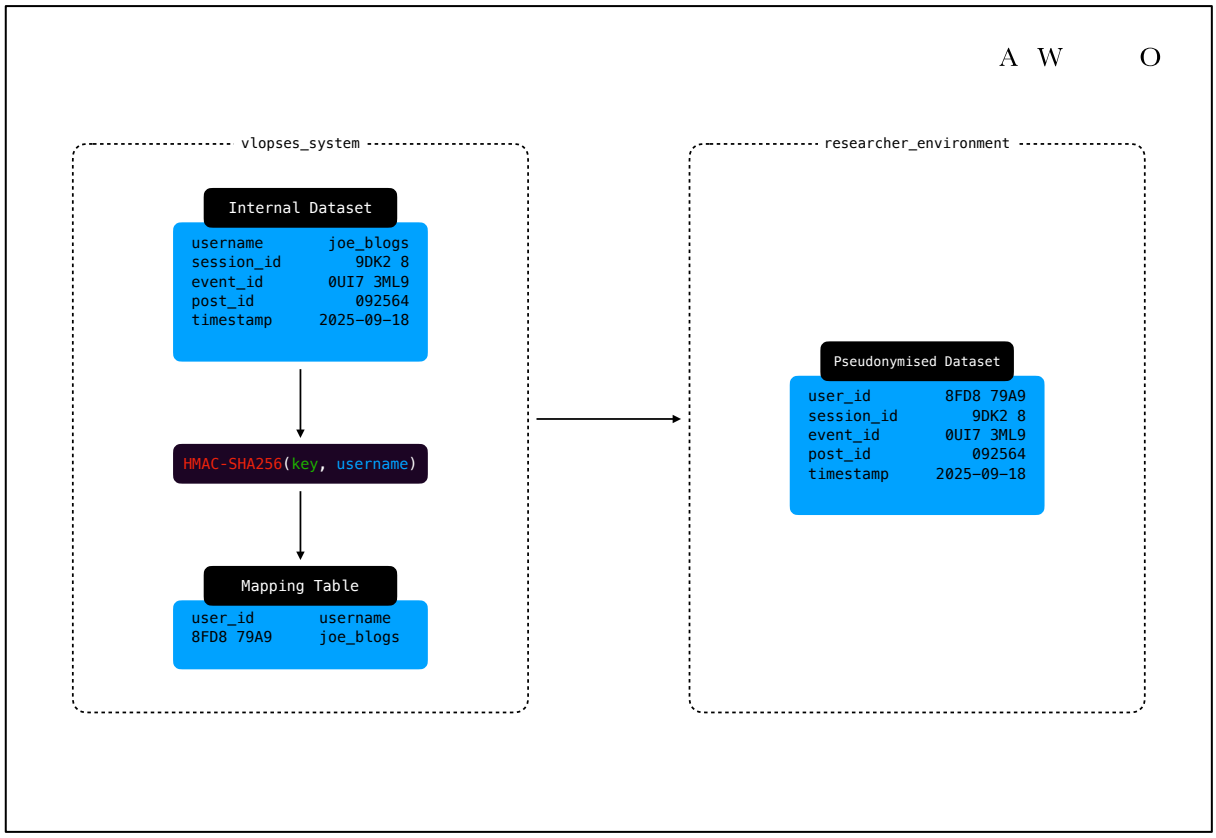
Identifiability refers to the ease with which information can be used to identify specific individuals. This means that identifiability exists on a spectrum – some information can be used to directly identify a specific individual (e.g., a name) whereas other information can indirectly identify an individual (e.g., a user ID). Indirect identifiability includes data points which can be linked to *a person* even if it is not known who *the person* is. Anonymity is the opposite of direct identifiability – this is where data cannot be linked to *any person* (e.g., aggregated statistics).

Pseudonymisation includes a range of techniques that reduce the identifiability of personal data. It consists of taking the original personal data and applying a pseudonymisation technique that produces pseudonymised data. The data is considered pseudonymised if the *pseudonymised data alone* cannot be

[4] See C-413/23 P, *EDPS v SRB* (2025), available at: https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:62023CJ0413.

used to directly identify the person. To identify the person, the pseudonymised data needs to be combined with *additional information*. As an example:

* A VLOPSE has a dataset of user activity with the column `username`.
* The VLOPSE replaces the username column with a `user_id` that is generated using a one-way cryptographic hash function,[5] so that `joe_blogs` becomes `8FD8 79A9`.
* The pseudonymised data is the hash value (i.e., `8FD8 79A9`) and the additional information is the table that maps `user_id` to `username`.



From the perspective of the VLOPSEs, the pseudonymised dataset with the `user_id` column is still personal data. This is because it still has access to the table mapping `user_id` to `username` and can use this table to identify users on its platform.

However, if the VLOPSEs shares with researchers *only* the pseudonymised dataset with the `user_id` column, this may not be personal data from the perspective of the researcher. That dataset will not be regarded as personal data if the researcher receiving the data cannot:

1. Lift the measures preventing re-identification, namely that the researcher cannot access the additional information (i.e., the table mapping `user_id` to `username`)
2. Otherwise perform reidentification through some other reasonably likely means, for example by cross-checking or combining the dataset with additional information the researcher could access

Regarding the second condition, the means used for reidentification are not considered *reasonably likely to be used* if the risk of identification "appears in reality to be insignificant", namely in cases where

[5] For example, HMAC-SHA256 with a secret key.

identification is prohibited by law (e.g., hacking or breaching the backend systems of the VLOPSEs) or impossible in practice.

If the above conditions are satisfied, then researchers do not need to treat the data received from the VLOPSEs as personal data. Such data falls outside the scope of the GDPR.

However, in practice, researchers will need to think carefully about whether the data they receive from VLOPSEs is pseudonymised in manner in which it cannot be linked to individuals. This means articulating how the data will be used throughout the data lifecycle of the project. For instance, if as part of the research project the VLOPSEs data will be combined with other datasets accessible to the researcher, then this could increase the identifiability of the data subjects in the VLOPSEs data and render it personal data under the GDPR.

Ultimately, if there is ever any doubt as to whether VLOPSEs data is personal data or not, it is best to treat the VLOPSEs data as personal data and implement the appropriate data protection measures.

2.3.3 Justifying the use of personal data

The GDPR provides a range of legal bases that could be relied on to process personal data. The Delegated Act suggests the legal bases that are appropriate for projects involving data from VLOPSEs accessed via data access process under the DSA, namely legitimate interest or public interest. For research projects involving the processing of sensitive personal data, researchers will also need to rely on an appropriate exception to the prohibition on the processing of such data under the GDPR. Again, the Delegated Act recommends either substantial public interest or archiving purposes in the public interest, scientific or historical research purposes or statistical purposes.

Regardless of the legal bases and exceptions relied on for processing, researchers will need to demonstrate that:

1. The research is in pursuance of a legitimate interest or is in the public interest, which in this context means showing that the main objective or purpose of the research must be to carry out research that helps detect, identify and understand systemic risks on VLOPSEs in the EU
2. The use of personal data is necessary to undertake the research
3. Any risks arising from the use of the data are addressed with mitigation measures

The other principles that are important for justifying the use of personal data include purpose limitation, data minimisation and storage limitation:

* **Purpose limitation.** As required by the DSA, the main objective or purpose of the project must be to carry out research that helps detect, identify and understand systemic risks on VLOPSEs in the EU. However, applicants may seek to undertake various different processing operations across the data lifecycle that contribute to the main objective or purpose (e.g., classifying or labelling content that may be illegal or breach the terms of service of a given platform). This is permitted so long as applicants can demonstrate that those processing operations are necessary for and compatible with that main objective or purpose.
* **Data minimisation.** Reducing the amount of personal data to that which is needed for the research project reduces compliance work and risk. Applicants should therefore, for each research project, consider and implement measures for achieving this. This could include different pseudonymisation techniques such as removing direct identifiers, generalising or masking certain data points or adding noise to datasets. The techniques applied will ultimately

depend on the data required for the research and ensuring a balance between identifiability and usability.

* **Storage limitation.** This is linked to the measures applied for data minimisation. Retention periods for data can be determined either by setting a hard date (i.e., 'data will be kept until x date') or using a more general criteria (i.e., 'data will be kept until the research project is complete). Regardless of the retention period set, applicants will need to demonstrate that the period is necessary for achieving the main objective or purpose of the research.

2.3.4 Assessing data protection risk

A key requirement specified under the DSA for data access applications is demonstrating that any risks arising from the use of personal data can be addressed by the researcher. In this regard, the GDPR lays down a risk-based approach, meaning that researchers are only required to implement measures that address the risks that are specific to the way personal data are used for a given project.

The exact risks that researchers may encounter when carrying out research using personal data will certainly vary. However, when identifying and mitigating risk, researchers should take a compliance-based approach whereby risks are defined as events that constitute non-compliance with the GDPR and lead to negative consequences for data subjects. For each risk, researchers should identify how the risk may manifest in the context of the processing activities carried out for the research and the measures that could be implemented to remediate the risk and its consequences.

The table below provides some examples of the risks researchers may encounter using data from VLOPSEs for DSA research and the potential remediation measures for each:

Risk	*Description*	*Potential Remediation*
Reporting potentially inaccurate assessments of violative content to VLOPSEs	In carrying out the research, researchers may identify content that potentially violates the terms of service of a platform and report it to the platform. This could result in the platform taking action against the offending account, such as suspending the account. However, if the analysis on the content being violative is inaccurate, then users may be unfairly moderated as a result.	✓ Evaluate any automated techniques used to ensure that they provide appropriate levels of accuracy ✓ Include a process for human review of the outputs of any automated tools/services used
Unauthorised use of personal data by service provider	The research may involve the use of products and services from other service providers. Researchers may therefore be processing personal data using these products and services. This may may involve the risk that such data is processed by service providers for their own purposes. For example, AI providers may use data for training their models or systems.	✓ Use open-source tools that can be deployed locally (if possible) ✓ Check the privacy-preserving settings available for the product or service to ensure that data cannot be processed by providers for their own purposes (e.g., opt-out of using data for training purposes) ✓ Enter into data processing agreements with service providers and review their terms to ensure they provide with an adequate level of data protection compliance
Lack of information provided to data subjects	Though personal data will not be collected from the data subject directly, it is still required that information is provided regarding how their data might be used for research purposes. A lack of such information makes it more	✓ Include on your website a privacy/data protection notice providing information about how you process social media data for your research projects

	difficult for them to exercise their rights with respect to their data.	✓ Provide a means of contact for people to get in touch to find out more about how personal data may be used for research projects and to exercise their rights

The Delegated Act mentions that DPIAs can be used to demonstrate that risks arising from the processing of personal data have been identified along with appropriate remediation measures. However, researchers are only required to complete a DPIA if any relevant criteria under the GDPR have been met. For researchers, this would include research projects that involve at least one of the following:

* Systematic and extensive evaluation of personal aspects of individuals based on automated processing on which decisions are based that produce legal effects concerning the individual or similarly significantly affect the individual. This could be the case for a research project which, for example, involves identifying potentially illegal content and reporting this to VLOPSEs or competent authorities.
* Processing sensitive personal data on a large scale. This could be case for research projects focused on election interference which involves the collection of content relating to the political opinions or affiliations of users.

2.3.5 Informing data subjects

The GDPR requires researchers to inform data subjects about the processing of their personal data. This includes where the data are collected via VLOPSEs rather than collected from the data subject directly. However, especially in the context of data access under the DSA, informing data subjects of the processing may involve a disproportionate effort or seriously impair the achievement of the objectives of the research.

It may therefore be sufficient for researchers to draft and publish on their website an information notice that includes the following information

* The identity and contact details of the researcher (this doesn't need to include family names, but just information on the organisation and means of communication that are monitored)
* The purposes of the processing for which personal data are intended as well as the legal basis for the processing – including the legitimate interest or public interest being pursued by the researcher using the personal data (i.e., helping to detect, identify and understand systemic risks on VLOPSEs in the EU)
* The categories of personal data processed
* The recipients or categories of recipients of the personal data
* Whether international data transfers are involved in the processing and the relevant transfer mechanism relied on for this processing activity
* The retention period for the personal data processed
* The rights that data subjects can exercise with respect to the personal data processed by the researcher
* The right of data subjects to lodge a complaint with a supervisory authority
* The sources from which personal data are collected and that they are collected from publicly available sources (i.e., social media platforms)
* Information related to automated decision-making

2.3.6 The approach for individual researchers

Individual researchers may submit data access applications under the DSA, but it is essential that they are supported by their affiliated research organisation. Key elements of the application process – particularly those relating to data protection compliance – require institutional input and cannot be fulfilled by an individual alone. For example, if you are using the IT systems of the affiliated research organisation, it will need to be ensured that such systems are equipped with the necessary security measures (encryption, network security, access controls etc). More specifically, certain of the controller obligations listed in Section 2.2 of this paper will require the support of the affiliated research organisation (i.e., on data security, handling data subject rights, implementing appropriate transfer mechanisms for international transfers, and accountability measures).

However, there are certain aspects of the data protection compliance process that individual researchers will need to lead or provide input on when working with their affiliated research organisation. Those aspects include the following:

* Identifying an appropriate legal basis for the processing of personal data (see Section 2.3.3 above)
* Implementing purpose limitation measures (see Section 2.3.3 above)
* Implementing data minimisation measures, for example appropriate pseudonymisation techniques (see Section 2.3.3 above)
* Implementing measures for verifying the outputs of automated tools used to conduct the research (see table in Section 2.2 above)
* Ensuring that any data collected, including any analysis done, is stored in a secure manner and is only kept for as long as is necessary for purposes of the research project (see Section 2.3.3 above)

2.3.7 Research regarding public figures

As mentioned in Section 2.3.2 above, the concept of personal data is broad. The GDPR therefore equally applies to the use of personal data of public figures such as politicians. While their public roles may affect the level of privacy they can reasonably expect, this does not alter the fundamental point: if the information qualifies as personal data, the GDPR applies.

2.4 Further Resources

Other helpful data protection resources include the following:

* Coimisiún na Meán (Irish DSC) Guidance for Applicants [Link]
* Center for User Rights Risk Assessment Tool for Researchers [Link]
* EDPB Guidance for Small Businesses [Link]
* EDPB Pseudonymisation Guidance [Link]
* Irish DPC DPIA Template [Link]
* ICO Legitimate Interest Template [Link]
* ENISA Recommendations for a methodology of the assessment of severity of personal data breaches [Link]

Annex 1 – Summary of Data Protection Requirements under the DSA and Delegated Act

Under Article 40(4) of the Digital Services Act (DSA), vetted researchers can obtain access to data from very large online platforms and search engines (VLOPSEs) for the purposes of helping detect, identify and understand systemic risks on the platform or search engine. Article 40(8) details the following criteria for becoming a vetted researcher:

* Affiliated to a research organisation
* Independent from commercial interests
* Capable of fulfilling data security and confidentiality requirements by implementing the appropriate technical and organisational measures
* The access to data and time frames are necessary for and proportionate to the purposes of the research and contribute to compliance with the DSA regarding systemic risks
* The planned research will be carried out for the purpose of helping detect, identify and understand systemic risks on VLOPSEs
* The results of the research are made freely and publicly available

Article 40(13) further states that the European Commission must adopt delegated acts specifying the technical conditions under which VLOPSEs are to share data with vetted researchers. This includes, among other things, the conditions under which data sharing with researchers can take place in compliance with the GDPR.

In July 2025, the European Commission adopted a Delegated Act laying down the technical conditions and procedures under which providers of VLOPSEs are to share data with vetted researchers.[6] The Act makes provision for how the data access process will be managed, the requirements for formulating reasoned requests to VLOPSEs for access to data, and the technical conditions for provision of access to data by the VLOPSEs.

Under the Delegated Act, applicant researchers need to apply to the relevant Digital Services Coordinator (DSC). The DSC must then review the application and determine whether it is sufficient for the DSC to make a reasoned request to the relevant VLOPSEs to access data. To make this determination, the DSC must consider, among other things:[7]

* The description of the data requested, including the format, scope and specific attributes (if possible), relevant metadata and data documentation (taking into account the DSA data catalogue provided by VLOPSEs which describes the data assets that can be accessed for the purposes of Article 40(4) DSA)[8]
* The information provided on the necessity and proportionality of the access to the data and the information on the time frames of the research for which the data are requested[9]

---

[6] Commission Delegated Regulation (EU) 2025/2050 of 1 July 2025 supplementing Regulation (EU) 2022/2065 of the European Parliament and of the Council by laying down the technical conditions and procedures under which providers of very large online platforms and of very large online search engines are to share data with vetted researchers, available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32025R2050.

[7] Delegated Act, Article 8.

[8] Delegated Act, Article 8(c).

[9] Delegated Act, Article 8(d).

* The information on confidentiality, security, and data protection risks related to the data that would be accessed, as well as a description of the technical, legal and organisational measures put in place, including suggested access modalities[10]
* The description of research activities to be carried out with the data[11]
* The summary of the application that includes: (i) the research topic, (ii) the VLOPSE from which data are requested and (iii) a description of the data requested[12]

Recital (15) of the Delegated Act further states that DSCs should verify for each data access application that:

* Information on the appropriate legal basis to be relied on for the processing of personal data (including special category data) is provided, including on whether the applicant seeks to rely on both or either public interest[13] or legitimate interest[14] for processing personal data and either substantial public interest[15] or archiving purposes in the public interest, scientific or historical research purposes or statistical purposes for special category data.[16]
* A sufficient indication has been given that the risks to personal data protection have been assessed, for example by the completion of a data protection impact assessment (DPIA) if required under the GDPR.

If the applicant researcher meets the above criteria, it can become a vetted researcher under the DSA, and the DSC can make a reasoned request to the relevant VLOPSE based on the application. Additionally, VLOPSEs cannot impose any measures on researchers other than those included in the reasoned request, including data protection measures.[17]

GDPR compliance is an essential part of platform data access in the EU. If researchers want to gain access to this data to carry out important platform research that contributes to the detection, identification and understanding of online systemic risks, getting data protection measures in order is a must.

[10] Delegated Act, Article 8(e).
[11] Delegated Act, Article 8(f).
[12] Delegated Act, Article 8(g).
[13] GDPR, Article 6(1)(e).
[14] GDPR, Article 6(1)(f).
[15] GDPR, Article 9(2)(g).
[16] GDPR, Article 9(2)(j).
[17] Delegated Act, Article 15(3) and (4).