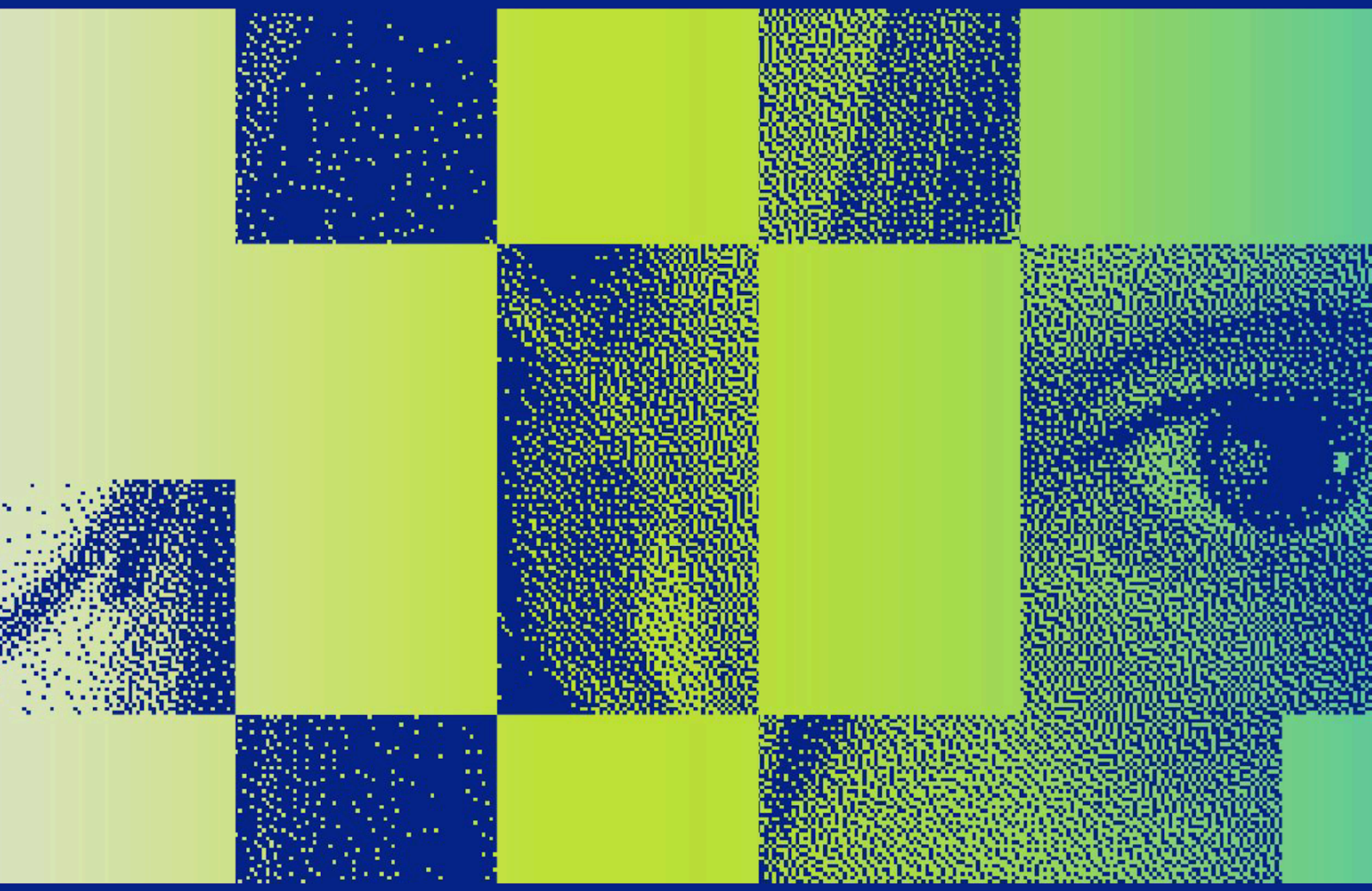

Building a better future with data and AI: a white paper

ODI Research
ADVANCING TRUST IN DATA

Contents

Executive summary

This white paper outlines the Open Data Institute's (ODI) vision for artificial intelligence (AI) in the UK, emphasising the need for robust data infrastructure, governance and ethical foundations to support the tech ecosystem. Much attention has been paid to the existential risks of AI, many of which relate to so-called AGI (artificial general intelligence) technologies, which are yet to be developed. Eighteen months into this 'AI Summer', these existential risks have yet to materialise, and experts disagree on when — and indeed, if — we will see AGI. Existing and imminent generative AI systems and large language models (LLMs) like ChatGPT, Llama, Palm and Claude present immediate opportunities and challenges. They require a sound AI infrastructure.

The ODI's 2024 research highlights significant weaknesses in AI infrastructure, particularly in data governance, with both present and future implications. AI's rapid evolution has led to overlooked data fundamentals, with governments and organisations neglecting critical data infrastructure. Data-centric AI is the response to this. As the issues become impossible to ignore, the community is slowly starting to work on the specific challenges around data quality, governance and data-centric model evaluations.

The ODI's research shows flaws that could cause model and business collapse, due to lack of data transparency and poor data-sharing practices threatening rights and livelihoods. If the UK is to benefit from the extraordinary opportunities presented by innovations in this field, it must look beyond the hype cycle and return to the data fundamentals that are so essential for building a robust data and AI ecosystem, and a strong, equitable market that benefits everyone, and minimises harm.

The ODI is one of the few civil society organisations working in the data-centric AI space in the UK today. In this white paper, the ODI presents its early findings, for the attention of the new UK Government and all those who have an interest in this topic — and who might like to contribute to our work — on how to place data at the centre and ensure the UK benefits from the AI revolution; growing economically, while mitigating any potential risks.

To achieve these outcomes, the ODI recommends five policy interventions:

- **Ensure broad access** to high-quality, well-governed public and private sector data to foster a diverse, competitive AI market;
- **Enforce data protection** and **labour rights** in the data supply chain;
- **Empower people to have more of a say** in the sharing and use of data for AI;
- **Update the intellectual property regime** to ensure AI models are trained in ways that prioritise trust and empowerment of stakeholders;
- **Increase transparency** around the data used to train high-risk AI models.

1: Realising the opportunities of AI with access to high-quality, well-governed data

Emerging AI technologies offer opportunities in many industries, from diagnostics and humanitarian mapping, to personalised education. To further these opportunities, there must be a focus on the data underpinning them.

Achieving the benefits from AI requires high-quality, well-governed datasets. While examples like the INSIGHT ophthalmic database drive social good through efficient data governance supporting medical research, there are major deficiencies in available data for purposes such as medical imaging for rare diseases. Compounding this problem, the ODI's detailed scoping review of data governance practices revealed a critical gap in the governance infrastructure supporting access to datasets throughout the development of AI systems.

Policy stakeholders worldwide are increasingly investing in accessible public datasets and supporting infrastructure for Responsible AI (RAI). For example, the UK's Competition and Markets Authority has called for greater access to AI model data. Similarly, the EU AI Act addresses data access by requiring providers of high-risk AI systems to ensure that national authorities have access to the data used to develop them. The newly introduced Product Safety and Metrology Bill in the UK allows for the use of the new product safety framework to introduce checks on those producing the most powerful AI models.

Beyond access, **good data governance requires resources to ensure data is and remains relevant**, and can be repurposed and linked to data from other sources. Without support, the rising price of high-quality AI training data excludes many potential innovators like small businesses and academic researchers. As proposed in the 2024 UK Labour Party manifesto, and supported by ODI's Executive Chair, Sir Nigel Shadbolt, emerging policy proposals to create a national data library should be extended to consider how the data could be made AI-ready to address these imbalances.

Recommended policy intervention: Ensure broad access to high-quality well-governed data, from both public and private sectors, to foster a diverse, competitive market for AI technologies.

Case study: Leveraging smart data and AI for societal benefits and innovation

Smart data refers to technologies that allow the safe sharing of customer data to design new digital services for economic growth and innovation. Smart data initiatives have emerged in several sectors in the UK, notably in banking, starting with Open Banking and more recently the intersectoral Smart Data Challenge Prize. Under these initiatives, customers can choose whether and how they share their data with 'named third parties' in exchange for more relevant, effective service choices.

The previous UK Government invested in AI-ready data through entities like the Office for AI, which worked to identify valuable datasets that the Government should publish as authoritative, open, machine-readable data for both public and private sector AI models. The previous UK Government also funded Smart Data Research UK and EPSRC AI Hubs, which benefited transport environmental solutions and planning.

Smart data has lots of potential applications. For example, in cities, in conjunction with AI technologies, it can enable a better understanding of mobility patterns and allow more tailored services to emerge. City planners can optimise transportation networks, Mobility Service Providers (MSPs) can improve their on-demand services, and local governments can make data-enabled decisions to reduce congestion and air pollution. This collaborative approach ultimately leads to a more sustainable and efficient urban transport system.

In July 2024, the new Labour administration in the UK announced at their first King's Speech the introduction of new smart data schemes via the Digital Information and Smart Data Bill, utilising newly established Digital Verification Schemes - digital identity products and services from certified providers that the Government wants to see introduced to help with things like moving house, pre-employment checks, and buying age-restricted goods and services.

2: The infrastructure to achieve AI benefits and mitigate harms

As the examples above demonstrate, making data available and AI-ready is foundational to delivering the value of AI effectively and responsibly, and **depends on the creation of a strong, reliable data infrastructure**. AI engineers spend 80% of their time preparing data to get it ready for use in AI applications. Recent investments in AI Hubs aim to make findings, technologies and data accessible and usable.

Data infrastructure is made up of key components such as data standards — eg W3C-recommended formats like RDF and JSON-LD — ontologies — eg Croissant, a metadata schema for datasets — and unique identifiers — eg Digital Object Identifiers - DOIs — that facilitate interoperability and reusability of data across multiple contexts.

Global advances in AI depend on numerous datasets, including those with potential data protection and labour rights issues. For instance, datasets like ImageNet and MS-Celeb-1M have faced scrutiny over privacy concerns and problematic labelling practices. ODI research has found that with a few exceptions, AI training datasets typically lack robust governance measures throughout the AI life cycle, posing safety, security, trust and ethical challenges.

Additionally, significant challenges related to interoperability between data sources must be addressed. In the UK, data quality issues and silos currently hinder AI applications in the NHS, as outlined in the case study below.

Case study: Data infrastructure at the heart of responsible AI use in the NHS

Health data, including electronic health records, medical imaging, genomics and wearable devices, is foundational for AI applications in the NHS. This data trains AI models to recognise patterns, make predictions and generate insights to enhance clinical decision-making and patient care. For example, natural language processing techniques applied to clinical notes and medical literature can support evidence-based practice within the health service.

While AI integration in the NHS promises benefits, the use of sensitive health data raises concerns about privacy and security. Ensuring patient consent and maintaining trust requires robust data infrastructure to support the ethical use of AI. This includes secure data storage, efficient data processing and strict adherence to data protection regulations. The importance of maintaining public trust is underscored by recent events: following the General Practice Data for Planning and Research (GPDPR) campaign in 2021, patient opt-outs from sharing health data nearly doubled, rising from 2.75% to over 5% of the total English patient population within just one month.

Examples of work to build patient trust and strengthen data infrastructure in health include INSIGHT, funded by Health Data Research UK, which uses anonymised eye scan data, overseen by a diverse Data Trust Advisory Board, co-developed by the ODI. The ODI also highlights the role of Privacy-Enhancing Technologies (PETs) in ensuring secure, ethical data access. For example, in a PETs explainer, federated learning — as demonstrated by the CURIAL-Lab team at Oxford University — enables AI models to be trained across multiple data sources to screen patients for COVID-19 without sharing data; enhancing privacy and collaboration. OpenSafely enables the linkage of patient health records in a Trusted Research Environment. Data analysts can use this capability to uncover patterns across a huge range of diseases, comorbidities and patient demographics. Well-curated data infrastructures are the foundations upon which AI capabilities and deployments must be based.

A substantial body of work outlines the requirements for robust AI-driven data infrastructure. Establishing AI-powered data institutions and governance structures is crucial for fostering trust in AI systems. Additionally, **enforceable protections for those impacted by AI-based data use** and the broader AI-enabled data supply chains are essential.

Hidden human labour in the data supply chain is an often-overlooked area that needs protection. Safeguarding labour rights is vital for advancing the UK's AI safety agenda. Privacy risks also persist throughout the data supply chain, necessitating stronger protections.

Data protection laws are designed to protect privacy as a fundamental right. Europe's General Data Protection Regulation (GDPR), for example, applies to all data that relates to an 'identifiable natural person', including names, pictures, identification numbers, and less obvious kinds of data like location information. Under the GDPR, the processing of personal data requires a clear legal basis and must be subjected to robust security measures.

Despite, or perhaps because of, its unprecedented data consumption, the Data Protection Enforcement Gap[1] regarding foundational AI must be avoided. During the previous UK administration, the ODI expressed disappointment that the then-proposed Data Protection and Digital Information Bill (DPDI) would have weakened transparency, rights and protections, and failed to recognise the importance of maintaining public trust. Twelve national data protection agencies, including the UK's Information Commissioner's Office (ICO), published a joint statement clarifying that even when personal information is publicly accessible, it is still subject to data protection laws and that mass scraping of personal information from the web to train AI can constitute a reportable data breach in many jurisdictions.

The G7 Hiroshima AI Process supports this endeavour, with guiding principles and a voluntary code of conduct that underscore the need for "appropriate data input measures and protections for personal data and intellectual property". In March 2023, the Italian data protection authority temporarily suspended ChatGPT over concerns about the processing of personal data to train the system. In the US, a Californian lawsuit claimed that OpenAI's foundational models have been illegally trained on "private conversations, medical data, and information about children". Moreover, researchers have shown that it is possible to cause ChatGPT to 'leak' data on which its underlying model has been trained, such as real email addresses and phone numbers, by using specific prompts.

As stated in a recent Policy Manifesto, **the ODI urges policymakers to redouble their efforts against the weakening of data protection rights**. Alongside discussing the risks of using personal data illegally to train AI models, steps must be taken to address the ongoing risks of models leaking personal data. New technologies have emerged in recent years that may help address these issues. Solid and other privacy-enhancing technologies have great potential to help safeguard people's rights and privacy as AI models become more prevalent.

[1] A wide disparity between the stated protections on the books and the reality of how companies respond to them on the ground.

Case study: Solid pod as a solution to AI-related data protection challenges

The Solid pod offers a promising solution to challenges presented by AI to data protection, particularly regarding personal data. Solid, led by ODI co-founder Sir Tim Berners-Lee, aims to give individuals control over their data by storing it in decentralised data stores called pods. These pods enable users to decide who can access their data, providing a mechanism to ensure data protection and privacy. By giving individuals, or groups, direct control over their personal data, Solid pods aim to mitigate the risks associated with the mass scraping of personal data from the web to train AI models, as users can grant or revoke access to their data at any time.

The decentralised nature of Solid pods aligns well with the principles outlined by the G7 Hiroshima AI Process, which emphasises the need for appropriate data input measures and protections for personal data and intellectual property. By storing data in decentralised pods, Solid ensures that personal information is not easily accessible for unauthorised use, reducing the likelihood of data breaches and unauthorised data scraping. This approach also supports compliance with data protection laws such as the GDPR, which require clear legal bases and robust security measures for processing personal data. Solid's model effectively empowers individuals to enforce their data protection rights, addressing the Data Protection Enforcement Gap.

Furthermore, Solid pods may help address concerns about AI models leaking personal data. Since users control access to their data, they can ensure that only trusted entities can use their information, reducing the risk of inadvertent data leaks through AI models. This approach enhances data security and helps maintain public trust in AI technologies by ensuring transparency and accountability in data usage.

Recommended policy intervention: Enforce people's data protection and labour rights in the data supply chain.

3: Prioritising trust and empowerment

Building trust in AI requires the engagement and empowerment of affected individuals. Additionally, by increasing AI literacy and engagement, a society can be created where deep fakes and other misleading technologies are less effective and easier to detect. Generative AI tools like ChatGPT and Claude can enhance digital literacy and improve public data infrastructure by automating the collection, processing, and analysis of data.

Citizen involvement in initiatives like OpenStreetMap and Wikipedia has created trusted foundational datasets. On a policy level, the AI Safety Institute should include broader participation; a concept publicly supported by ODI's Executive Chair Sir Nigel Shadbolt, following various initiatives to engage more individuals in public sector data and AI projects.

It is essential to **promote and engage with strong, independent civil society organisations** to represent the interests of the public across this domain — including those organisations with expertise in the delivery of participatory approaches — to match the powerful voices of government and big tech.

Empowerment relates not only to the general public but also to professional communities. For example, in the creative industries, high-profile lawsuits and industrial action demonstrate what collective action can achieve. The protection of intellectual property is only one part of ensuring trustworthy data and an AI governance regime that protects public interests.

Recommended policy intervention: Empower people to have more of a say in the sharing and use of data for AI.

Case study: Intellectual property in the creative industries

Alongside their wider benefits to UK society, the UK's creative industries contribute £108 billion to the economy annually. As AI technologies advance, questions about copyright and intellectual property (IP) rights for data used in AI development are emerging, particularly in music, visual arts and literature.

The UK Intellectual Property Office (IPO) conducted a consultation on AI and IP rights in 2021, recognising the need to adapt existing frameworks, and exploring topics such as a licensing regime or exceptions to copyright for text and data mining (TDM). Current UK law restricts unauthorised TDM of copyrighted works, with exceptions limited to 'non-commercial' research. Balancing innovation with protecting creative professionals' rights and livelihoods remains crucial.

AI challenges are emerging in real-time, and current AI has provided a shock to the IP system highlighting the need for legal infrastructure that addresses these issues, especially in creative industries. For UK creative industries, updating the IP regime to ensure fair AI model training is crucial.

The Getty Images vs. Stability AI case, where Stability AI allegedly used over 12 million Getty photos without permission, exemplifies the risks of AI potentially exploiting creative content without compensation. This case underscores the need for robust IP protections to prevent misuse and ensure fair compensation for artists.

Recommended policy intervention: Update the intellectual property regime to ensure AI models are trained in ways that prioritise trust and empowerment of stakeholders.

As highlighted in this white paper, growing concerns about AI systems necessitate broader intervention. Policies should address incidents like harmful deep fakes and inequalities in the justice system. However, as discussed in the ODI's Policy Manifesto, UK data and digital policy have recently prioritised industry growth, often neglecting community concerns.

A key policy concern is the consensus that opaque models require significant intervention. EU law now mandates transparency for most AI models, responding to global calls for policy change. Users, deployers and AI practitioners must know when to trust AI predictions and when to use their own judgement, supporting system accountability and research.

Recent research has demonstrated the extent of poor AI transparency, identifying transparency about data as lagging significantly behind most other areas. ODI research highlights that key transparency information about data sources, copyright and inclusion of personal information and more is rarely included by AI systems flagged within the Partnership for AI's AI Incidents database.

The ODI is developing an 'AI data transparency index' to highlight and monitor the state of AI data transparency and incentivise improved practice. To address current harms and restore trust, emerging international transparency efforts must be strengthened along with introducing stringent governance and explainability practices.

Recommended policy intervention: Increase transparency around the data used to train high-risk AI models.

Concluding remarks

Building a better future with data and AI requires a collective commitment to ethical, transparent and innovative practices. This white paper outlines the critical steps required to strengthen the data infrastructure on which AI is built, enabling it to reach its full potential as a force for economic and social good in the UK; from ensuring broad access to high-quality, well-governed data for key datasets, to enforcing data protection and intellectual property rights.

Sectors such as healthcare, finance and transport have significant AI opportunities. But individuals and communities need to be empowered to ensure they have a say in how data is used in these sectors and beyond. While transparency alone doesn't build trust, ethical interventions and inclusive practices are essential in fostering public confidence and driving innovation. The ODI is committed to advancing these principles, conducting research and advocating for policies that balance innovation with the rights of data providers and users.

To ensure that the journey towards making data AI-ready benefits everyone, the ODI invites your comments and insights on the findings and recommendations presented in this white paper. In the next phase of this work, the ODI will investigate designing and implementing strategies for overcoming risks and maximising the benefits of new AI technologies. The ODI will also continue to foster partnerships that enable impactful AI innovation.

The ODI looks forward to working together to create a robust data ecosystem that maximises the benefits of AI for all.