



Data Access for Recommender Systems Research: leveraging the EU's Digital Services Act

João Vinagre
European Commission, Joint Research
Centre
Seville, Spain
joao.vinagre@ec.europa.eu

Lorenzo Porcaro
Sapienza University of Rome
Rome, Italy
lorenzo.porcaro@uniroma1.it

Silvia Merisio
European Commission, DG Connect
Brussels, Belgium
silvia.merisio@ec.europa.eu

Erasmus Purificato
European Commission, Joint Research
Centre (JRC)
Ispra, Italy
erasmo.purificato@ec.europa.eu

Emilia Gomez
European Commission, Joint Research
Centre (JRC)
Seville, Spain
emilia.gomez-
gutierrez@ec.europa.eu

Abstract

The European Union (EU) Digital Services Act (DSA) has introduced a novel set of rules for online platforms and search engines, with significant implications for the Recommender Systems community. Through its data access mechanisms, the DSA invites researchers to request both publicly available and private data from Very Large Online Platforms (VLOPs) and Very Large Search Engines (VLOSEs) – those with more than 45 million active recipients in the EU – to investigate systemic risks associated with the dissemination of illegal content, risks to the exercise of fundamental rights, and negative effects on electoral processes, public health, and gender-based violence. This tutorial is aimed at researchers who are interested in submitting such data access requests and will provide them with the knowledge to do so by introducing the relevant definitions and provisions of the DSA, and addressing the most important procedural steps to obtain data access and will provide attendees with a comprehensive understanding of the DSA's data access implications for RecSys research. The tutorial targets researchers, practitioners, and students in understanding current developments in online platform regulation in Europe and their impact on RecSys research.

ACM Reference Format:

João Vinagre, Lorenzo Porcaro, Silvia Merisio, Erasmo Purificato, and Emilia Gomez. 2025. Data Access for Recommender Systems Research: leveraging the EU's Digital Services Act. In *Proceedings of the Nineteenth ACM Conference on Recommender Systems (RecSys '25)*, September 22–26, 2025, Prague, Czech Republic. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3705328.3748004>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

RecSys '25, Prague, Czech Republic

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-1364-4/25/09
<https://doi.org/10.1145/3705328.3748004>

1 Motivation and Objectives

The Digital Services Act [2] regulates how online platforms and search engines may operate within the EU. Platforms and search engines with more than 45 million active recipients are designated Very Large Online Platforms (VLOPs) and Very Large Online Search Engines (VLOSEs) and are subjected to specific obligations, given their size and potential societal impact. These obligations are related to the systemic risks identified in the DSA which fall under four categories: a) the dissemination of illegal content; b) risks to the exercise of fundamental rights [1]; c) negative effects on civic discourse and electoral processes, as well as public security; and d) negative effects related to gender-based violence, protection of minors and public health, and negative effects to physical or mental well-being of individuals.

Recommender Systems are at the core of most VLOPs and VLOSEs, playing a central role on their design and operation. In the ACM RecSys community there is a growing interest on societal impacts of algorithmic recommendation, with an increasing number of papers, workshops, tutorials and other contributions related to topics such as fairness, bias, normative approaches, recommendation for vulnerable users – including minors –, gender perspectives, health and the overall influence of recommender systems in informative and cultural spaces.

Given the lack of consolidated scientific knowledge about the impacts of VLOPs and VLOSEs on society, and the complexity of systemic risks, the DSA invites researchers to study and contribute to the understanding of risks that may stem from the design and functioning of VLOPs and VLOSEs. Researchers can request data from VLOPs and VLOSEs if their research (i) is clearly connected to systemic risks in the EU, and (ii) satisfies a set of conditions. With this, the DSA provides an unprecedented legal mechanism for addressing a wide range of research questions that have previously largely been left unanswered due to the unavailability of data. The goal of this tutorial is to encourage researchers to submit DSA data access requests and to equip participants with the necessary knowledge to do so.

2 Tutorial length and outline

The tutorial has a duration of 3 hours including breaks synchronized with the conference agenda.

I The DSA and the Data Access Provisions (60 minutes)

- **Introduction:** Tutorial background, motivation, objectives, and relevance to RecSys community.
- **The Digital Services Act:** Brief introduction to the EU legal system; History and objectives of the DSA; Similar initiatives worldwide; Definitions and stakeholders; Definition of systemic risks.
- **Discussion: Mapping research questions to systemic risks**

II Types of Data Access (60 minutes)

- **Access to publicly available data:** Eligibility; Application process; Technical aspects.
- **Access to private data by vetted researchers:** Data access request workflow; Eligibility and research team; Project description and duration; Evaluation and amendments; Technical means to access data; Data protection, privacy and confidentiality; Publication of results.
- **Discussion: Data requirements - public vs non-public data**

III Interactive session (60 minutes)

- Selection of a research question relevant to systemic risks;

- Collaborative drafting of a mock data access application for access to non-public data: focus on the main elements of a data access application for the ‘vetted’ status;
- Final discussion and wrap-up.

3 Targeted audience and prerequisites

The intended audience of this tutorial are researchers focused on leveraging online platform data for addressing systemic risks under the EU’s DSA, particularly those within the RecSys community and the broader recommender systems domain. The tutorial welcomes researchers, practitioners, and professionals from diverse fields and levels of experience, including those new to the EU legal framework and the DSA.

4 Related Tutorials

The European Commission has organised multiple dissemination initiatives related to Data Access for researchers (Art. 40 DSA), that include other similar tutorials, such as the ones held at ICWSM 2025, UMAP 2025 and ECML PKDD 2025.

References

- [1] European Union (EU). 2012. Charter of Fundamental Rights of the European Union. https://eur-lex.europa.eu/eli/treaty/char_2012/oj/eng
- [2] European Union (EU). 2022. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act). <https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng>