Digital Policy Lab
Policy Digest #19
23 April 2025

Policy Digests offer an overview of recent digital policy developments in Digital Policy Lab (DPL) member countries, including regulatory and non-regulatory initiatives aiming to combat online harms such as disinformation, hate speech, extremist or terrorist content. In addition to general updates, each Policy Digest provides a snapshot of topic-specific schemes relevant to the upcoming DPL session.[1]

Section 1 Digital policy developments

Australia: eSafety Commissioner publishes transparency report summarising service providers' responses on measures relating to online terrorist and extremist content

Type Regulatory (report)
Status Published

On 6 March 2025, the eSafety Commissioner published a Basic Online Safety Expectations (BOSE) report summarising industry responses to transparency notices under the Online Safety Act. The notices were issued under the Online Safety Act (section 56(2)) in March 2024 for Google, Meta, WhatsApp, Reddit, Telegram and X and required each to outline detailed information on how their service is meeting its Basic Online Safety Expectations. Each service provider had 49 days to respond to the notice or face a fine of AUD 782,500 (around EUR 435,000) per delayed day. The report examines the service providers' responses about the tools, policies and processes used by each to address terrorist and violent extremist material and activity (TVE) on their services from 1 April 2023 to 29 February 2024. A summary of the findings can be found here.

Canada: Office of the Privacy Commissioner opens investigation into X on data used to train AI

Type Regulatory (investigation)
Status Ongoing

On 27 February 2025, the Office of the Privacy Commissioner opened an investigation into X on whether it violated Canadian privacy laws to train its artificial intelligence (AI) models. The investigation will determine whether X is meeting its obligations under the Personal Information Protection and Electronic Documents Act (PIPEDA) regarding how X collects, uses and discloses Canadian's personal information to train AI models. A similar investigation conducted by Switzerland can be found below.

European Union: European Commission launches Apply AI Strategy consultation

Type Regulatory (consultation)
Status Open

On 9 April 2025, The European Commission's Artificial Intelligence Office launched a call for evidence and a public consultation regarding its Apply AI Strategy. The Apply AI Strategy aims to position Europe as a frontrunner in AI innovation. Its primary objectives include empowering European businesses to lead globally in AI, enhancing the quality of public services, and integrating AI technologies across industries. It will also connect with the forthcoming European Strategy for AI in Science,

---

[1] We welcome any feedback from DPL members regarding additional developments, as well as own submissions from DPL members who wish to be featured in the digest.

which is set to be adopted concurrently. Stakeholders, including industry representatives, civil society, and public authorities, are invited to respond to the public consultation to identify priorities and challenges related to AI implementation, including aspects of the AI Act. This consultation will close on 4 June 2025.

European Union: Regulation to establish independent scientific expert panel on AI enters into force

Type Regulatory (enforcement)
Status In force

On 30 March 2025, the implementing regulation on the establishment of a scientific panel of independent experts in the field of AI came into force, under the EU AI Act. The panel is tasked with issuing qualified alerts and seeking support from the EU AI Office, with a focus on the impacts and enforcement of the Act. Experts are responsible for overseeing conflicts of interest, ensuring compliance, and may be appointed by the Commission to carry out thematic hearings or provide evaluations as required. The Commission will shortly issue a call for expressions of interest to select the experts.

European Union: European Commission publishes third draft of the AI Code of Practice

Type Voluntary Code of Practice
Status In consultation

On 11 March 2025, the European Commission published a third draft of the General-Purpose Artificial Intelligence Code of Practice. With contributions from civil society, academia and industry, this draft Code presents a revised structure and provisions and an enhancement of the outlined commitments. The Code applies to AI models classified as posing systemic risks, as well as general-purpose AI models. Its purpose is to support compliance with the EU AI Act and its provisions on general-purpose AI, which will take effect in August 2025. Following the written feedback from stakeholders submitted by 30 March 2025, along with feedback gathered from stakeholder workshops, the Chairs and Vice-Chairs of the Code will convene to finalise the text by May 2025.

Germany: Federal Court of Justice ruling about Meta's Facebook App Center on data use

Type Litigation
Status Decision issued

On 27 March 2025, the Federal Court of Justice (Bundesgerichtshof, BGH) issued a ruling in Federation of German Consumer Organisations v Meta Platforms Ireland on General Data Protection Regulation (GDPR) violations relating to how Facebook's App Center handles data processing. The court ruled that the service provider did not adequately inform its users about the scope and purpose of the collection and use of their personal data. The ruling gives rise to claims for injunctive relief under competition law and can be pursued by consumer protection organisations by way of an action before the civil courts.

Italy: Senate passes AI bill focusing on domestic data security and transparency

Type Legislative
Status Passed

On 20 March 2025, the Italian Senate (Senato) passed the Bill on Provisions and Delegation to the Government on AI, which outlines economic development principles and requires public procurement platforms to prioritise AI suppliers that process

and store strategic data within domestic data centres. The obligations aim to ensure disaster recovery, business continuity and high security and transparency standards. Only in exceptional cases should AI systems used in the public sector be hosted outside of Italy. Additionally, the Bill enhances collaborative research between industry and research institutions to commercialise AI research outcomes. Next, the Bill must be approved by the Chamber of Deputies for it to become law.

Netherlands: House of Representatives approves motion for age limits for social media platforms

Type Legislative (motion)
Status Passed

On 4 March 2025, the Dutch House of Representatives (Tweede Kamer) of the Dutch Parliament approved a motion which tasks the government to consider varying age limits for social media platforms. The motion highlights increasing issues around social media usage among children, including increasing signs of addiction and exposure to harmful content and recommends a minimum age of 15 to access social media. The motion also encourages the government to work with other countries to develop consistent regulations and enforcement of age restrictions across jurisdictions to safeguard young people.

Republic of Korea: Supreme Court issues KRW 6.7 billion fine on Meta

Type Litigation
Status Decision issued

On 13 March 2025, the Supreme Court of the Republic of Korea upheld a ruling by the Personal Information Protection Commission (PIPC). This ruling rejected Meta's legal challenge to overturn a fine of KRW 6.7 billion (around EUR 4,130,000) and an associated corrective order, which were imposed due to the alleged unauthorised sharing of personal data belonging to approximately 3.3 million South Korean Facebook users with around 10,000 external app operators. Meta is now obliged to implement the corrective measures set out by the commission.

Switzerland: Federal Data Protection and Information Commissioner concludes investigation into X

Type Regulatory (investigation)
Status Concluded

On 20 March 2025, the Federal Data Protection and Information Commissioner (FDPIC) concluded its investigation into X regarding the use of personal data for training its AI system Grok. Initiated in response to concerns raised in summer 2024, the inquiry assessed the transparency of data practices and user opt-out options. X appointed a representative in Switzerland and confirmed that it had implemented an opt-out mechanism on 16 July 2024, allowing users to exclude their public posts from training datasets. This aligns with the Federal Act on Data Protection (FADP), enhancing user control over data preferences. The FDPIC determined that X is now compliant with the relevant data protection regulations.

United Kingdom: Apple takes legal action against the UK Government on encryption backdoor

Type Litigation
Status Ongoing

On 7 April 2025, the Investigatory Powers Tribunal confirmed that Apple is contesting a directive from the UK Government which seeks to provide backdoor access to its encrypted cloud storage service. The court revealed the existence of this legal

dispute after rejecting a request from the Home Office to conduct the proceedings in private. In February 2025, Apple stated its intention to cease offering the end-to-end encryption feature for UK users following reports of the Government's demand, which raised concerns about potential electronic surveillance by national security officials.

United Kingdom: Ofcom fines OnlyFans GBP 1.05 million for non-compliance with age verification duties

Type Regulatory (enforcement)
Status Decision issued

On 27 March 2025, Ofcom imposed a GBP 1.05 million (around EUR 1,230,000) fine on Fenix International Limited, the owner of OnlyFans, for failing to comply with age verification requirements. After requests for information in June 2022 and June 2023, Fenix reported inaccurately on its age assurance measures. Specifically, it misrepresented the age setting of its facial estimation technology, stating it was set at 23 years when it was actually set at 20 years since November 2021. This information affected Ofcom's October 2022 report on video-sharing platform regulations. Although Fenix was informed of the discrepancy by its technology provider on 4 January 2024, the company did not notify Ofcom until 22 January 2024, more than two weeks later. As a result of its failure to provide accurate and timely information and the delay in reporting the error, Ofcom launched its investigation in May 2024.

United Kingdom: Illegal content duties under Online Safety Act enter into force

Type Regulatory (enforcement)
Status In force

On 17 March 2025, the illegal content duties under the Online Safety Act came into force, mandating that in-scope online platforms implement robust measures to protect users from illegal content. These obligations require services to promptly remove illegal content upon detection and to proactively reduce the likelihood of 'priority' criminal material being disseminated. With enforcement led by Ofcom, platforms must demonstrate compliance through suitable illegal content risk assessments and may face stringent penalties for non-compliance, including significant fines or court action. Certain large service providers, as well as providers of smaller but potentially harmful services, must submit these assessments to Ofcom by 31 March 2025.

United Kingdom: Ofcom opens consultation on guidance on 'a safer life online for women and girls'

Type Regulatory (consultation)
Status Open

On 25 February 2025, Ofcom opened a consultation on draft guidance to improve the online safety of women and girls. The nine areas covered in the draft guidance include conducting risk assessments, increasing transparency, setting safer defaults, and giving users more autonomy over their online experience. The guidance interlinks with existing Codes and guidance on risk assessments for addressing illegal content, as well as the current draft Codes and guidance on the protection of children, while also outlining additional voluntary steps that services can take to further mitigate risks to women and girls online. The consultation is open until 23 May 2025.

United States: TAKE IT DOWN ACT proceeds to vote in the House of Representatives

Type Legislation
Status Pending House of Representatives vote

On 8 April 2025, the House Committee on Energy and Commerce advanced the TAKE IT DOWN Act with a 49–1 vote. The Act seeks to make it a criminal offence to distribute non-consensual intimate deepfakes and would require online platform providers to remove such content under threat of enforcement by the Federal Trade Commission (FTC). The Senate had previously passed the legislation with unanimous consent in February. The next step involves the TAKE IT DOWN Act moving to the House of Representatives floor for a vote, after which it could be sent to the President for approval. In his address to Congress in March 2025, President Trump expressed his support for the legislation.

United States: Federal court grants preliminary injunction against California Age-Appropriate Design Code Act

Type Temporary Injunction
Status Implemented while case is under review

On 13 March 2025, the U.S. District Court for the Northern District of California granted a preliminary injunction, preventing the enforcement of the California Age-Appropriate Design Code Act (CAADCA) in NetChoice v Bonta. CAADCA was passed in California in 2022, requiring online platforms that are likely to be accessed by children to conduct a survey to assess potential risks and limit the collection of minors' personal information, among other duties. The law has faced legal challenges, with opponents arguing it violates the First Amendment because it imposes content-based speech regulation. The law is now blocked until a judge can review its constitutionality. Parts of the legislation were originally set to start on 5 April 2025.

United States: President Trump signs memorandum to investigate foreign digital services taxes and fines on American companies

Type Executive Order
Status Implemented

On 21 February 2025, President Donald Trump signed a memorandum titled "Defending American Companies and Innovators From Overseas Extortion and Unfair Fines and Penalties". The memorandum directs the Administration to investigate whether "any act policy or practice in the European Union and the United Kingdom incentivises U.S. companies to develop or use products and technology in ways that undermine free speech or foster censorship". The Trump administration will consider "responsive actions" including tariffs if countries utilise digital service taxes or fines against American companies.

United States: Federal Trade Commission opens inquiry on censorship practices by technology platforms

Type Public inquiry
Status Open

On 20 February 2025, the Federal Trade Commission opened a public inquiry into how online platforms deny or deplatform users based on the content of their speech, and how this could have violated U.S. law. The FTC specifically encourages "tech platform users who have been banned, shadow banned, demonetized, or otherwise censored" to share a comment. The FTC is accepting comments until 21 May 2025.

Section 2 Topic-specific snapshot: "Algorithmic Recommender Systems: Assessing Risks and Potential Mitigation Measures"

This section summarises selected analyses and responses published by government agencies, civil society organisations and academia on algorithmic recommender systems, their risks and potential mitigation measures.

Pulling Back the Curtain: An Exploration of YouTube's Recommendation Algorithm,
Aoife Gallagher, Lucy Cooper, Rhea Bhatnagar and Cooper Gatewood, Institute for Strategic Dialogue, 2024

This report by ISD summarises a series of four investigations into YouTube's algorithmic recommendations and outlines policy recommendations for platforms. In the investigations, ISD analysts examined the video recommendations presented to US-based YouTube accounts over a period of one month. The accounts represented varying profiles, differing in age, gender, and interests. As 70% of all video views on YouTube stem from recommendations, each of the four investigations aimed to understand the type of content recommended to specific profiles.

Analysts found no conclusive evidence to suggest that age and gender of profiles impact the type of recommended content. YouTube's recommender system thus appears indifferent to age, assigning it low weighting importance in the recommendation process. However, this can result in younger users being exposed to inappropriate and potentially harmful content, including videos related to self-harm and suicide. This also underscores the unknown variables that recommender systems operate upon. Additionally, analysts discovered that YouTube recommends a considerable number of videos unrelated to the profile's interests, which may lead to greater exposure to harmful content, including health misinformation and misogynistic material by influencers such as Andrew Tate, who has been banned from the platform. The report also notes that Christianity-related content was recommended across all four investigations, raising questions about how YouTube's recommendation system accounts for religion when generating suggestions.

Initial Analysis of the First Round of Risk Assessments Reports under the EU Digital Services Act,
DSA Civil Society Coordination Group, 2025

This report was published by the DSA Civil Society Coordination Group, an informal coalition of civil society organisations, including ISD, academics, and public interest technologists. The group advocates for the protection of international human rights, adherence to the rule of law, and the implementation of human rights due diligence concerning the development, application, and enforcement of the DSA. The report provides an initial, high-level analysis of the first round of published risk assessment reports under the Digital Services Act (DSA).

By focusing on Google's risk assessments for its services, specifically Google Search and YouTube, as well as those of Meta for Facebook and Instagram, TikTok, and X, the report delineates four key aspects. These include best practices that have emerged from the risk assessment reports, the necessity for transparency and data disclosure to foster trust, the importance of meaningful stakeholder engagement, and the significance of design-related risks, which the report suggests the selected risk assessment reports insufficiently address. Algorithmic recommender systems are

identified as one such design-related risk. However, the authors of the report note that while the risk assessment reports (indirectly) acknowledge these risks, the primary focus remains on user-generated and content-related risks.

Overall, the report finds that the selected risk assessment reports offer little consideration of the integral role that platform design, including recommender systems, play in the creation of risks. Instead, they tend to position these systems as potential mitigation measures. This approach downplays the research evidence that highlights how such design features may contribute to the risks analysed, including those specifically related to the amplification of harmful content.

Finally, the report presents three key recommendations in its conclusion to promote public interest, foster trust, and ensure effective compliance in future risk assessment reports:

1. There should be a stronger emphasis on the risks inherent in **platform design,** particularly regarding recommender systems that can exacerbate harmful content.

2. Platforms must enhance **transparency** by providing verifiable data that demonstrates the effectiveness of their mitigation measures, along with improved reporting formats for greater clarity.

3. Meaningful **engagement with stakeholders** is essential; platforms should consult civil society, researchers, and affected communities, and incorporate independent research to ensure that risk assessments accurately reflect user experiences and enhance accountability.

Better Feeds: Algorithms That Put People First. A How-To Guide for Platforms and Policymakers,
Expert Working Group on Recommender Systems, Knight Georgetown Institute, 2025

The report prepared by the KGI Expert Working Group on Recommender Systems offers policy recommendations aimed at enhancing recommender systems for lasting user benefit and improved experiences. Drawing from a diverse research background and industry expertise, it identifies significant challenges in the current design and regulation of these systems while suggesting practical solutions for policymakers and product designers.

A primary concern highlighted is that some platforms prioritise engagement metrics, such as clicks and likes, which may overshadow more meaningful indicators of long-term user value. This focus on short-term engagement aligns with the financial interests of tech companies that earn revenue through advertising, often leading to detrimental effects like the proliferation of low-quality information, increased user dissatisfaction, excessive use, and heightened societal polarisation. The evidence presented in the report underscores the urgent need to transition toward designs that foster long-term user satisfaction and societal well-being. To facilitate this shift, the Working Group proposes several key recommendations:

1. **Transparency in Design:** Platforms should openly disclose the input data, value model weights, and metrics used to assess long-term user value, along with the internal criteria employed to evaluate product teams in charge of recommender system design.

2. **User Choices and Defaults:** Users must be given straightforward options to select from various recommender systems, ensuring at least one option is designed for long-term benefit. Minors should receive default systems prioritising their long-term value, or default to non-personalised systems if there's insufficient data.

3. **Long-Term Holdout Experiments:** Platforms should conduct ongoing experiments that last at least 12 months, assessing the long-term impacts of algorithm changes. These results should be aggregated and made publicly available, with independent audits for accountability.

The report serves as a guide for both policymakers and product designers, promoting algorithmic systems that prioritise users' long-term interests while still maintaining the ability to personalise content. The authors make these recommendations for stakeholders to address the negative impacts of current recommender systems and enhance user experiences and societal value.

Auditing Recommender Systems. Putting the DSA into practice with a risk-scenario-based approach,
Dr Anna-Katharina Meßmer and Dr Martin Degeling, Interface, 2023

This report focuses on the evaluation of recommender systems in the context of the DSA. It emphasises the DSA's aims to enhance transparency and accountability among large online platforms by imposing due diligence obligations on their algorithmic systems. However, the authors point out that the DSA lacks guidance on how to operationalise abstract systemic risks. To address this issue, the report proposes a risk-scenario-based audit process consisting of four steps:

1. **Plan:** Understand the platform and identify relevant stakeholders, including developers, researchers, legal experts, and affected parties.

2. **[Define] scenarios:** Identify and prioritise specific issues related to systemic risks, breaking down abstract concepts like 'mental well-being' into concrete, testable scenarios. Such scenario details the affected party (individual/group/institution) defined by a characteristic, experiencing harm in relation to the platforms' involvement and the potential broader consequences (macro impact).

3. **Measure:** Develop and prioritise measurement methods to assess the scenarios effectively, using different kinds of measurements, such as crowdsourced or automated methods.

4. **Evaluate:** Analyse results and produce an audit report that promotes reproducibility and suggests mitigation measures. This structured auditing approach aims to ensure the DSA is effective in evaluating platform recommender systems.