

DSA Article 40: The EU's tool for research access to online platforms

Ali Knott

Victoria University of Wellington, New Zealand
Social Data Science Alliance, Global Partnership on AI



Human-AI interaction ‘in the large’

Until recently, the field of ‘human-computer interaction’ has involved the study of *individual users*, interacting with *separate instances* of a computer system.

The kind of questions asked:

- How well can the user achieve their goal?
- How good is the interface?

But these days, a single computer system can be deployed to a huge *population* of users.

- Each user has their own front-end interface to the system.
- But there’s a single back-end. . .
 - Which *aggregates information* from many users. . .
 - And uses this combined information to *respond* to each user.

Human-AI interaction 'in the large'

These larger computer systems can be called **Very Large Online Platforms** (or **VLOPs**).

- VLOPs are powered by **AI**, which is why their influence is so huge.

With a VLOP, a whole *population* of users is effectively interacting with a single (complex) AI system.

- Of course, we can still study how individuals interact with their instance of the system.
- But we can also study *broader effects*:
 - How the population of users influences the system
 - How the system influences the population of users.

In this broader domain, a new process of **cultural evolution** is taking place.

There are two evolutionary processes under way.

1. **The users evolve their behaviour**, because the AI system offers powerful new opportunities and incentives.
 - New ways become available to make money, make friends, work efficiently (etc).
 - Users who pick up on these have more of an influence on the system, and therefore on other users. . . and so on the culture.
2. **The AI system also evolves**—by its own learning. . . and because its designers *tweak* it, to make it better for their purposes.
 - The competing entities here are *alternative versions of the system*.
 - Many versions are trialled. The most effective ones are most widely deployed. . . and have most influence on the culture.

We urgently need to understand these processes!

They seem to be having huge effects on our society.

- We have a new object of scientific study: **human-AI coevolution**.

How should we study human-AI coevolution?

To properly study the process, we need new forms of access to VLOPs.

1. VLOPs are the only agencies who have all the *data* needed to observe the evolutionary process.
 - Interactions between users and platforms are *private to VLOPs*.
 - VLOPs are typically secretive about the AI systems they deploy.
2. VLOPs are the only agencies in a position to *experiment* with the algorithms they deploy.
 - They certainly do experiment, to find the best versions.
 - But their expts *serve their own interests*, not those of society.
 - And expts run by individual VLOPs are *limited in scope*.

We need a new public science of tech impacts!

The new science:

- Must have access *new data*, gathered from within VLOPs.
- Must deploy *new methods*, that operate within VLOPs.

The new science must also be a *public one*, where results are made available in journals, conferences.

- A public science is vital to unite the efforts of the whole scientific community, and to address social concerns.
- But the VLOPs don't want to do their part.

Enter the DSA!

The EU's **Digital Services Act (DSA)** came into force in Nov 2022.

- The Act 'introduces rules for online services used by EU citizens in their everyday life'.

There are four main categories of provider:

- **Intermediary services** that host user content
 - That's social media platforms and online marketplaces.
- **Online platforms**, that disseminate information to users
 - Very broadly defined.
- **Online search engines** (Google search and similar).
- **Very Large Online Platforms and Search Engines (VLOPSEs)**
 - Those engines with more than 45M monthly EU users.
 - Currently the **EU's list** includes 26 VLOPSEs.

Obligations imposed by the DSA

A key obligation for all ‘intermediary services’:

- They must produce a yearly **transparency report** on ‘any content moderation that they engaged in’ (Art 15).
 - This should include reports on the **accuracy of automated content moderation tools**.

Some key obligations on all ‘online platforms’:

- They must act swiftly to remove **illegal content** (Art 16), and enact measures that prevent its reappearance (Art 20).
- They should explain the design of their **recommender systems** (Art 27).
- They must enact special measures for **protection of minors**: in particular, no profile-based advertising (Art 28).

Special obligations for VLOPSEs

VLOPSEs must:

- Make their own yearly **assessments of 'systemic risks'** they may produce (Art 34) and their mitigation (Art 35).
- Provide users with at least one **non-ML-based recommender system** (Art 38).
- Arrange for **independent audits** of their compliance with other DSA obligations (Art 36).
- Provide **transparency about online advertising**, through a public tool (Art 39).
- More frequent & detailed **transparency reports** on content moderation (Art 42).
 - For human moderators: information about qualifications/skills
 - For automated moderation: accuracy breakdowns by language.

DSA Article 40

Article 40 has some key obligations relating to research access to VLOPSEs.

- Art 40.4 requires VLOPSEs to 'provide access to data' to 'vetted external researchers', to do research on the 'detection, identification and understanding' of 'systemic risks in the Union', and to the 'mitigation of these risks'.
- Art 40.12 requires VLOPSEs to give vetted researchers access to 'real-time data publicly accessible data', to do research on the same systemic risks and their mitigation.

My talk (and Virginia's!) will focus on DSA 40.4.

Support for prospective DSA 40.4 researchers

Negotiating access under DSA 40.4 will be a *battle!*

- The current US government is disposed to protect US platforms against EU regulation.
- All VLOPSEs have huge teams of lawyers and lobbyists, to defend their interests.
- To coordinate Article 40.4 access requests, it would be useful for researchers to *act collectively*, rather than individually.

A few organisations have been set up to assist researchers.

- The [DSA observatory](#) (University of Amsterdam) is an *observer* of DSA 40 access processes, with a focus on tech law.
- The [DSA40 collaboratory](#) (Weizenbaum Inst, Euro New School) is a *support hub* for researchers considering access requests.

The Social Data Science Alliance

The Observatory and Collaboratory are both focussed on *legal* aspects of DSA access.

- We felt there was room for a more *technical* organisation, focussing on the *substance* of DSA-enabled research.
- So we set up the [Social Data Science Alliance \(SDSA\)](#).
 - It was launched at the Paris AI Summit in Jan 2025.
 - It has over 80 members, and 5 supporting organisations.
 - Founding members include Yoshua Bengio, Stuart Russell, Gillian Hadfield, Jonathan Stray, Jeff Allen, Virginia Dignum, Albert-Laszlo Barabasi, Sandy Pentland, Ricardo Baeza-Yates.
 - Closer to home: Dino Pedreschi, Fosca Giannotti, Luca Pappalardo, Luca De Biase, Virginia Morini, Roberto Trasarti, Emmanuele Ferragina, Andrea Passarella, Luca Belli, Stefano Cresci.

The SDSA's focus

Recall: studying human-AI coevolution requires researchers to have *access VLOPs*.

- To see the relevant data, to run the relevant experiments.

DSA 40.4 gives researchers a good deal of the necessary access.

- It has the potential to create a **new public science** of tech impacts, and their mitigation.

The SDSA's aim is to identify the *shape* of this new public science.

- *What questions* should researchers ask, with DSA40 access?
- *What methods* should researchers use, in DSA-enabled work?

Some collective thought on these matters will help researchers coordinate applications, and collaborate.

The SDSA's first project

Our first project was to ask members *what questions they thought should be asked*, under DSA 40's new access provisions.

- Over 100 questions were suggested!
- You can find our report [here](#).
 - Our report presents a **taxonomy** of research questions.

A taxonomy of research questions

Our taxonomy focuses on risks associated with three basic *human activities* conducted in platforms, each relating to *platform algorithms*.

- Our focus is on *platform algorithms*, because platforms are largely controlled by algorithms.

Three platform activities

Technical teams *develop* and *deploy* algorithms, and regularly *train* and *adjust* them.

- These are the *actions* through which the platform can be controlled.

The platform is intensively *monitored*, to observe the *effects* of the deployed algorithms on its traffic and users.

- Monitoring provides *feedback*, which informs how the algorithms are adjusted.

At the highest level, platform executives create *policies* about how algorithms should be adjusted, in the light of feedback.

- These define what state platform owners are *working towards*.

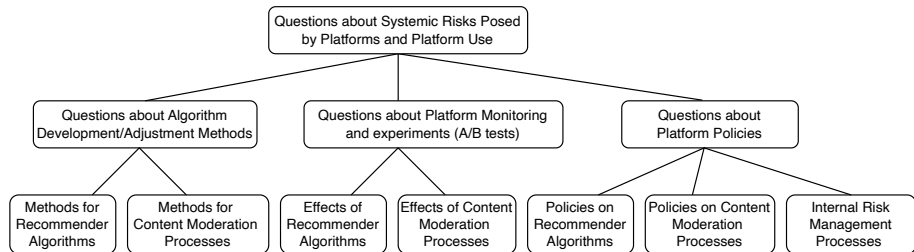
Two types of platform algorithm

Following the Integrity Institute's [2024 report](#), we distinguish two key types of platform algorithm.

- **Recommender algorithms** *push* content at platform users, ranking items for presentation in content feeds, or lists of search results.
- **Content moderation algorithms** *withhold* content from users, by identifying harmful or inappropriate content in various categories, and moderating it in various ways.

The three activities, and two algorithm types, give us the top level of our taxonomy.

The top level of our taxonomy of research questions



Other structure in the taxonomy

We also documented the different types of *user*, *content* and *external event* mentioned in member questions. Some examples:

	User group	Content category	Event type
Politics	political dissidents	disinformation, 'news content', 'legitimate democratic discussion'	elections, political crises
Society	immigrants, users of a particular race, gender, class	sexually explicit material, misinformation, hate speech	outbreaks of violence
Economics	poor people	ads, 'commercial content'	'seasonal shocks', 'demand spikes', cut-price sales
Health	young women	health content	pandemics

Organisation by platform type

We also organised questions by *platform type*.

- Social media platforms (Sec 3) ← most member qns were here!
- Retail / commercial platforms (Sec 4)
- Mapping platforms (Sec 5)
- Search engines (Sec 6)
- Adult content platforms (Sec 7)

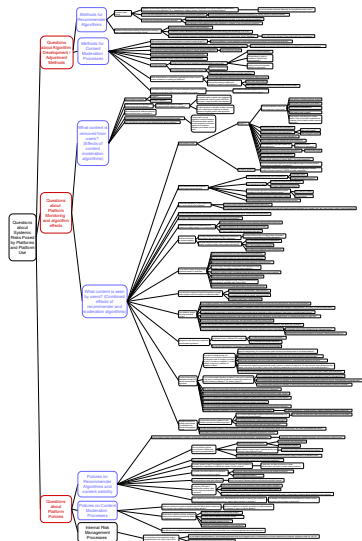
We separately grouped questions about *Gen AI methods in VLOPSEs* (Sec 8).

- These are in scope for the AI Act, but also for the DSA.

Member questions for social media platforms

Members suggested *many* research questions!

- Each question is tagged with the name of the researcher who asked it.



Some sample member questions

Under 'algorithm monitoring':

How does platform content influence real-world events?

Does 'dangerous speech' (as defined by Susan Benesch) at a given time and place correlate with real-world political violence at that time/place?

Does sustained exposure to misinformation, propaganda, or extremist content online translate into measurable changes in voting behavior, civic trust, or political radicalization?

Do interactions with misogynistic or otherwise extremist communities online contribute to offline harassment, hate crimes, or gender-based violence?

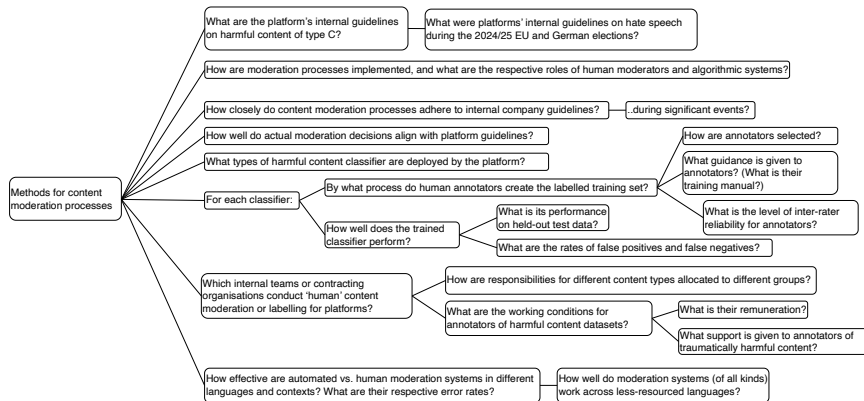
Does being subjected to targeted harassment or hate speech online lead to social withdrawal, mental health issues, or reduced participation in public life?

How do platforms affect mental health, body image, and radicalization, especially among youth?

What are the long-term effects of exposure to platform content on political participation, trust in institutions, and democratic discourse?

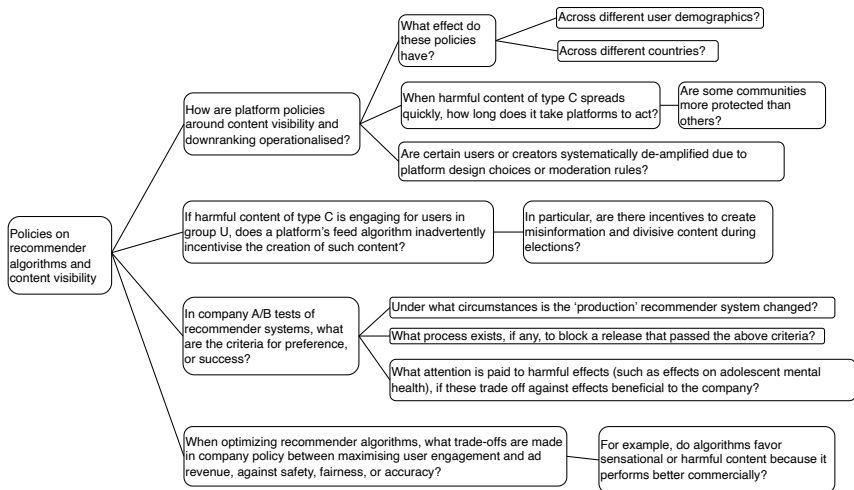
Some sample member questions

Under 'algorithm development and adjustment':



Some sample member questions

Under 'policies':



Which questions are most supported by members?

Support could be useful, in directing the DSA 40.4 process.

- It might help prioritise research topics, and help choose the most useful datasets for companies to prepare.
- It might help negotiations between researchers and platforms.

An exercise for you!

Get into groups, and take a look at the questions in Sections 3–8 of our report.

- Or browse through the samples on pp23–25 of these slides.
1. Choose a question you think is particularly important.
 - Why is it important? What method would you use to answer it?
 2. Are there questions you don't like? Are there questions missing?

A 'leaderboard' of supported research questions

The most supported questions from members at present:

- How is user data employed to train, fine-tune, and update recommender algorithms?
- How do LLMs personalize responses based on users' characteristics, and what are the implications for linguistic homogenization and cultural bias?
- How does the design of a recommender system influence the dynamics of content of type C for users in category U , over time?
- How does introduction of LLM-powered chat or summarising devices affect user behaviour?
- What impact are 'AI summaries' in search results having on user traffic to news providing sites?