



An ODI European Data and AI Policy Manifesto

July 2025

Contents

Contents	1
About this paper	2
Introduction	3
The ODI European Data and AI Policy Manifesto - our six principles	5
Principle 1: Strong data infrastructure	5
Principle 2: Open data as a foundation	7
Principle 3: Trust	8
Principle 4: Trusted, independent organisations	10
Principle 5: A diverse, equitable and inclusive data ecosystem	11
Principle 6: Data knowledge and skills	12

About this paper

This policy paper – an ODI European Data and AI Policy Manifesto – was prepared by the Open Data Institute (ODI), a UK-based organisation with an international perspective. The ODI is a non-profit research institute founded in 2012 by Sir Tim Berners-Lee and Sir Nigel Shadbolt. Its vision is for a world where data works for everyone. This work has been conducted independently but is sponsored by Google, and we thank them for their support.

The ODI is headquartered in London, but we have always been interested in policy globally, including in Europe. For example, we participated in three of the four working groups involved in developing the EU's General-Purpose AI Code of Practice; we recently submitted [feedback on the Code](#) and have contributed to [previous consultations](#). A recent ODI project looked at the [use of secondary health data across Europe](#), and we are contributing, with Google and more than 30 other organisations, to the [Croissant metadata format](#), which simplifies how ML models use data. We are a member of the [Big Data Value Association](#), an industry-driven members association with the mission to promote European innovation in the data and AI ecosystem.

In preparing this paper, the ODI explored the EU regulatory and policy landscape and applied a principles-based approach to our EU policy recommendations. These principles were set out in the [ODI's five-year strategy](#), published in 2023. The ODI [launched](#) a [manifesto](#) for data and AI policy in the UK in March 2024 ([updated in September, following the UK general election in 2024](#)). We have focused on similar topics in this European version. This means that some topics – for example, copyright and the broader intellectual property regime – are not included here (but [are covered elsewhere](#) in ODI's research and consultation responses).

As the EU takes forward its ambitious agenda on data and AI, we recognise the depth of expertise, careful deliberation, and democratic processes that underpin EU policymaking. Our aim in this paper is not to prescribe or direct, but to offer constructive reflections through the lens of key principles that hold relevance globally, including within the EU.

Our perspective is informed by our interest in data and AI policy at an international level. While we are based outside the EU, we share many of the same aspirations, including fostering innovation, safeguarding fundamental rights, ensuring responsible data stewardship, and supporting the development of AI that serves society. In this spirit, we offer these perspectives to contribute to the ongoing conversation about effective implementation, fully respecting the EU's sovereignty in shaping its regulatory and policy landscape.

Introduction

2025 is a pivotal year for the future of Europe. As the world grapples with the transformative power of data and the AI it feeds, the EU can shape a future where these technologies and their applications benefit citizens. Yet inaction or overreaction risks harming citizens and society and leaving the continent's economy behind its competitors.

The rapid rise of AI, fuelled by vast datasets and increasingly complex algorithms, presents immense opportunities and significant challenges.

2025 brought a new Commission, which faces the task of promoting innovation to ensure Europe's global competitiveness while safeguarding the rights and values that define the EU's relationship with its citizens. Building on last year's [Draghi report](#) on the future of European competitiveness, the Commission starts with a new [Competitiveness Compass](#). This 'sets a path for Europe to become the place where future technologies, services, and clean products are invented, manufactured and put on the market, while being the first continent to become climate neutral'. It increasingly frames its ambition to 'reverse the trend' of the EU not keeping pace with other economies within the context of European 'tech sovereignty' amid global geopolitical shifts.

The year ahead will also test how the various overlapping pieces of legislation and regulation on data and AI will work together. We summarise some of the key initiatives in an annexe and await the outcome of a review of the digital package of legislation and regulation announced in the Commission's latest work programme.

The world is also contending with a challenging geopolitical situation. The Trump Presidency in the US, its rivalry with China, and conflicts around the globe – including in Europe – bring uncertainty. Tech and AI policy will be both affected by, and shape, these global trends.

As the EU seeks to harness AI's immense potential to drive economic growth and address societal challenges, it must do so in a manner that protects citizens' rights, ensures fairness, and maintains public trust.

This manifesto aims to support these efforts. It challenges all those in the European data ecosystem to leverage the Union's collective strength to pioneer a distinctly European, data-centric and people-centric model of AI development. It should balance innovation with responsibility, economic growth with societal well-being, and technological advancement with unwavering respect for human rights and European values.

Furthermore, as the EU seeks to ensure its digital acquis works effectively in practice, the principles outlined below – focusing on strong and open infrastructure, trust-building through participation and assurance, independent scrutiny, equity, and skills – offer a framework to support coherent and effective implementation, complementing the Commission's planned assessment.

The approach should support a strong and ethical data ecosystem to underpin AI development, ensure transparency and accountability, empower citizens to help shape how their data is used, develop skills and data literacy, and address global challenges through international cooperation.

This approach, grounded in the core values of openness and trustworthiness, can help address pressing challenges and create a future where data – and AI – work for everyone.

The ODI European Data and AI Policy Manifesto - our six principles

Principle 1: Strong data infrastructure

*The ODI believes that a **strong data infrastructure** is the foundation for building an open, trustworthy data ecosystem on a global scale, and that this can help address the world's most pressing challenges.*

As part of this, we would like to see:

A legislative and regulatory regime that is future-proof and has people and society – as well as economic competitiveness – at its heart

- Headline EU initiatives – including the Competitiveness Compass – emphasise innovation and competition. However, regulation, protecting people and their data, and building trust in the use of data and AI, are also crucial. Innovation and regulation should not be seen as a trade-off, but as complementary necessities to ensure growth is purposeful, directional, aligns with European values and needs, and supports the EU's ambitions for greater competitiveness and technological sovereignty in a way that benefits society.
- Achieving this will involve building on existing measures. These include:
 - strong data governance frameworks that ensure interoperability and that can support AI development while protecting citizens' rights and European values (like [Common European Data Spaces](#) and [Gaia-X](#));
 - ensuring broad access to high-quality, well-governed data from both public and private sectors (the intention of the Data Act, Data Governance Act, Data Spaces, the Open Data Directive, and the [AI Continent Action Plan](#)) that is [AI-ready](#) to foster a diverse, competitive market for AI technologies;
 - a tiered compliance model (like that in the EU AI Act) that scales requirements based on an organisation's size and the risk level of their AI applications, reducing burdens on SMEs while ensuring comprehensive coverage of AI-related risks;
 - the European Data Protection Board (EDPB) continuing to consider [data protection in the context of AI](#), perhaps developing guidelines

for AI-appropriate data minimisation;

- and continuing to support investment in data and computational infrastructure, including secure, scalable data storage and processing capabilities across the EU, to support greater technological sovereignty.
- One theme from our research is uncertainty about how different parts of the legislative landscape fit together (for example, the GDPR and Data Governance Act regimes) and a lack of clarity around definitions and how legislation might be enforced (especially where Member States differ in approach). This reflects in part the challenge of striking the right balance between providing certainty, especially for businesses, while keeping rules flexible. A related theme is the need for the Commission to monitor and measure how different pieces of legislation and regulations are operating and whether they are working as intended against a definition of success. The Commission's latest work programme, '[A Simpler and Faster Europe](#)', includes an assessment of the digital acquis. This provides an opportunity to review how well existing legislation is working together, explore any contradictions, clarify any uncertainties, identify gaps, and critically assess practical implementation effectiveness and enforcement challenges across Member States.
- The Commission has already confirmed that the assessment of the digital package will include GDPR, with [a focus on SMEs](#). The Commission should also consider strengthening the rights of the public, communities and society - for example, in our UK policy manifesto, we called on the UK government to consider collective as well as individual rights, and we have [also called](#) for data protection impact assessments to consider group- and societal-level impacts as well as individual ones.

Principle 2: Open data as a foundation

*Strong data infrastructure includes data across the spectrum, from open to shared to closed. **The best possible foundation is [open data](#)**, supported and sustained as data infrastructure. Only with this foundation will people, businesses and governments be able to realise the potential of data infrastructure across society and the economy.*

As part of this, we would like to see:

Further support for opening up data

- We suggest that the Commission [continues to examine](#) the effectiveness of its current measures for opening up data (as outlined in ‘[A European strategy for data](#)’ and supported by platforms like [data.europa.eu](#)), and build on the experience of Data Spaces, the Open Data Directive, and recent measures in the Data Act and Data Governance Act. This could involve acting upon the results of the recent survey on High-Value Datasets, giving greater clarity and exploring further incentives for private sector organisations to share non-sensitive data under the Data Act.
- We recommend that the Commission considers what support might be available for data altruism organisations that allow individuals and organisations to voluntarily share data for the benefit of society, and data intermediaries, which allow data sharing, under the Data Governance Act. Currently, [only one organisation is registered](#) as a data altruism organisation, and there are [data intermediaries registered](#) in only four Member States. The Commission should seek to understand why this is the case and take action accordingly.
- The Commission and European AI Office might also explore how best to ensure [broad access](#) to data for training AI models (potentially leveraging regulatory sandboxes to test approaches), which might include incentives and other activities to create and regulate high-quality datasets and infrastructure that enables AI systems to use data responsibly, exploring new approaches to open up access to, and steward, public data; and ensure that new AI initiatives pay more attention to the role of data (for example, the [new OpenEuroLLM](#) focuses on models). The Data Labs recently announced in the AI Continent Action Plan are a step in the right direction, but details on what data they will entail and how this data will be made available across the user ecosystem are yet to be defined.

Principle 3: Trust

*For data to work for everyone, it needs to work across borders – geographic, organisational, economic, cultural and political. **For this to happen ethically and sustainably, there needs to be trust** – trust in data, and trust in those who access and share it.*

As part of this, we want to see:

A more participatory data future

- We need to move beyond transparency and accountability to a world [where people can meaningfully participate](#) in how data is used by the government and others, empowering individuals and communities to help shape how data is used for society, the environment, the economy, and the public good.
- Nonetheless, transparency reporting remains important in its own right and to underpin meaningful participation. In [our submission](#) to the development of the General Purpose AI Code of Practice, we emphasised the need to expand the scope of documentation requirements to include a range of other categories, such as channels through which datasets have been distributed and the full legal entity name of any organisation involved in data annotation or other outsourced preprocessing services, and detailed transparency about the characteristics of training datasets. We also stressed the importance of clear, machine-readable metadata in verifying results and judging risks of models, and the need for the Code's guidelines to ensure datasets adhere to new metadata standards specifically for AI datasets (like Croissant). Open standards would help make data – and metadata – widely available, and enable collaboration across the whole ecosystem
- We recommend that the European AI Office monitors the operation of the transparency regime under the EU AI Act, producing guidance and standards as necessary (the Commission review of the digital package provides an opportunity to evaluate transparency reporting across other regulations). The Office could also explore supporting an AI data transparency index ([like the one developed by the ODI](#)) and consider developing a comprehensive AI incidents database to inform policy and improve AI system design or work with existing ones (such as the OECD) to ensure consistency in reporting.
- Rights and protections are also critical for building trust, especially as the development of AI accelerates. In particular, the Commission might

consider how best to recognise data supply chains and [protect and enforce people's data protection and labour rights](#) within them.

- We recommend that the Commission should promote the use of privacy-enhancing technologies (PETs) such as federated learning and differential privacy in AI development to retain compliance with the data protection principles and maximise the use of data while reducing risks inherent to data use. Open standards for managing digital identities and storing personal data for reuse (like [Solid](#)) also have an important role to play in building trust. There may be opportunities to set global standards, while establishing an EU fund to support research and development of PETs could help firms comply with legal requirements while opening up more data.

Investing in assuring data practices across the economy

- Data is now fundamental to all organisations, but to unlock its true value, organisations must be confident in their [data practices](#). At the ODI, we believe we need to go beyond the current legal requirements to build trust in data and unlock the significant economic benefits of data sharing through [data assurance](#).
- The European AI Office might consider what an AI assurance framework, building on the AI Act and incorporating standards for certifying high-risk AI systems, might look like. Developing world-leading standards for AI assurance and auditing could help position the EU as a global leader in responsible AI development and develop a new sector of the data and AI economy.

Principle 4: Trusted, independent organisations

*There is a greater need than ever for **trusted, independent organisations** to help people across all sectors, economies and societies to benefit from better data infrastructure.*

As part of this, we would like to see:

Strategic funds to support organisations working to ensure we all benefit from the use of data, and scrutinise and challenge the use of data (including in AI)

- Credible, independent civil society organisations play a crucial role in our data ecosystem. They can convene actors from across sectors, develop research and policy for beneficial uses of data (and data-centric AI), and scrutinise (and challenge) the use of data and AI by powerful organisations in government and other sectors. These organisations require support to survive and continue to provide this essential role for the benefit of everyone in the data and AI ecosystem.
- The previous US administration showed what could be possible. In November 2023, it announced an initiative with philanthropic organisations, with more than \$200m to ‘advance AI designed and used in the best interests of workers, consumers, communities, and historically marginalised people in the United States and across the globe’. The Commission should consider how to support civil society organisations and establish accessible long-term funding mechanisms for independent organisations working on responsible AI and data governance.
- The Commission might also explore support for data altruism organisations and data intermediaries (see Principle 2), and for publicly funded incubators for data for AI projects. These could even foster partnerships between different sectors, building on examples such as the [Open Data Incubator for Europe](#) (ODINE) - and [Data Pitch](#).

A clear commitment to including civil society in decision-making on the use and scrutiny of data and AI

- We recommend that the Commission and other bodies take decisive action to include diverse civil society voices in data and AI decision-making, reflecting the EU's focus on public services and accountability. This could involve ensuring diverse representation from civil society, academia, and industry on existing bodies' boards and advisory panels. It could also include understanding what participatory methods work in bringing in different voices and providing central support, guidance and infrastructure for such initiatives.

Principle 5: A diverse, equitable and inclusive data ecosystem

*For data to work for everyone, those collecting and using it need to be highly alert to inequalities, biases, and power asymmetries. AI organisations working in data must take proactive steps to ensure that they contribute fully and consciously to creating a **diverse, equitable and inclusive data ecosystem**.*

As part of this, we would like to see:

Ensuring equitable access to data for growth and innovation purposes

- For the data ecosystem to work for everyone, we need as many sectors, communities and people as possible to access the data, data infrastructure, and skills they need. This is particularly important for developing AI systems and use cases with significant societal implications, such as misinformation, climate, and infectious diseases, [where fair and equitable access must be mandated](#).
- It is also critical for competition and economic growth that SMEs and start-ups can compete with big tech firms that will benefit from [data asymmetries and network effects](#) (where being able to link vast amounts of data leads to the extraction of more value). Providing equitable access to data for growth and innovation purposes is the intention of several recent European regulations - the current review of the digital acquis offers an opportunity to assess whether they are working as intended. The AI Factories and their Data Labs referenced in the AI Continent Action Plan need to ensure they cater for SMEs and start-ups.
- We note some of the work already happening that could be built upon further, including the [EU Startup and Scaleup Strategy](#), the [GenAI4EU project](#) ('Creating European Champions in Generative AI'), the [European Digital Infrastructure Consortium \(EDIC\)](#), and the [InvestAI initiative](#).

Greater researcher access to data from social media platforms

- In our UK policy manifesto, we underlined the importance of researchers being able to access data from social media platforms to understand online harms better, particularly those affecting marginalised communities and people with particular characteristics. We cited section 40 of the EU's Digital Services Act, which allows vetted researchers access to data from large online platforms and search engines, as a precedent. The Commission has [drafted a delegated act](#) setting out the specific conditions. As these provisions come into force, the digital services coordinators in each Member State should monitor how well the new regime is working - not only to improve researcher access to data in the EU, but to continue to help shape international norms and approaches.

Principle 6: Data knowledge and skills

*The world needs a new cohort of data leaders – individuals who **have data knowledge and skills and are equipped to understand the value, limitations and opportunities offered by data, data practices and data sharing.***

According to [an overview of digital literacy](#) in the EU, published by data.europa.eu, ‘more than 90% of jobs in Europe require basic digital knowledge alongside traditional skills like literacy and numeracy’, but ‘32% of Europeans still lack basic digital skills’. The EU’s digital strategy also notes that [digital skills are ‘key’](#) to ‘ensuring technological sovereignty, bridging the digital divide and promoting a human-centred digital environment across the EU.’ The [Digital Decade Compass](#) directly recognises skills as one of its four focuses with a target of 80% of citizens having at least basic digital skills and the EU having 20 million ICT specialists.

In our work looking at [approaches to data literacy](#) in the UK, we highlighted the AI Council’s three-pronged approach on skills and diversity, which focused on developing high-level skills in data and AI, promoting general data and AI literacy across the population, and prioritising inclusion and diversity, such that those using and making decisions about data and AI come from a range of backgrounds and experiences. That approach should inform all data and AI skills initiatives.

As part of this, we would like to see:

Greater data literacy and skills among leaders, including in business and the public sector

- We suggest that the European AI Office should be empowered to take the lead in developing professional standards for those working in data and AI.
- The European AI Office could also then work with others to ensure existing EU-wide digital skills initiatives, such as those under the [Digital Skills and Jobs Platform](#), adequately address data and AI literacy for policymakers and regulators, helping them understand the field, leading to better policy and more effective implementation and enforcement of existing and future legislation.
- The Commission could draw on the expertise of the European AI Office and others to enhance the Digital Education Action Plan (2021-27), which ‘aims to support the adaptation of the education and training systems of Member States to the digital age’, to further develop data and AI skills components, and include a focus on retraining and upskilling workers affected by AI-driven automation.

Greater data literacy among the public

- Beyond the [Digital Education Action Plan](#), the Commission might consider how best to integrate data and AI literacy into educational curricula across all levels, from cradle to grave.
- It might also consider launching a public awareness campaign to improve the general understanding of data and AI and its societal implications.
- As per building a more participatory data future (within Principle 3), people having more of a say in the sharing and use of data for AI would foster citizen involvement and improve AI literacy across society.
- We suggest a particular focus on involving and supporting historically underrepresented groups; otherwise, those making decisions with, and about, data and data-driven technologies will continue to come from the same privileged backgrounds, with marginalised groups most likely to be harmed.