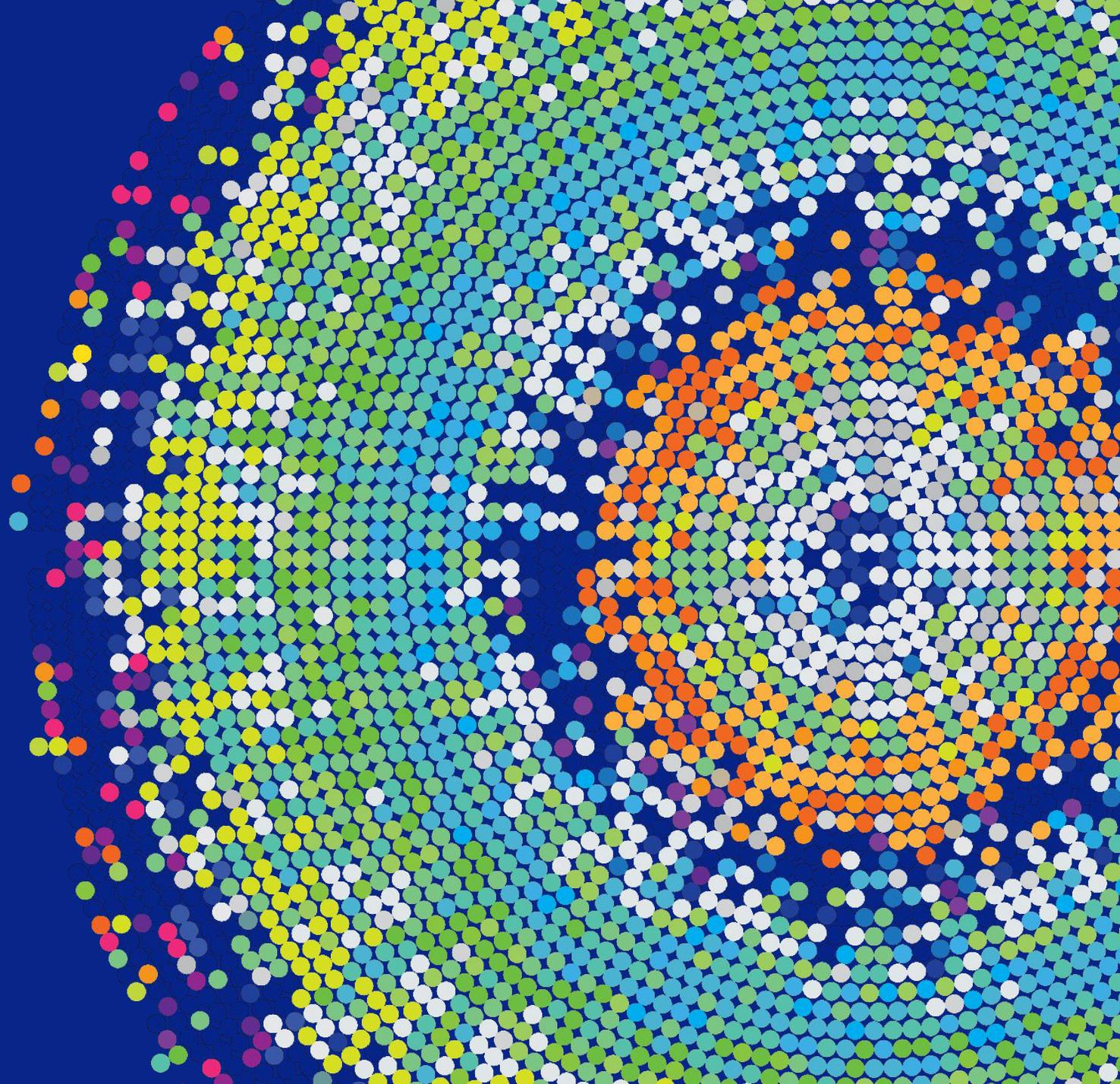# ODI

**The ODI in 2025:
Advancing trust in data**

Foreword by Sir Nigel Shadbolt and Louise Burke

Throughout 2025, artificial intelligence (AI) increasingly became part of most people's day-to-day lives, and every part of this rapid expansion is underpinned by data. As AI technologies spread, they exposed global weaknesses in data infrastructure. Governments and organisations have been forced to confront uneven data quality, patchy legal safeguards and growing concerns about the exploitation of personal and creative content. These tensions have brought new urgency to long-standing questions of trust, ethics and accountability in how data is collected, used and shared.

Our work at the ODI is shaped by the strategy we published in 2023. In the face of rapid AI advances, our mission, vision and six strategic principles remained constant. Our principles champion the importance of strong data infrastructure; the foundational role of open data; the need for trust in both data and those who share it; the continued presence of trusted, independent organisations - like the ODI - in the data ecosystem; one that is diverse, equitable and inclusive; and the creation of a new cohort of data leaders, through data skills and

knowledge. This Annual Review demonstrates these principles in action and show how they allow us to adapt and tailor our approach in a world of rapid change.

The coming years will see us centre on a number of areas that have become increasingly important to the stewardship of trustworthy data infrastructure, including data-centric AI, data assurance, data ethics, data standards (for data across the spectrum, from open to closed to shared) and the value of data. These themes extend the priorities we set out in our 2023 strategy and mark a shift towards a more action-oriented ODI. We remain anchored to the same vision, mission and principles, but the way we deliver has become more applied and closely tied to the practical realities facing governments, industry, civil society and individual members of the public.

As we look ahead, our focus will be on strengthening the infrastructure, standards and institutions that allow data to serve the public good - nationally and internationally - while ensuring that our work reflects the world as it is, not as it once was.

The ODI in 2025

Data underpins every part of the rapid expansion of AI. Every significant advance in automation, decision-making and digital services depends on the quality, openness and governance of data. The progress made also highlights long-standing weaknesses in data stewardship, raising urgent questions about the trust, ethics and accountability surrounding how data is collected, used and shared.

We are working at the cutting edge of research, delivering insights on data-sharing practices, the role of data work in AI supply chains, and the properties that data should adhere to in order to be AI-ready. A key focus of this work is the role played by governments around the world, not only as legislators, but as providers of data, benchmarks, or transparency mechanisms that can encourage innovation. In 2025, we embarked on a project to assess the AI-readiness of local councils in the UK to uncover where poor-quality data could hinder the implementation of AI in public services.

We examined the work that lies behind AI models, and developed a framework for AI-ready data. We contributed to discussions about the National Data Library (NDL), and responded to a call by the Wellcome Trust and the Economic and Social Research Council to set out a technical vision and architecture for the NDL that will make public sector datasets more accessible and benefit data-driven public services.

We played an active role in the data ecosystem and ensured that data was a key talking point with policymakers, business leaders and civil society. In February, we were a partner in the AI Fringe, which marked the Paris AI Action Summit with a two-day event at the British Library Knowledge Centre in London. Our Executive Chair and Co-founder, Sir Nigel Shadbolt, took part in a panel looking at the role of AI in public services, alongside speakers from the Ada Lovelace Institute and techUK, and we organised a panel on labour rights in the AI supply chain, with guest speakers from the Tony Blair Institute for Global Change and Full Fact. Throughout the year we've curated our own online events, with over 3,000 people attending to listen to experts talking on a range of topics, from the NDL to the growing use of "dark patterns" that can compromise user privacy and personal data security online.

We further consolidated our work on data standards throughout the year. As the steward of Solid, an open standard for managing digital identities and storing personal data for re-use across online platforms, we addressed barriers to adoption by improving user experiences, expanding hosting infrastructure and creating compelling real-world applications. We secured continued funding from Sport England to steward OpenActive, the open data standard for sport and physical activity opportunities data. We undertook significant technical modernisation of OpenActive with a new datastore and dashboard, showing over 450 data feeds with more than 10 million activities, and explored integration with health and community data initiatives. We are also working with various government departments to explore opportunities to implement data standards that make an impact in adult and child social care, as well as the volunteering sector.

Responding to the growing need for data and AI skills, we refined and expanded our data literacy offerings. We developed a transparent AI learning assistant, which successfully boosted attainment and AI literacy, proving that well-designed AI can enhance learning and promote safe, confident tool use. We streamlined internal operations by improving systems and integrations, and automating processes such as course booking, student enrollment, and certificate issuance and recertification. We trained more than 450 people through Government Campus and almost 50 people on the Data Ethics Professional programme.

The current climate of widespread online misinformation makes the role of trusted, independent organisations like the ODI more vital than ever. With many good initiatives being scaled back or closed - both in the UK and around the world - it is increasingly critical to maintain organisations that embody independence and reliable neutrality. It is crucial that we advocate for diverse, equitable and inclusive approaches to data collection and use. We continue to work with GIZ (Deutsche Gesellschaft für Internationale Zusammenarbeit) to support data-driven decision-making for policymakers worldwide, working directly with policymakers in Iraq, Egypt and India. Our Global Policy Observatory for Data-centric AI is very important in much of this work, as it reveals disparities in how countries emphasise data within their AI policies and looks to address the global data divide. We have used this in our work with the Foreign, Commonwealth and Development Office (FCDO) in Mexico and Chile, helping policymakers understand the extent to which data is a key topic in their AI policies.

There was much to engage with the UK government on across 2025. We contributed to key policy debates, including the AI Opportunities Action Plan, the National Data Library, smart data, and the Data (Use and Access) Act. We responded to the Department of Health and Social Care consultation on the NHS 10-Year Health Plan for England, and the Department for Business and Trade's consultation on Invest 2035: the UK's modern industrial strategy green paper. And we further strengthened our relationships with the government through our membership of the Smart Data Council and our involvement in the UK Open Government Network.

Despite the challenging economic and political landscape of 2025, the ODI successfully advanced its core mission: an open, trustworthy data ecosystem. We achieved considerable progress and have much to be proud of through our work across research, policy advocacy, training and consultancy. As technology continues to evolve and the world changes, the importance of data is only increasing, and we look forward to the opportunities the coming years will bring.

Contents

Principle one

We believe 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 our most pressing challenges.

What we achieved in 2025

In 2025, we contributed to developing a strong data infrastructure by developing open data standards, stewarding nationally significant data infrastructure, publishing original research, and contributing to UK and EU policy consultations.

We responded to the UK government's industrial strategy consultation and were encouraged to see data explicitly recognised as a strategic national asset when the Modern Industrial Strategy was published. This recognises data as a lever for growth, and was backed by funding for the National Data Library and the expansion of Smart Data schemes.

We became the steward of Solid in October 2024, and this year, we successfully transferred all appropriate infrastructure to the ODI's oversight. We are now working to scale Solid globally and to enable users to store their data and share it selectively. This decentralised approach allows data to be reused across platforms, rather than

being locked in proprietary systems. We also addressed current barriers to adoption by improving user experiences, expanding reliable hosting infrastructure, and creating compelling real-world applications that demonstrate immediate value to individuals, communities and organisations.

Data standards are an integral part of data infrastructure, providing assurance and allowing data to be shared across organisational and geographical boundaries. We have worked increasingly on data standards across the year, particularly with government departments. We are developing open data standards for volunteering with the Department for Culture, Media and Sport, and child and adult care standards in partnership with Social Finance for the Department for Education and the Department of Health and Social Care. This sits alongside OpenActive, an open data standard to help find physical activity opportunities, which we have stewarded for eight years, and which is now becoming more integrated with public healthcare systems.

Spotlight on: The National Data Library

The UK holds a vast amount of national data, presenting an opportunity to empower researchers and benefit data-driven public services. The National Data Library (NDL) has the potential to be transformative, especially in the new era of AI.

Over the course of 2025, we made recommendations for the design of the NDL, leveraging more than a decade of experience in designing and building open data infrastructure. We proposed an NDL blueprint in the Wellcome technical data challenge, focusing on AI readiness. Then, following the release of the AI Opportunities Action Plan, we hosted an event looking at the requirements of the NDL from different sectors. We also shared our vision of an AI-ready NDL at the NDL Symposium at King's College London and at London Data Week.

As the year came to an end, we started developing a proof of concept, informed by our AI readiness framework, to understand what a tangible, responsibly governed NDL asset could look like, ensuring that by design it respects the terms of use of each data source, captures data provenance and lineage, and organises the data in a way that is natively accessible to AI engineers.

We look forward to working with government, academia and private sector organisations from across the world towards an NDL with robust design, user-centric architecture and innovation-focused results.

EXPLORE

Our vision for an AI-ready NDL

Can the AI Opportunities Action Plan deliver?

TBI's vision for a National Data Library: an ambitious approach

Shaping the National Data Library: key considerations for the AI age

Spotlight on: AI-ready data infrastructure

Zooming out from a prospective National Data Library, most public sector data remains unstructured or inaccessible, creating "readiness gaps" that prevent AI innovation and adoption. Organisations lack clarity on whether their datasets are suitable for safe, effective automation or if they carry hidden risks, such as re-identification or severe bias.

We partnered with Nortal to pilot our AI Readiness (AIR) framework. We analysed 10 high-impact use cases across UK local government (including social care and housing), assessing data quality, metadata and governance to determine true AI compatibility and costs, and identify necessary infrastructure upgrades. This data included examples as varied as demographic data, unstructured social care notes and fire incidence reports, but critically, not personal data.

The pilot successfully identified critical data gaps in 10 council use cases. It revealed a mixed landscape where councils have made progress with structured data for predictive modelling, but often lacked the consistent metadata, standards and infrastructure required to support advanced generative AI applications. These findings also inform our work on the Croissant vocabulary, whose new release is expected at the start of 2026. Furthermore, we can provide actionable roadmaps for data maturity, enabling local governments to move from theoretical AI exploration to practical, risk-aware service implementation.

EXPLORE

The state of AI-ready data in local government

Case study: Data standards for public services

Public services, community organisations and the health system face increasing strain from rising demand, limited funding, and staff and volunteer shortages. Fragmented data about opportunities and services compounds these challenges, preventing systems from working together effectively.

We have developed and maintained OpenActive, an open data standard for discovering physical activity opportunities, for many years. With renewed support from Sport England, we strengthened the data infrastructure, expanding the standard so small providers can easily publish open data. OpenActive now provides near-real-time access to over 10 million activities from almost 5,000 providers, covering approximately 65% of local authorities in England. We published Data infrastructure for a healthy nation, showing how better access to physical activity via shared data supports preventive care,

improves patient outcomes and reduces NHS demand. We also began integrating OpenActive with healthcare systems, with further tests planned in 2026 to refine its impact.

Our experience delivering OpenActive directly informed our work in the volunteering sector. Working in partnership with the Department for Culture, Media and Sport, we convened senior leaders to explore how open data infrastructure could improve the volunteering experience. This led to the co-design and testing of open standards for volunteering data, alongside Do IT and Team Kinetic, addressing long-standing gaps in how opportunities are described and shared. Standardised, openly published volunteering data makes it easier for people to find opportunities and enables new approaches. Our hackathon proved that standardised data could be used in agentic search tools and crisis response scenarios.

Taken together, these initiatives highlight the potential to align and connect open standards for physical activity and volunteering, creating a more joined-up data ecosystem that supports healthier, more resilient communities.
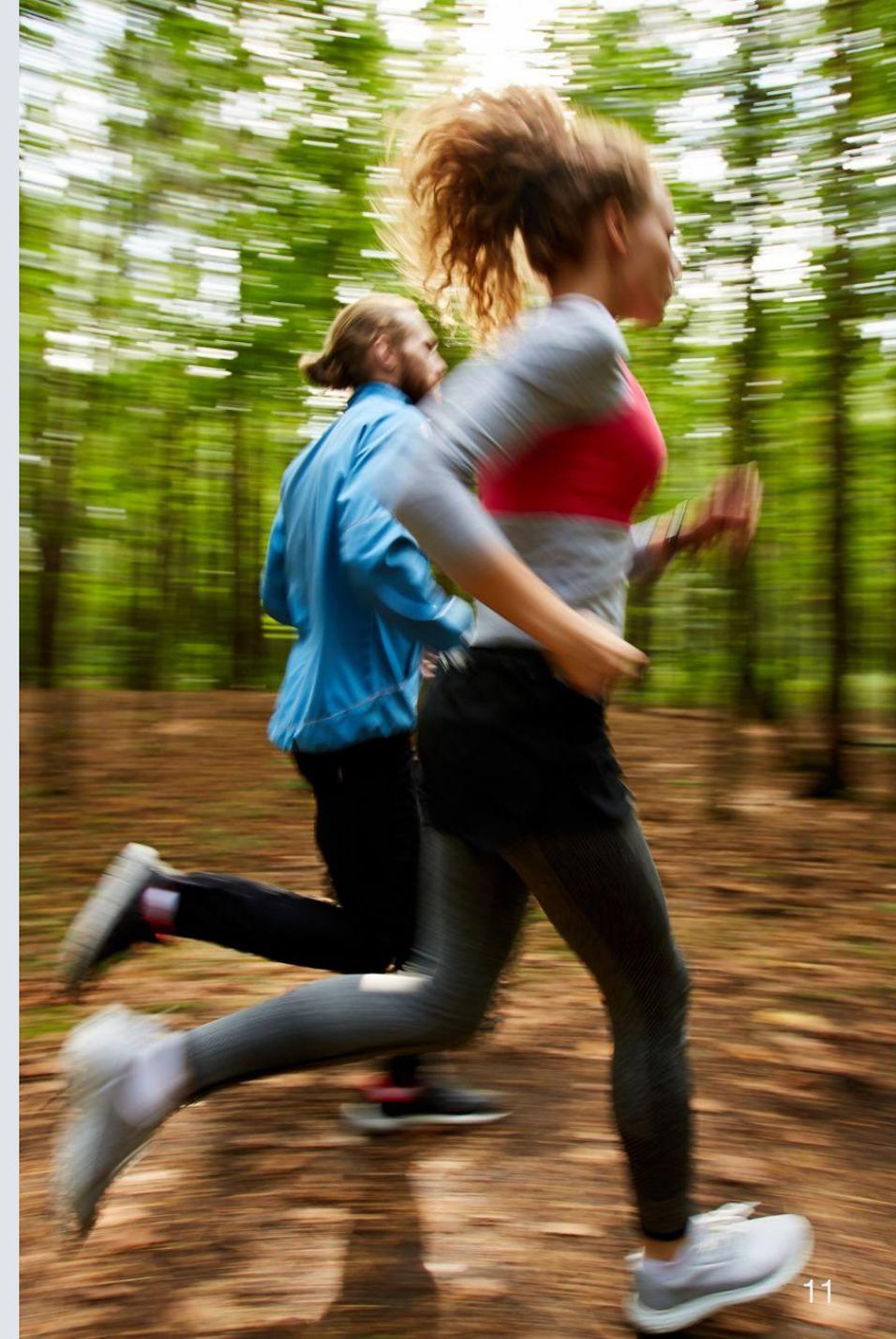
Case study: Solid

The way most people interact on the web involves handing over data to organisations that retain it. It is such an accepted way of interacting on the Web that it is rarely questioned. But in recent years, this relationship has received a lot more scrutiny as people become more cautious about handing their data over to organisations.

Solid is a decentralised infrastructure for the Web, and aims to fundamentally change how data is managed, decoupling data from the applications that use it. Instead of continuously providing data for every new service that people use, they can keep this data in their own personal online data store. Each time users interact with a new app or service, they only allow it to access the data needed to perform the service required. This gives them greater control and reduces friction when using new services or switching between them.

Throughout 2025, we recruited a full technical team to work on Solid. We created documentation to help new developers build their first Solid apps, and created a guide for people to run their own Solid server. We built developer tooling to make it easier to build with Solid, and we built new applications, such as one that enables users to port conversations between different chatbot providers. We continue to support the wider Solid community, including regularly hosting the Solid World webinar, and we look forward to co-hosting the Solid Symposium in 2026 with the University of Strathclyde.

EXPLORE

solidproject.org

Case study: Smart Data Challenge Prize

Smart data, the secure sharing of customer data with authorised third parties, has great potential value for the economy, and could save consumers both time and money. However, without access to the data and prototyping real services, proving that value is challenging.

Alongside the Department for Business and Trade, Challenge Works, bigspark and NayaOne, we ran the Smart Data Challenge Prize, using synthetic data (artificially created information that mimics real-world data), in a secure, isolated test environment, enabling 10 teams to develop prototype services that could deliver value to consumers in the "real world", if the real data were available. These services used data across multiple industries, including finance, property, energy, retail and transport, to help individual and business customers cut bills, reduce their carbon footprint, or shorten the time it takes to buy a property.

With a range of prototype services, a recognised winner (a Digital Property Pack developed by Moverly and the Open Property Data Association) and two runners-up (Open Transport Initiative, and Nigel by Beyond Encryption), there is a clear argument for the potential of smart data to transform people's lives for the better. Smart data will enable people to save money on regular bills, live more eco-friendly lives, get access to new forms of financing, and greatly reduce the stress of buying and selling homes.

EXPLORE

About the Smart Data Challenge Prize

Launching the Smart Data Challenge open call

Principle two

Strong data infrastructure includes data across the spectrum, from open to shared to closed. But 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.

ODI

What we achieved in 2025

We published the Open Data Use Case Observatory for Microsoft, a collection of 15 case studies from across subject matters and applications that presents a unique usage of open data with real-world impact. This work demonstrated the role that open and publicly available data has played in recent important innovations.

We play an active role in Stream, unlocking the potential of open data in the water industry that was historically siloed across different companies in different formats and standards. Stream is now live and publishing critical UK water-sector open data from Brighton to Inverness. This open approach now benefits customers, society and the environment, and we will continue to help individual water companies on their open data journey.

We worked with Global Fishing Watch to strengthen the foundations of their data practice, helping them to use open data to improve understanding of global fishing,

supporting efforts to address illegal, unreported and unregulated activity. We supported this work by strengthening the foundations of data practice, helping ensure that openness at scale can be sustained responsibly and relied upon across jurisdictions.

In highly regulated and fragmented environments, data often lacks the necessary coherence between initiatives. Our collaboration with Arup, a long-standing partner of the ODI, on whole-life carbon in the built environment addresses this challenge. Multiple organisations and projects generate data across the asset lifecycle, but without consistent standards or shared reference points. By analysing data flows, governance models and value exchange within these ecosystems, we have supported efforts to improve interoperability and reuse. This work shows how open data can act as connective tissue, enabling comparison and scale in complex systems.

Case study: Stream

Historically low trust in water companies, combined with high environmental pressures on the industry as a whole, has led to calls for greater transparency and innovation from the sector. Data is the key to achieving this, but it has been siloed across different companies in different formats and standards.

Stream is a novel data institution that publishes data from water companies in an open and standardised way. Stream is making data available for the very first time, from domestic water quality and regional water consumption, to the National Storm Overflow Hub that provides near-real-time updates on every storm overflow in England, in both an interactive map and an open API.

With downloads in the tens of thousands, and hackathons as part of Stream's Open Data Day and Northumbrian Water's Northumbrian Innovation Festival, we know our data is being used across industry, government, academia and civil society, for a wide range of purposes. From local planning to examining river and drinking water health, Stream is the data pipeline supporting water use in the UK.

EXPLORE

Stream Water Data

Enabling innovation across the water sector with Stream

An open data strategy for the water industry

The National Storm Overflow Hub | Stream - Portal

Case study: Arup

Addressing climate impact and managing flood risk requires coordination across many stakeholders. Data sharing is complex, and new initiatives often lack reliable practices. Open data can support effective climate action, but governance gaps limit its full potential for collaboration and resilience.

Our co-delivery work with Arup has strengthened data sharing and maturity, and has involved developing a data institution model to improve flood risk management in the North East of England, enhancing data sharing in Greater Manchester, and improving infrastructure at EnTrade and Coastal Partnership East to ensure reliability and compliance. Our partnership work has promoted open data through tailored training, research on its role in climate resilience, and facilitated discussions on a whole-of-society approach to data collection.

Our joint delivery has strengthened trustworthy data practices and stewardship models across the water and climate sectors, improving decision-making, collaboration and environmental risk management. The partnership has provided training and raised awareness of the benefits of open data for Arup's water teams. It has also produced research on best practices and enabled discussions on the enablers and barriers to open data infrastructure for a whole-of-society approach to Blue-Green Infrastructure planning - where green spaces and water systems are strategically integrated, rather than considered in isolation.

EXPLORE

Arup partnership

ARUP

Spotlight on: The open data observatory

Data is a critical part of tackling the biggest challenges we face as society, and has been the foundation of many of the important technological innovations of recent years. However, recent trends around the use of data by some AI businesses have led to increased concerns over the risks of sharing or opening data.

Microsoft and the ODI have partnered for a number of years to advance the cause of trustworthy data sharing and collaboration, so that any organisation of any size can more easily collaborate around data and realise the benefits of the data economy. As some were predicting an open data winter, we worked together to build the open data use case observatory, to communicate with renewed urgency the importance of open data in tackling society's biggest challenges, innovation, and better business.

The observatory collated a wide range of examples of best practice to demonstrate the role that open data has played in these use cases. We described in-depth how the Protein Data Bank, an open data repository for bioinformatics, has matured over 50 years of operation to the point that it could fuel the development of Nobel prize-winning AI from Google, and also explored how open data enabled AI startups to power electrification in rural Uganda. We hope to continue to add to the observatory in the future, with additional examples of important datasets as they emerge.

EXPLORE

Open Data Use Case Observatory
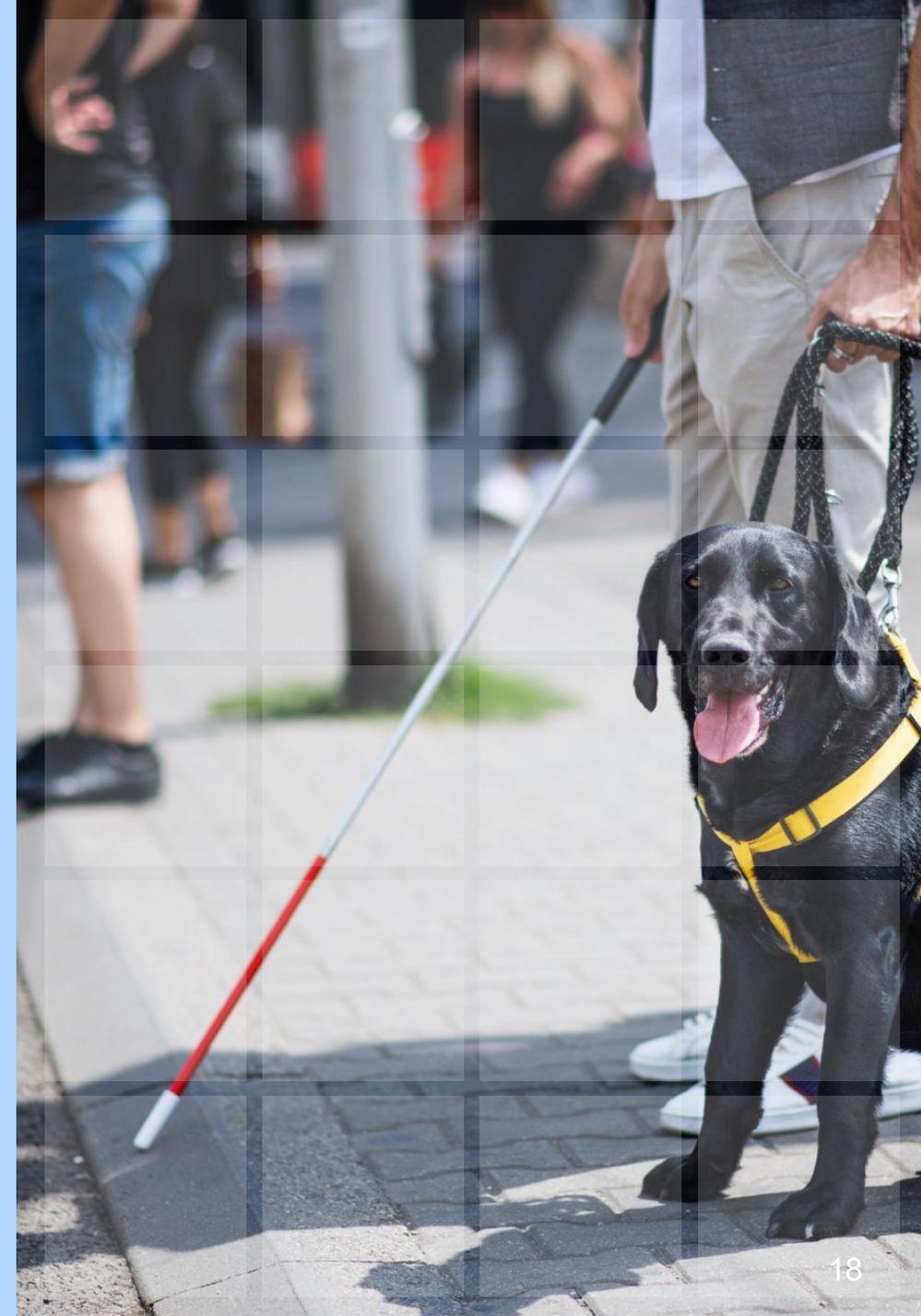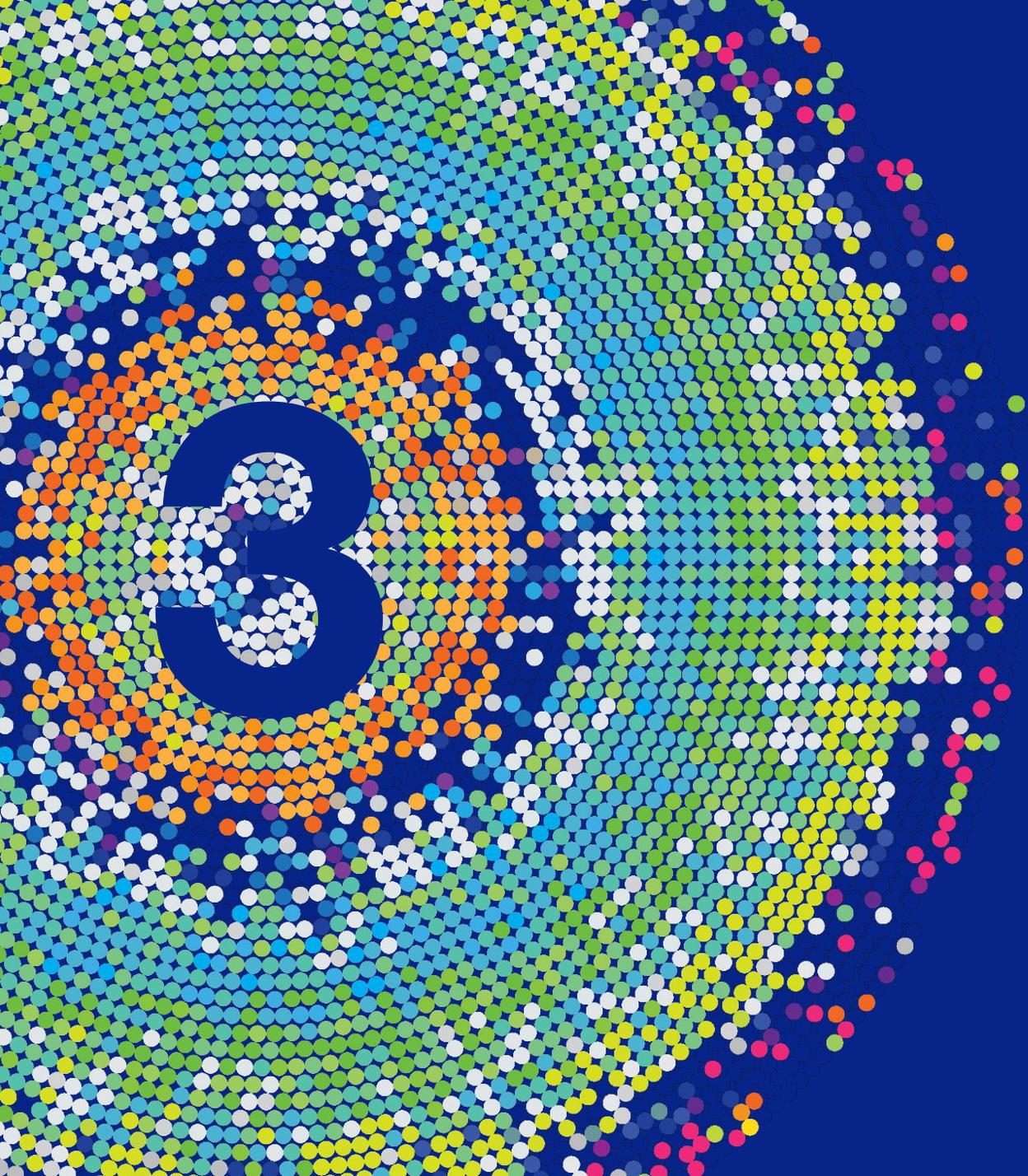
Principle three

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 share it.

ODI

What we achieved in 2025

Throughout the year, we engaged with the government on policy issues related to trust. We responded to the Department of Health and Social Care consultation on the Change NHS 10-Year Health Plan for England, and recognised the need to commit to investing in expansive technical foundations and careful management of data when their final plan was published. We responded to the Department for Business and Trade's consultation on Invest 2035: the UK's modern industrial strategy green paper, and were encouraged to see that data was explicitly treated as a strategic national asset, and also that the National Data Library is to be AI-ready from the outset; both important priorities we championed.

We developed a range of practical tools and, across 2025, these were increasingly adopted in a number of sectors in the UK and globally, including water, energy and the environment. Organisations such as Northumbrian Water, Northern Powergrid and Global Fishing Watch

are using the tools such as the Data Practices Assessment, the Data Sharing Risk Assessment Workbook, the CARE (Context, Accountability, Rights, Ethics) tool, and an Open Data Maturity Assessment tool, helping them to evaluate and improve their data governance practices. Our work on data-centric AI continued, examining the work that lies behind AI models, and developing a framework for AI-ready data, which has been used as a baseline to assess the AI-readiness of local councils in the UK.

We continued our work with ML Commons to develop Croissant, a standard for describing machine-learning datasets consistently. We joined a new consortium called CoCoDa, which brings together technical and legal expertise to investigate systemic risks in society arising from dominant online platforms, and to address the root causes.

Spotlight on: Data assurance tools

Existing data governance frameworks and tools often aren't effective for people responsible for creating and implementing data governance in organisations. They lack easy, practical tools focused on specific use cases, making it challenging to implement data strategies across the spectrum, from open data to data sharing.

We developed a number of products which address this need, combining research-led methodologies, a user-friendly interactive interface and innovative use of large language models to enhance the tools' capabilities. The source code for our digital products is openly licensed, and access to the digital interface is currently available for registered ODI members. There is no charge for registration.

ODI's tools are driving measurable impact. Global Fishing Watch adopted the CARE tool to assess and mitigate ethical risks in data and AI use, strengthening marine governance. UK water sector organisations were part of a group of 300 users who completed the Open Data Maturity Model using the Maturity Assessment Tool, gaining comparable insights that inform investment, publication, and long-term data capability building.

EXPLORE

List of free tools available to ODI members

Putting data ethics principles into practice

Global Fishing Watch

Case study: Lease

Leasehold ownership in England and Wales is complex and has important legal and financial considerations. When a lease tenure falls to below 80 years, extension costs increase substantially, and securing a mortgage becomes difficult, restricting sales. Leaseholders need accurate, timely information to make informed decisions.

We conducted a robust analysis of HM Land Registry datasets, including leasehold information about over 7.6 million properties, to provide the LEASE team, regulators and citizens with evidence-based information on existing leases and their lengths. This better informs their advisory services. We also developed a free, public-facing tool accessible on the LEASE website that allows consumers to check the length of their lease by entering their address.

Since publishing the tool in September, it has been used by over 3,500 people who have looked for lease information on over 6,000 properties. One member of Parliament reported that they were able to advise a constituent that they had 10 years left on their lease, due to using the tool.

EXPLORE

Check your lease length
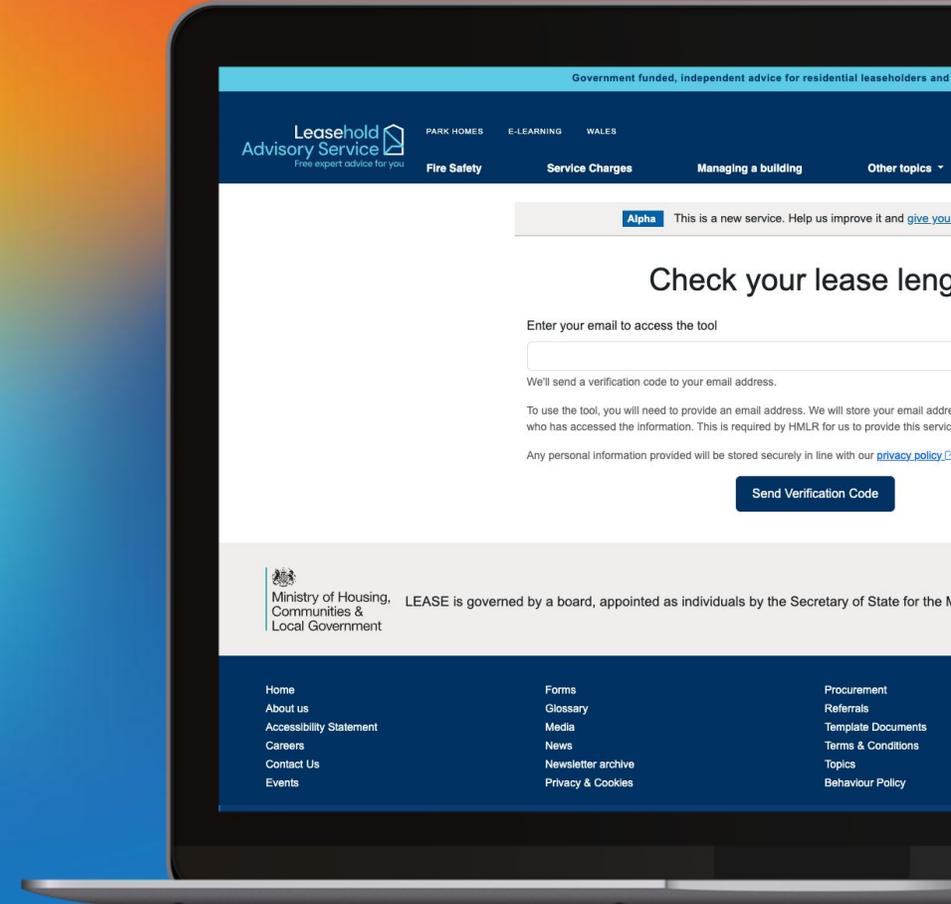
Spotlight on: Data-centric AI tools

The AI landscape suffers from a lack of visibility on multiple fronts. Externally, "black box" models hide critical risks regarding bias and copyright. Internally, organisations lack clarity on whether their own data is safe for automation. Furthermore, policymakers struggle to navigate a fragmented global regulatory landscape, making it difficult to create effective standards for data governance.

We developed a suite of tools to address these visibility gaps. The AI Data Transparency Index benchmarked 22 major foundation models to measure exactly how open (or secretive) they are about their training data. The Global Policy Observatory for Data-centric AI contains a dashboard of 512 policy documents from 64 countries to track international divergence in AI governance. With the Fr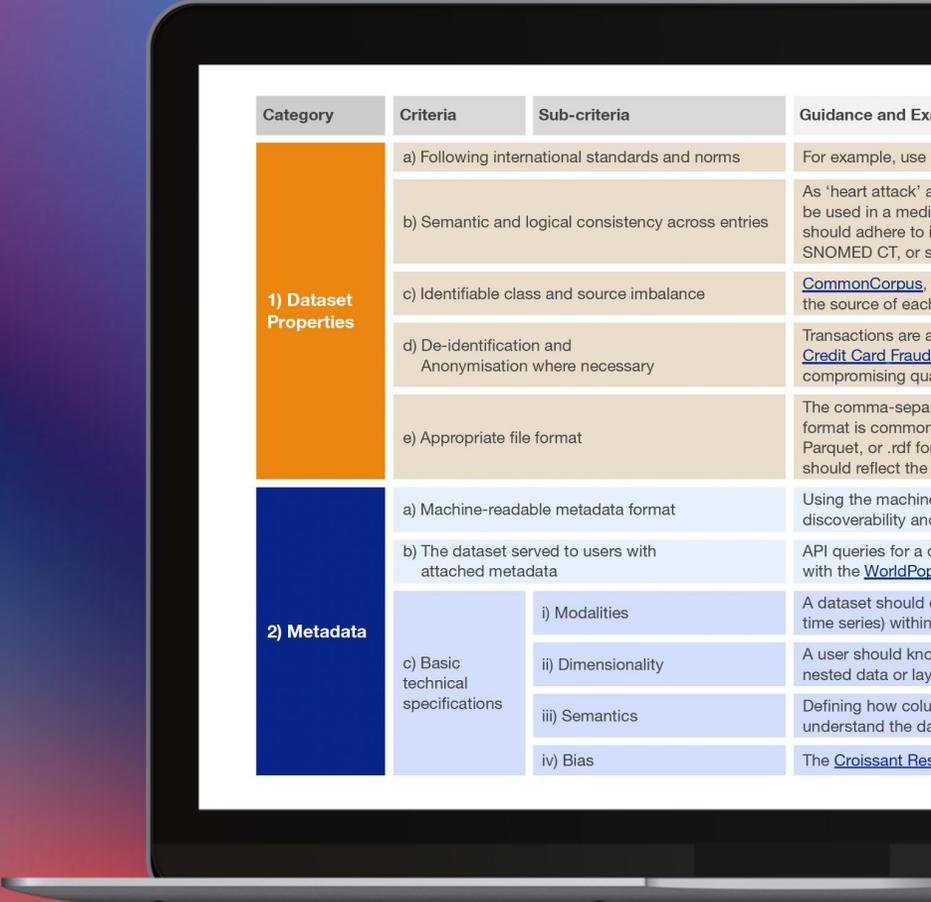amework for AI-ready Data, we partnered with Nortal to pilot this readiness tool, stress-testing it on 10 high-impact use cases in UK local government to diagnose actual infrastructure maturity.

These resources have evolved from research outputs into active consulting and engagement tools across 2025. The Global Policy Observatory has been deployed in international projects with the Foreign and Commonwealth Development Office (FCDO), serving as a critical benchmark to assess how nations prioritise data within their AI strategies. The Framework for AI-ready Data is driving industry standards, having been presented at major forums including NeurIPS 2025 and the British Computing Society to demonstrate practical pathways for data readiness. The AI Data Transparency Index has been presented at webinars involving hundreds of attendees, sparking spirited discussion about what data transparency should look like and for whom.

EXPLORE

The AI Data Transparency Index

Global Policy Observatory for Data-centric AI

Case study: Data licences

The AI ecosystem lacks standardised data licensing. Existing frameworks, such as Creative Commons licences, were not designed for AI contexts, leaving critical terms like 'non-commercial use' or attribution requirements open to conflicting interpretations when applied to model training. Organisations have to navigate a fragmented landscape of bespoke, incompatible agreements, which requires significant legal resources to review and ensure compliance across multiple datasets. This uncertainty discourages data sharing, hinders innovation, and complicates compliance with ethical standards and copyright laws for researchers and smaller entities.

In collaboration with Duke University and MLCommons, we conducted interviews and ran a survey to research motivations for data sharing and the barriers posed by current licences. We proposed a suite of modular, standard data licences designed to clarify rights, streamline sharing, and embed responsible AI principles directly into legal terms.

The work validated the industry need for standardised terms. These tools reduce legal friction, promote equitable access for smaller entities, and ensure data sharing aligns with ethical governance and transparency requirements.

EXPLORE

Unlocking Data Collaboration with AI-Ready Licenses

How might standard contract terms help unlock responsible AI data sharing?

ODI

Spotlight on: Enabling responsible data practices for AI

Machine learning datasets are often challenging to find and use because their descriptions can be unstructured. This can slow innovation and obscure critical information, such as bias and provenance, making ethical and responsible AI development difficult.

We co-chair the Croissant working group, and have developed it to act like a digital "nutrition label" for datasets. We ensured this standard includes critical safety information (such as known biases, ownership, or usage limits). This allows AI developers to instantly understand what is inside a dataset and automatically filter out data that isn't safe or legal to use.

Croissant has become a common language for data, now used by major platforms like Hugging Face to describe over 700,000 datasets. This level of adoption reduces the time spent cleaning data and, crucially, stops developers from accidentally training AI models on flawed or unethical material.

Building on its widespread adoption, Croissant focused on expanding its ecosystem in 2025. For example, adding support for CKAN and the Model Context Protocol (MCP). Now moving toward its 1.1 release, the standard continues to reduce data cleaning time while safeguarding developers from accidentally training AI models on flawed or unethical material.

EXPLORE

Transforming AI data governance with Croissant: a new standard for ML metadata

Case study: CoCoDa

Very large online platforms, such as Google, Amazon and social media sites, play a key role in our day-to-day lives, and hold vast amounts of data that enable them to do so. However, these companies are protective of information about how their platforms work and the data that they hold about people, making it very difficult for researchers and others to address the systemic risks they create, or to generate public benefit from this accumulation of data.

We are working with the University of St. Gallen, the University of Lausanne, and Maastricht University to bring equity to the high **co**ncentration of **co**ntrol and **da**ta (CoCoDa) held by dominant online platforms. This project extends a programme of work we started at the ODI in 2023 exploring how well designed, governed and resourced global data infrastructure can enable research and

innovation to address pressing global challenges. By combining technical data access methods with novel legal approaches, the CoCoDa project seeks to create "techno-legal" tools that can support regulators, researchers and civil society in studying and mitigating systemic risks that arise from important data being kept private.

After launching in July 2025, the project will last four years. Our work includes examining the importance of structured data approaches for helping researchers to access data from online platforms. This means ensuring that when tech companies need to give data to researchers, as regulated under article 40 of the Digital Services Act, they provide this information in a way that is useful to the researchers.

EXPLORE

Technolegal solutions for online platform accountability: announcing the CoCoDa project

ODI and CoCoDa: Technolegal Solutions for Online Platform Accountability

Spotlight on: Making AI use more responsible in practice

In 2024, Constant Washing Machine, by Blast Theory, emerged as part of the Data as Culture curated artists' residency on the University of Sheffield's Framing Responsible AI Implementation & Management (FRAIM) Project. FRAIM instigated world-leading research on responsible AI (RAI) in the policies and practices of key organisations in different sectors of the economy to ask what we actually mean when we talk about 'Responsible AI?'. In partnership with the ODI, Sheffield City Council, the British Library and Eviden, the research revealed the fragmentation of RAI into many different principles and ideas across different sectors. It brought out the cacophony of different understandings and approaches to making AI use more responsible in practice.

Using the research as its source, Constant Washing Machine materialised and creatively problematised the findings through the creation of eight bespoke soap bars embossed with words and phrases contributed by RAI researchers, portraits of the researchers with their soaps, and performance. Through discursive workshops (including at conferences and symposia such as UKAIRS25, AI4LAM Fantastic Futures 2025 and REMIX 2026), modest and impermanent objects – bars of soap – are transformed into a reminder that it is we, our daily habits and cultures, rather than policies or regulations, that make responsible AI happen.

Constant Washing Machine was honoured with the UK AI Research Symposium (UKAIRS) inaugural Impact Award, recognising outstanding research that drives impact and real-world change. Constant Washing Machine is available for exhibition, public engagement, research workshops and events.

FRAIM is a Bridging Responsible AI Divides (BRAID) scoping project funded by the Arts and Humanities Research Council as part of the broader UK Bridging Responsible AI Divides (BRAID) programme to bring out the essential role of arts and humanities perspectives in shaping Responsible AI dialogues.

Launch of Constant Washing Machine by Blast Theory. Photo by Melanie Pollard.
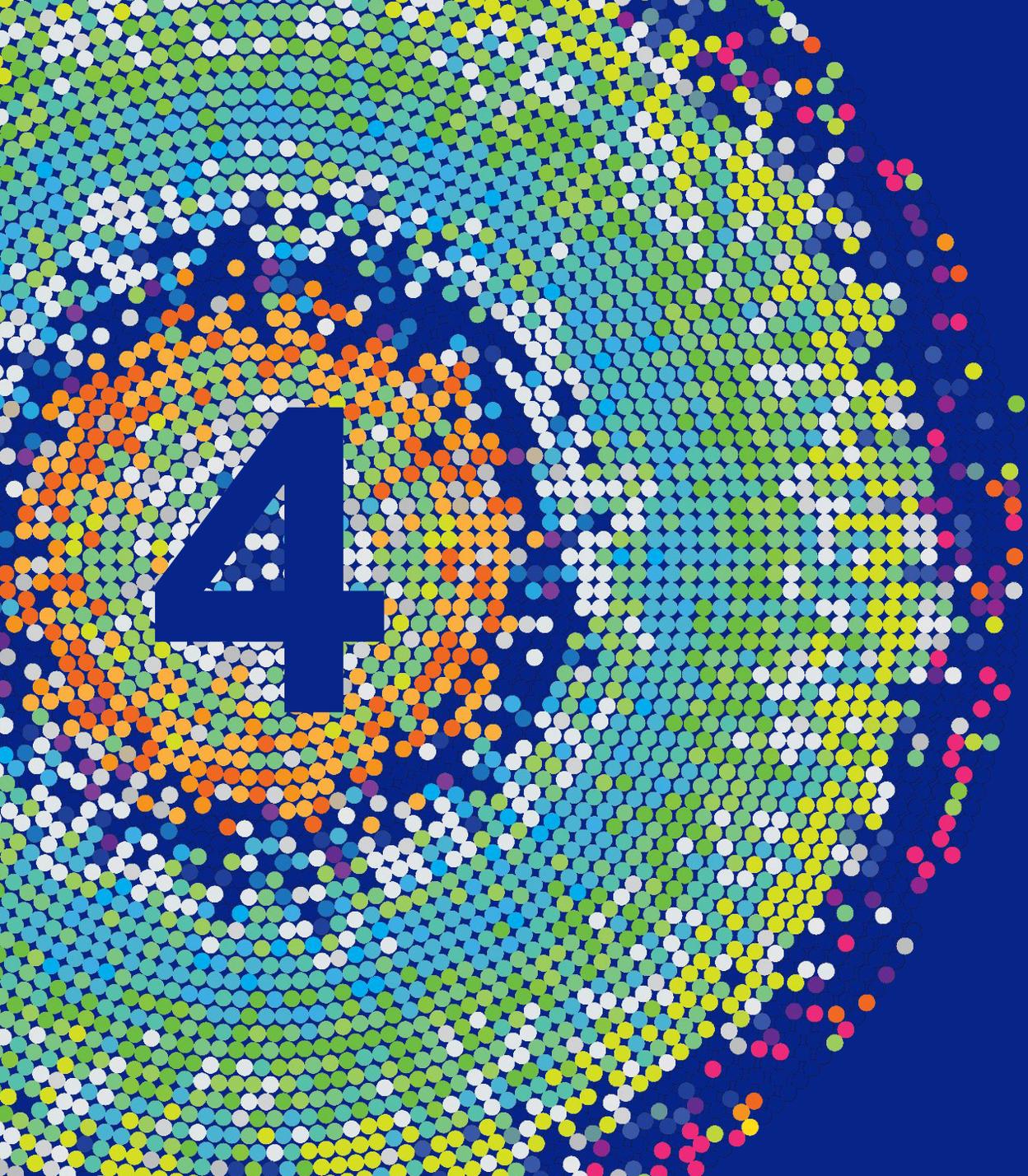
Principle four

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

ODI

What we achieved in 2025

As a trusted, expert, independent organisation, we have a recognised ability to convene diverse stakeholders. This year, we hosted roundtables for organisations exploring the opportunities of data across the spectrum. This included a national regulator, the Department for Culture, Media and Sport together with volunteering-involving organisations, the pharmaceutical sector, and organisations interested in exploring how data infrastructure can get more people physically active. As part of our work on Solid, we convene the global Solid community, which includes people with a range of technical skills, such as developers and researchers, as well as those with non-technical backgrounds, including policy and social science. We also host Solid World, a quarterly webinar providing updates on the project and a platform for community members to share their work on applications.

Across 2025, we've hosted 29 webinars, with 7,238 bookings and 3,151 attendees. We continued our Data-centric AI series and added a series where Data Ethics Professionals share their experiences and offer practical insights with actionable takeaways, which has grown a committed audience.

We continued to refine and align our membership offer with our audiences, delivering more of what they need, and ensuring it remains both relevant and sustainable. We have attracted new members through our Data Ethics Professional webinar series, and have drawn interest from national bodies and companies who want to offer our combined membership and learning opportunities as a benefit for their employees.

Spotlight on: Advocacy and convening

Throughout 2025, we continued to see the increased use of AI and data-enabled technologies as they became embedded across sectors, economies and public services. This spread exposes weaknesses in data stewardship and transparency, raising the urgency of long-standing questions about trust, ethics and accountability in how data is collected, used and shared.

The ODI is a trusted institute with expertise and insights in open data and data across the spectrum. We are well placed to facilitate discussions with policymakers, civil society and businesses to explore best practices in a technologically evolving world.

This year, we convened several roundtable discussions across different sectors. We hosted an industry roundtable for a national regulator to explore the possibilities of digitally and data-enabled services in their industry. We also convened a roundtable in Manchester, bringing together people from regional government, local NHS digital teams, and community sport organisations to explore how OpenActive can address specific regional challenges. We had 40 people attend the event, and established relationships that could help OpenActive become integrated with the healthcare service, ultimately helping more people get physically active.

In partnership with the Department for Culture, Media and Sport, we convened senior stakeholders from volunteer-involving organisations and industry bodies, data and technology leaders, academics, and volunteers to explore how a modern, open, data infrastructure could transform volunteering. This led directly to our programme of work to develop and test open standards for volunteering data, which will continue into 2026 with a series of pilot tests.

EXPLORE

Let's transform volunteering with data

Data infrastructure for a healthy nation: reflections from our roundtable

Case study: Webinar series

Large-scale in-person events are costly to put on, and attracting audiences can be challenging. This can make it challenging to engage audiences and share our work.

Webinars have become a very effective way of engaging our audiences, and we have found them to be a popular and cost-effective way to share our work. We provide our audiences with expertise from our team and guest panelists in a format that is easy to attend without much time commitment, or which they can catch up with on demand. It also provides our experts with a platform to take a deep dive into topics, particularly data-centric AI, and Data Ethics Professionals.

The two webinar series, Data-centric AI (DCAI) hosted by Prof Elena Simperl, and Data Ethics Professionals (DEP) hosted by Dr Kay Achenbach, shared a consistent,

hour-long format across 2025. Each session featured a single guest speaker, with DCAI hosting research experts, while DEP showcased alumni who were data ethics experts or practitioners.

To date, the DCAI series has had 11 webinars, six taking place this year. In total, it has had 3,199 bookings and 1,237 attendances, with 1,619 bookings and 677 attendance this year. The series has an international reach with 62% of bookings from the UK, 7% from the US, 7% from Canada and 24% from the rest of the world. The DEP series started in 2025, and to date, there have been nine webinars with 2,251 bookings and 785 attendances. Like the DCAI series, it has attracted sign ups from the UK and around the world.

EXPLORE

Learn from world experts on the data in AI

Data Ethics Professional

ODI open data institute | Nortal

Data-centric AI webinar #10
Local authorities, data, and AI
Online Thursday, 12 June
(16:00 - 17:00 GMT+1)

Case study: Public policy

With shifting global politics, a new UK government settling in, and major European elections, policy developments have accelerated. The UK introduced the AI Opportunities Action Plan, passed the Data (Use and Access) Act and published the Modern Industrial Strategy and 10-Year Health Plan - all of which required timely engagement.

We launched a European data and AI manifesto to clarify our vision for a trustworthy, well-governed data ecosystem. We contributed to key UK policy debates, including on the AI Opportunities Action Plan, National Data Library and Smart Data. We took part in panels, roundtables, political briefings and thought leadership, submitted evidence in consultation responses, and joined parliamentary working groups. We also participated in the three main political party conferences, taking part in debates and chairing

sessions. We strengthened government relationships through our membership of the Smart Data Council and involvement in the UK Open Government Network, and also with other civil society organisations in the ecosystem, ensuring a united approach on critical issues.

Our work increased the ODI's visibility and influence, with stronger engagement with senior policymakers and advisors, and growing trust from partners. We have also had direct invitations to brief government leaders, regular requests to speak at high-profile events and the opportunity to give evidence to a Parliamentary Select Committee. Together, these demonstrate meaningful impact on national data and AI policy discussions.

EXPLORE

An ODI European Data and AI Policy Manifesto

The National Security Strategy - Oral evidence

Conference season on data & AI

The importance of data quality and interoperability in addressing NEETs and social mobility

## Spotlight on: Strategic councils and advisory forums

UK government, industry and arms-length bodies require clearer approaches to data governance and capability. Fragmented initiatives and inconsistent standards drive demand for trusted, expert guidance. They seek help to align priorities, improve coordination, and build resilient, well-stewarded data infrastructure, especially with AI's impact and new smart data legislation.

We provided strategic counsel through our participation in major advisory forums. As a member of the Data for London Board, we help guide the development of a joined-up city data platform and strategy. Through the Smart Data Council, we advise the Department for Business and Trade on coherent, cross-sector smart data schemes.

Through the Department for Science, Innovation and Technology's Digital Skills Council, we support efforts to strengthen national digital skills and workforce readiness. As a member of the OpenUK AI Advisory Board, we provide expertise for this key area of open source.

The ODI's advisory role has strengthened the coherence and ambition of UK data policy, helping shape frameworks that enable responsible innovation, improved public services, and better use of shared data. Our guidance has supported more coordinated smart data schemes, a more effective London-wide data ecosystem through the Data for London Library, and a clearer approach to digital skills, contributing to a more confident and capable digital economy.
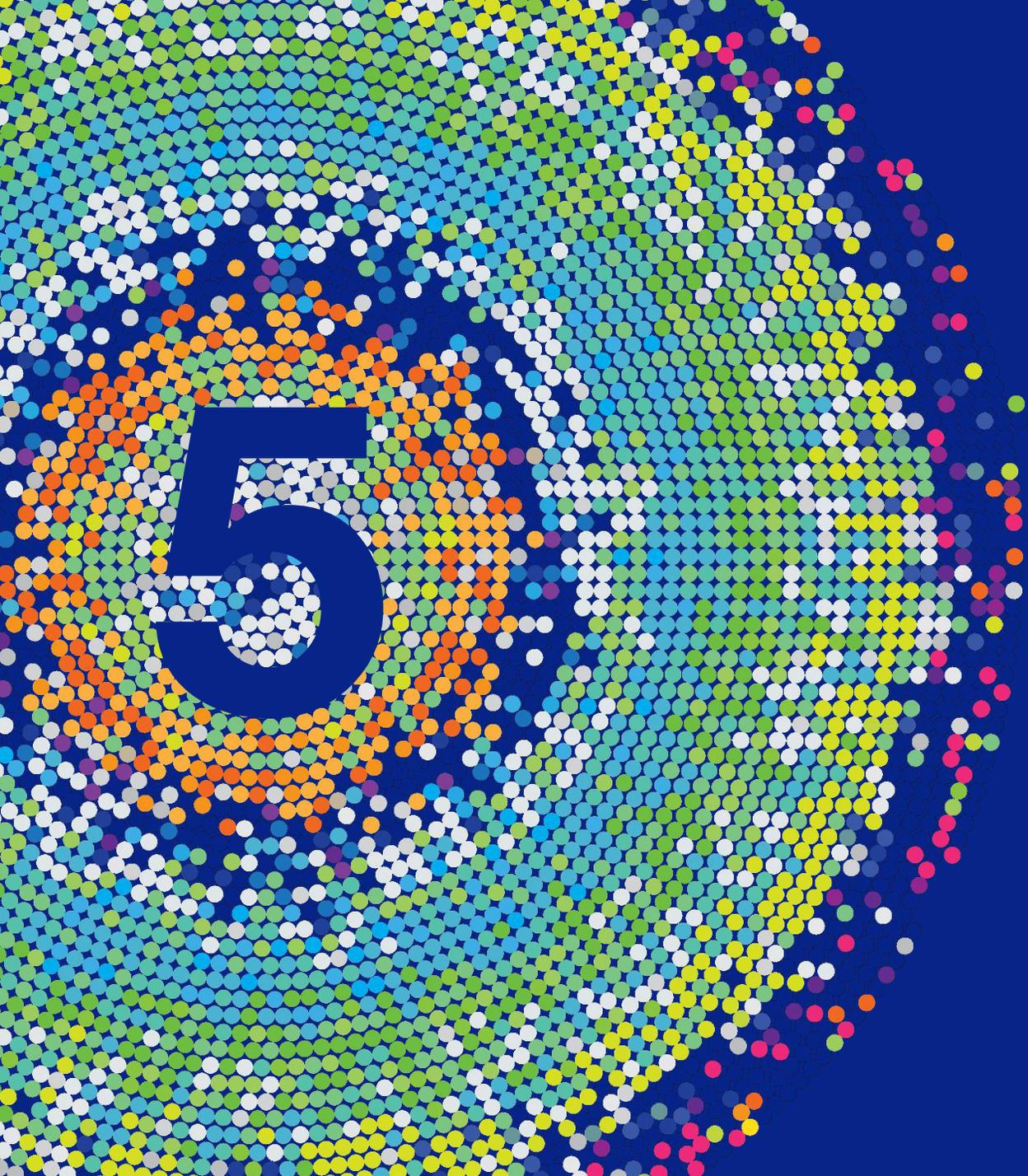
Principle five

For data to work for everyone, those collecting and using it need to be highly alert to inequalities, biases and power asymmetries. All 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.

ODI

What we achieved in 2025

We have endeavoured to work globally throughout 2025 to encourage an equitable approach to data and AI adoption so that everybody can see the technological benefits and have the same access to information and innovation.

Early in the year, we hosted a delegation of data and AI policy and regulatory representatives from Argentina, Brazil, Chile, Colombia and Mexico, supported by the UK Foreign, Commonwealth and Development Office (FCDO). Together, we worked to forge a collective understanding of data and AI governance and how to join up UK approaches with those of Latin American governance communities.

We have continued our relationship with the FCDO in a number of geographies, including helping to usher in a new generation of AI regulation in Chile, and working together with the government of Mexico City to develop an AI-powered chatbot to improve how residents and visitors access public and cultural information, particularly in the lead-up to the FIFA World Cup in 2026. Much of this work is supported by our framework for AI-ready data and our Global Policy Observatory for Data-centric AI, which reveals disparities in how countries emphasise data within their AI policies and looks to address the global data divide.

We have partnered with GIZ (Deutsche Gesellschaft für Internationale Zusammenarbeit) to support data-driven decision-making for policymakers worldwide, working directly in Iraq, Egypt and India to help them use data and AI to deliver on their policy goals.

Spotlight on: Working globally

The problems that face society from public health issues to environmental impacts are not confined to national boundaries. International collaboration is needed when challenges are global, sharing insights and best practices, or working together to solve issues facing us all.

Data and digital solutions aren't always constrained by geographical boundaries and the Foreign, Commonwealth and Development Office (FCDO) has been connecting governments, businesses, and civil society around the world to collaborate on tackling common challenges. Whether through supporting UK-Southeast Asia Tech Week, or working closely with city and national governments on how to maintain data best practices in the age of AI, the ODI has supported the FCDO's global vision.

In Mexico City, we are embedding data and AI best practices to improve digital public services. In Chile, we are helping usher in a new generation of AI regulation. In South Africa, we are embedding data ethics in city data ecosystems, while across Southeast Asia, we are sharing research on cutting-edge topics from data-centric AI to data institutions and Smart Data. We are a key partner to the FCDO, delivering projects and sharing insights globally. Combined, this promotes a collective understanding of data and AI governance, and helps to coordinate approaches around the globe.
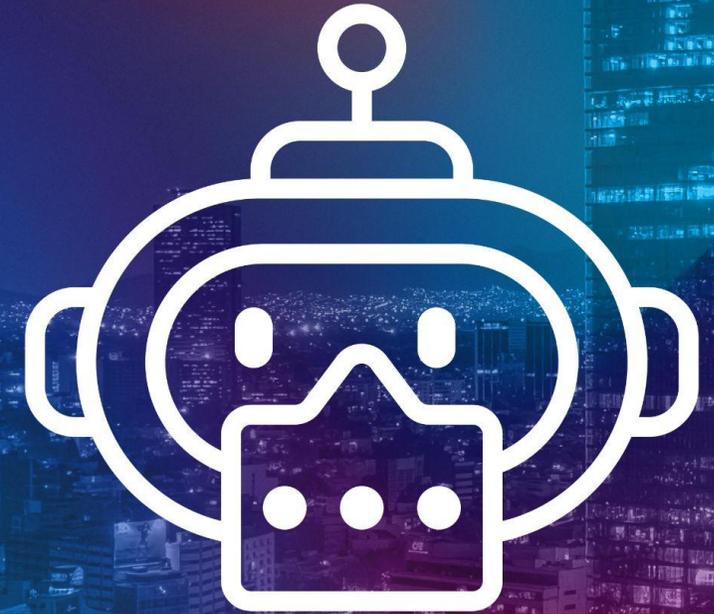
EXPLORE

A GenAI-powered chatbot for Mexico City

Co-creating the future of data-centric AI governance in Mexico

Empowering Thailand's digital government with open data

Open Bangkok: Turning the capital of Thailand into an 'open city'

ODI

Spotlight on: Supporting policymakers worldwide

Governments around the world are increasingly faced with challenges, including climate change, labour market shifts and gender inequality. These need to be addressed through effective policymaking informed by high-quality data. However, the available data is often undermined by gaps, limited resources and inadequate frameworks

GIZ (Deutsche Gesellschaft für Internationale Zusammenarbeit) launched the Data2Policy initiative, aiming to address these challenges by equipping policymakers with tools and frameworks to overcome their policy hurdles. This includes the Data to Policy Navigator, a resource co-designed by the ODI, GIZ and the United Nations Development Programme specifically for government officials with no, or basic, prior data science knowledge and experience. The tool was used in key sessions with policymakers and for co-creating new guidance with GIZ, the Good Practices Guide, as a durable resource.

This year, we used the Data to Policy Navigator directly with policymakers in Iraq, Egypt and India, supporting them in using data and AI to deliver on their policy goals. It maintains an active and credible knowledge platform with 1,800 monthly active users and a strong global audience seeking practical guidance on data-driven policymaking, and has been officially recognised as a Digital Public Good.

EXPLORE

Data2Policy: GIZ and the ODI working together to improve data-driven policymaking

Data to Policy Navigator

Data to Policy Navigator DPG Profile

Case study: Synthetic data in health research

Synthetic data (automatically generated information that mimics real-world data) can be used to facilitate health research, both to address gaps in datasets and/or to preserve the privacy of sensitive data. Using synthetic data without appropriate processes and safeguards, however, can undermine the validity and reliability of findings, compromising the robustness of the research and the products that might result from it.

In the absence of specific guidance on the use of synthetic data in health research, we undertook desk research, interviews, and a workshop to establish the challenges associated with the use of synthetic data, as well as recommended measures that should be considered by those looking to use synthetic data in health research settings.

Our work found ample evidence for synthetic data to be very useful in health research. A number of recommendations emerged that could promote the ethical use of synthetic data in health research. These include the need for clearer, regulator-led guidance in the UK on the use of synthetic data, clear definition and labelling of the data in machine-readable formats, and ethical frameworks on its use to adequately address bias and systemic issues in health data. Since publication, efforts towards this end have been underway. The report we produced has been cited by organisations including HDR UK and the UK Government Office for Science.

EXPLORE

Towards the ethical use of synthetic data in health research
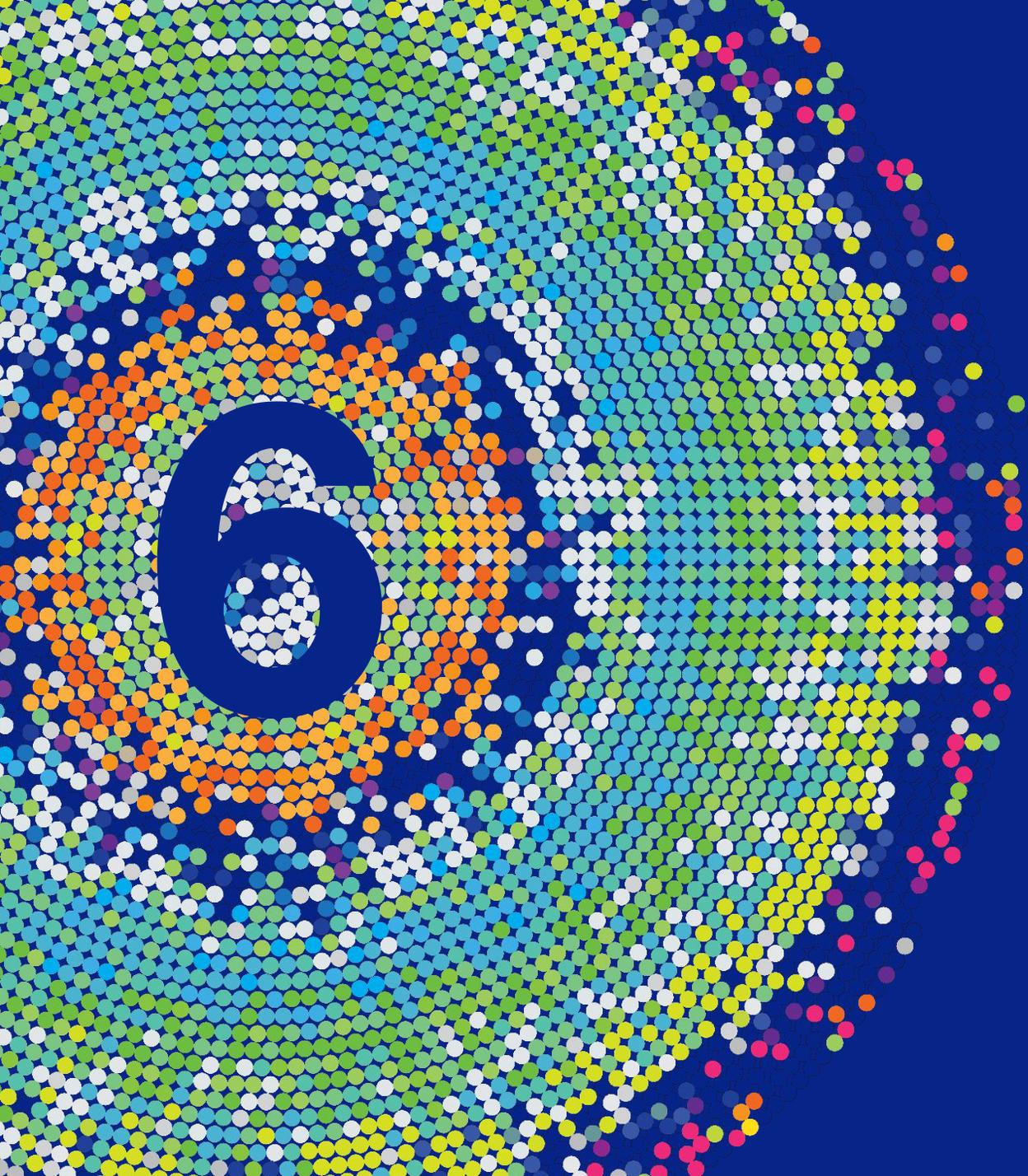
Principle six

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.

ODI

What we achieved in 2025

With much of our learning online, we continue to improve our offering as the data and AI ecosystem changes and adapts. This year, we developed a transparent AI learning assistant to boost attainment while strengthening learners' understanding of how AI works. The results of this have demonstrated that well-designed AI can enhance learning performance while also building the literacy people need to use such tools safely and confidently. We've also improved our systems and integrations, automating workflows for booking courses, enrolling students, issuing certificates, and recertification.

We continue to provide training for Government Campus (via KPMG), which offers learning opportunities for civil servants, and we have enrolled two cohorts for our Data Ethics Professionals course. We are also trialing introductory sessions with organisations that take membership to demonstrate what they can get as part of their membership, such as access to free learning and tools.

Our series of webinars featuring Data Ethics Professionals alumni has proven to be very popular, providing people with a platform to demonstrate how they've applied their learning in their workplace. So far, we've hosted nine webinars, which garnered 2,251 bookings and 785 attendees, with audiences primarily drawn from the UK, Canada and the USA.

Case study: Smart Assessment and Guided Education with Responsible AI (SAGE-RAI)

Personalised coaching can significantly improve learning outcomes, yet most programmes cannot provide individual support at scale. We set out to test whether responsibly deployed AI could close part of the "2-sigma" gap identified by Bloom, without creating opaque or untrustworthy learning tools.

Smart Assessment and Guided Education with Responsible AI (SAGE-RAI) was developed with the Open University and supported Responsible AI UK. It combines retrieval-augmented generation with a transparent design that reveals its prompts, data sources and reasoning. Positioned as a collaborative coach, it helps learners explore ideas, receive tailored guidance and understand the mechanics behind AI assistance. By lifting the lid on the system, the project advances AI literacy while providing accessible, personalised support throughout the course.

In a 10-week study, learners reported high satisfaction (mean = 4.62/5), and qualitative feedback highlighted the value of source-linked responses for checking understanding and maintaining confidence in the learning process. The project provides early evidence that responsible, well-designed AI can both enhance learning performance and build the literacy needed for people to use such tools safely and confidently.

EXPLORE

SAGE-RAI

Learn more about deploying Generative AI responsibly

Understanding data governance in AI

Case study: Data ethics at Global Fishing Watch

Global Fishing Watch is a global non-profit using satellite imagery, VMS tracking and AI to monitor human activity at sea. They provide a free, public and near-real-time view of global commercial fishing activity, safeguarding the global ocean commons. Their challenge is balancing the benefits of their work with potential harm their tools and insights could cause to people and communities.

We partnered with Global Fishing Watch to formalise their data ethics approach to collecting, analysing, modelling and publishing data. We benchmarked their data ethics maturity, upskilled staff through data ethics training, and developed organisational data ethics principles alongside a practical framework to embed data ethics reviews across their Tech and Innovation projects. The framework defines clear roles and responsibilities to make sure

that data ethics reviews are carried out consistently and managed through the right channels in support of responsible decision making.

Global Fishing Watch now has the tools to operationalise data ethics and actively prevent harm, including risks such as data colonialism or bias and discrimination arising from predictive modelling. Benchmarking, clear goals and training has strengthened internal capacity. Organisational principles have created a foundation for consistent, values-driven decisions, while the framework has ensured deliberate assessment of data use to safeguard communities and embed responsible, ethical practice across all key projects.

EXPLORE

Putting data ethics principles into practice

Acknowledgements – our achievements in 2025 would not have been possible without the following partners and funders:

Get in touch

For more information about how the ODI can support your data needs:

- ○ Research
- ○ Consulting services
- ○ Policy
- ○ Membership
- ○ Learning

Open Data Institute   info@theodi.org
Kings Place   +44 (0)20 3598 9395
4th Floor
90 York Way
London N1 9AG
United Kingdom