Ada Lovelace Institute

Input to the Global Digital Compact

About the Ada Lovelace Institute

The Ada Lovelace Institute is an independent research organisation with a mission to ensure data and AI work for people and society. Ada was created by the Nuffield Foundation in early 2018, in collaboration with the Alan Turing Institute, the Royal Society, the British Academy, the Royal Statistical Society, the Wellcome Trust, Luminate, techUK and the Nuffield Council on Bioethics.

We work to create a shared vision of a world where data and AI are mobilised for good, to ensure that technology improves people's lives. We take a sociotechnical, evidence-based approach and use deliberative methods to convene and centre diverse voices. We do this to identify the ways that data and AI reorder power in society, and to highlight tensions between emerging technologies and societal benefit.

Through research, policy and practice, we aim to ensure that the transformative power of data and AI is used and harnessed in ways that maximise social wellbeing and put technology at the service of humanity.

Consultation response

Our submission to this consultation is grounded in the expertise we have built in data protection, use, governance and management over the last four years, and draws heavily on the many public-engagement initiatives we have conducted during this time, including surveys of public attitudes about the use of data and data-driven technologies, and long-form public deliberations ('citizen juries').

Our responses to this consultation are grounded in our understanding of public interest, as evidenced by public attitudes to, and expectations and experiences of, data practices; as well as well-documented understandings of the incentives, benefits and harms of the data economy.

Our response focuses on the five areas below:

- 1. Protect data
- 2. Apply human rights online
- 3. Accountability for discrimination and misleading content
- 4. Regulation of artificial intelligence
- 5. Other areas: Rebalancing digital power for the benefit of people and society

The relevant research for this consultation is listed below.

On public attitudes and experiences of data-driven technologies:

- <u>Who cares what the public think?</u>, May 2022
- <u>The Citizens' Biometrics Council</u>, March 2021
- <u>The role of good governance in building public trust in data-driven</u> responses to public health emergencies, July 2022
- <u>Public dialogue on the ethics of location data</u> (public dialogue, in partnership with the Geospatial Commission), November 2021
- <u>The data divide</u> (survey of public attitudes towards and experiences of COVID-19 technologies), March 2021
- <u>Confidence in a crisis?</u> (public engagement on COVID-19 technologies), August 2020
- <u>No green lights, no red lines</u> (public perspectives on COVID-19 technologies), November 2020
- <u>Beyond face value</u> (survey of public attitudes to facial recognition), September 2019

On data and AI regulation:

- <u>Countermeasures: The need for new legislation to govern biometric</u> <u>technologies in the UK</u>, June 2022
- <u>Regulating Al in Europe: four problems and four solutions</u>, expert legal opinion commissioned by the Ada Lovelace Institute, March 2022
- <u>Regulate to innovate</u>, November 2021
- <u>Getting data right: perspectives on the UK National Data Strategy 2020</u>, November 2020
- <u>Rethinking data and rebalancing digital power</u>, November 2022

On data stewardship: responsible and trustworthy data use, governance and management:

- <u>Participatory data stewardship</u>, September 2021
- Exploring principles for data stewardship, September 2020
- Exploring legal mechanisms for data stewardship, March 2021

On transparency and accountability for algorithms and AI:

- <u>Algorithmic impact assessment: a case study in healthcare</u>, February 2022
- <u>Meaningful transparency and (in)visible algorithms</u>, October 2020
- <u>Technical methods for regulatory inspection of algorithmic systems</u>, December 2021
- <u>Examining the Black Box: Tools for assessing algorithmic systems</u>, April 2020
- <u>Transparency mechanisms for UK public-sector algorithmic</u> <u>decision-making systems</u>, 2020

On the use of data and data-driven technologies as part of the COVID-19 pandemic response:

- Learning data lessons: data access and sharing during the COVID-19 pandemic, January 2021
- <u>Checkpoints for vaccine passports</u>, May 2021
- <u>What place should COVID-19 vaccine passports have in society?</u>, February 2021
- Exit through the App Store? (rapid evidence review of COVID-19 technologies), April 2020

1. Protect data

a) Core Principles

Building trustworthy uses of data as the underlying principle for both data protection and data 'liberation'

The guiding principle for harnessing the potential of data should be building trustworthy uses of data (instead of building trust in the use of data). This would shift focus from a 'deficit model' (which implies the problem that needs solving is the public's lack of trust or understanding of the benefits from data)¹ and

¹ Sturgis, P. and Allum, N. (2004). 'Science in Society: Re-Evaluating the Deficit Model of Public Attitudes', Public Understanding of Science, 13(1), pp. 55–74.

recognises that the challenge lies in creating data practices which can be trusted. This means drawing on principles and mechanisms for responsible data practices, fairness, transparency, and accountability.

Today's data economy is characterised by exploitative data practices which have resulted in widespread erosion of public trust in the use of data, and of trustworthiness in the institutions that use and govern data, including in the public sector.² Creating a trustworthy data ecosystem is therefore critical to ensuring data is used in a lawful, fair and transparent way, but also as a way to realise the social benefits of data, where innovation proceeds with public legitimacy.

Realising the societal benefits of data through effective governance and public participation in data practices

Examples such as the UK's NHS General Practice Data for Planning and Research (GPDPR) initiative (which resulted in over three-million people opting out of data sharing) warn that public anxiety about weak data governance can lead to withdrawal from data sharing, minimising opportunities for socially beneficial innovation.³ This erosion of trust undermines the potential collective benefit that could be extracted through responsible, trustworthy data stewardship in the public interest.

Better use of data by governments, the private sector and not-for-profit organisations requires trust. Trust will not be built merely by providing the public with more information about the reasons why data is used or ways in which it is used. Instead, trust is achieved through trustworthy practices such as creating mechanisms for diverse participation in defining what good uses look like⁴, and ensuring all actors involved in the use of data employ mechanisms for

<u>-sharing/addressing-trust-in-public-sector-data-use</u>

³ NHS Digital (2023). National Data Opt-out, January 2023. Available at: <u>https://digital.nhs.uk/data-and-information/publications/statistical/national-data-opt-out/january-</u> 2023

² For example, the UK Centre for Data Ethics and Innovation (CDEI) report observed that public trust in data sharing by the public sector can be described as 'tenuous'. Centre for Data Ethics and Innovation (2020). Addressing trust in public sector data use. Available at: https://www.gov.uk/government/publications/cdei-publishes-its-first-report-on-public-sector-data

⁴ The Ada Lovelace Institute has developed a participatory data-stewardship framework, which demonstrates that practices that empower people to help inform, shape and govern their own data are preferable to extractive models, as they contribute towards the increasing legitimacy – and public confidence in – the use of data and AI that works for people and society. Ada Lovelace Institute (2021). Participatory data stewardship: A framework for involving people in the use of data. Available at: <u>https://www.adalovelaceinstitute.org/report/participatory-data-stewardship</u>

transparent, accountable and responsible data practices, as well as make available streamlined ways to raise concerns and efficiently receive redress.

b) Key Commitment/ Pledges/ Actions

Strengthening data regulation

Strong data regulation is essential for a trustworthy data ecosystem. Alongside protecting individuals and collectives and supporting the exercise of fundamental rights, one of the main roles of data regulation is to proactively incentivise accountability and responsible behaviours that drive best market practices. There should be commitment from companies to engage with data-protection regulation early in the process, stimulating an organisational culture which is attentive to data rights and making sure clear guidance and clarification is available to innovators. When better regulatory understanding is achieved, this upskills organisations to innovate more effectively and efficiently and brings the best out of technological advancements, offering competitive advantage, confidence and stability.⁵

Comprehensive biometric data regulation

The Ada Lovelace Institute has undertaken robust research on the creation and deployment of biometric technologies.⁶ The recommendations we identified for legislative reform in the UK could be more widely considered by other jurisdictions:

- Pass primary legislation to govern the use of biometric technologies. Legislation must address uses of biometrics for both categorisation and identification, and should apply to the deployment of biometric technologies by both public and private actors.
- 2. The oversight and enforcement of this legislation should sit within a new regulatory function focused on biometric technologies, which is national, independent, and adequately resourced and empowered. The regulatory function should publish a register of public sector uses of biometric

⁵ Ada Lovelace Institute, Centre for Public Data, the Institute for Government, the Open Data Institute and the Royal Statistical Society. (2020). Getting data right: perspectives on the UK National Data Strategy 2020. ODI. Available at:

https://theodi.org/article/getting-data-right-perspectives-on-the-uk-national-data-strategy-2020/ ⁶ Ada Lovelace Institute (2022). Countermeasures: The need for new legislation to govern biometric technologies in the UK. Available at

https://www.adalovelaceinstitute.org/report/countermeasures-biometric-technologies/

technologies, monitor trends and have an ombudsperson to receive public complaints.

- 3. This regulatory function should oversee the assessment of biometric technologies on two levels:
 - It should require that all biometric technologies meet scientifically based and clearly established standards of accuracy, reliability and validity.
 - It should assess the proportionality of biometric technologies in their proposed contexts, prior to use, for those that are used by the public sector, in public services, in publicly accessible spaces, or that make a significant decision about a person.
- 4. There should be a moratorium on the use of biometric technologies for one-to-many identification in publicly accessible spaces and for categorisation in the public sector, for public services, and in publicly accessible spaces until comprehensive legislation is passed.

4. Apply human rights online

a) Core Principles

Fundamental rights need to be reflected in the design of data-driven systems

The Ada Lovelace Institute's Rethinking data and rebalancing digital power report highlights that at the core of making digital systems work for people is the principle of protecting individuals and society from abuse by corporations' or governments' use of data and algorithms. What underlines this principle is that fundamental rights such as privacy and non-discrimination are both protected in law and reflected in the design of computational processes that generate and capture personal data, including artificial intelligence technologies. This means embedding a privacy by design and by default approach in processes from the very beginning and throughout the entire technological lifecycle.

Ex ante fundamental rights impact assessments considered as part of the global regulatory model for AI systems and new technologies

The expert legal opinion commissioned by the Ada Lovelace Institute on the EU Al Act⁷ discusses ex ante fundamental rights-based impact assessment as an

⁷ Edwards, L. (2022). Regulating AI in Europe: four problems and four solutions, expert legal opinion commissioned by the Ada Lovelace Institute. Available at:

important basis for a global regulatory model on AI systems. Comprehensive fundamental rights impact assessments should:

- consider group and societal values as well as fundamental rights and environmental impacts;
- demonstrates efficacy and be determined by whether an AI system is in fact needed at all and if it is consonant with human dignity
- consider the views of individuals, groups and communities actually and potentially affected. In particular, designers should be required by participatory processes to examine how AI systems might be misused in deployment contexts to harmfully impact the vulnerable
- be made public to encourage accuracy, enable scrutiny and provide templates for other providers, especially SMEs. Research should consider if external scrutiny can be supplied by regular audit after the system is deployed instead of, or to supplement, ex ante assessment.

b) Key Commitment/ Pledges/ Actions

Preventing harm at present and in the future

The commitment to protect people from harm needs to operate on a preventative basis. Preventing harms is realised in practice by using trusted mechanisms such as impact assessments and audits, by creating resilient institutions, supporting efficient enforcement action and independent oversight, and strengthening (individual and collective) legal representation mechanisms. A commitment to investment and practical operationalisation of these tools increases the chances of effectively preventing future risks and potential impact that can't be fully anticipated.

There should also be a collective commitment for preventative accountability that builds in participatory structures for citizens to be involved in scrutiny and decision-making, with more explicit consideration and consultation built in with individuals who may face particular risks from algorithmic systems - for example, in regulatory Al risk assessments and the development of standards.

https://www.adalovelaceinstitute.org/wp-content/uploads/2022/03/Expert-opinion-Lilian-Edward s-Regulating-Al-in-Europe.pdf

Shift in mentality and incentives

The paradigms of 'more data is better' and 'move fast and break things' have dominated the data and technology space in the last decades. There is an opportunity now for governments and the private sector to make a commitment for innovative processes that respect fundamental rights and principles such as purpose limitation, necessity and proportionality, with consideration for long-term investments that build robust digital infrastructure and services.

A culture shift will be necessary for more sustainable, trustworthy data use, with a focus on how technology can turn the balance towards innovating and using data responsibly, instead of 'effectively'. Overall, responsible data use should be understood and operationalised to involve taking responsibility not only for how data is used, but also how it is maintained, kept accurate and representative, used appropriately for well-defined concepts of public benefit, and involving citizens in decision-making.

5. Accountability for discrimination and misleading content

a) Core Principles

Transparency and accountability through an 'ecosystem of inspection'

Transparency, accountability, scrutiny and redress are key principles for setting high standards for data processing, minimising risks through appropriate and effective policies, procedures and organisational measures.

The aim should be actively cultivating an ecosystem of inspection where developers and deployers must be open to scrutiny. In practice, this would mean access for researchers and other stakeholders to algorithms and cutting edge AI models, and the ability to carry out external inspections (e.g. independent audits, red-teaming) without running the risk of prosecution (by offering safe harbour for researchers). These provisions would build an 'ecosystem of inspection', which can help overcome the challenges of finding adequate capacity for technical oversight⁸. External researchers representing a diversity of institutions, cultures, demographic groups, languages, and disciplines need to be able to critically examine data driven systems and foundation models from different perspectives.

⁸ Ada Lovelace Institute (2021). Technical methods for regulatory inspection of algorithmic systems. Available at

https://www.adalovelaceinstitute.org/report/technical-methods-regulatory-inspection/

Robust transparency and accountability measures for organisations also need to be coupled with mechanisms for people to critically engage and scrutinise different uses of data and to question the use of technologies with efficient routes for redress. These are key elements of shaping the ambitions for achieving wide public benefits from responsible use of data and Al. Some of these mechanisms are described in the commitments section below.

b) Key Commitment/ Pledges/ Actions

Algorithm audits and impact assessments

There is now greater awareness of the broader societal effects of data-driven misinformation, disinformation, and trolling on democratic participation and debate, with political events, and news scandals bringing previously niche concerns into mainstream reporting.

The Ada Lovelace Institute's Examining the Black Box report⁹ highlights that algorithm audits and impact assessments are a useful tool for regulators and third-party auditors to evaluate the potential harms and benefits of an algorithmic system both prior to launch and after deployment, including for testing for discrimination and misleading content. Our research also proposes that algorithmic impact assessments can foster accountability between institutions developing or deploying algorithmic systems and the people affected by them¹⁰.

An important commitment is setting out requirements for audits and impact assessments, and requiring documentation standards for datasets and algorithmic decision-making systems to be made accessible to third-party assessors.

Access to data for research

Access to data for research is a mechanism that allows accredited researchers and civil society organisations to study and examine the consequences and harms from algorithmic systems such as those deployed on social media

⁹ Ada Lovelace Institute (2020). Examining the Black Box: Tools for assessing algorithmic systems. Available at:

https://www.adalovelaceinstitute.org/report/examining-the-black-box-tools-for-assessing-algorit hmic-systems/

¹⁰ Ada Lovelace Institute (2022) Algorithmic impact assessment: a case study in healthcare. Available at:

https://www.adalovelaceinstitute.org/project/algorithmic-impact-assessment-healthcare/

platforms. A commitment to independent scrutiny over opaque systems is essential for discovering risks and regulating effectively. For example, this mechanism would allow regulators to draw on a wider community of expertise in their regulatory duties, particularly in enabling specialist auditing that regulators may not have the capacity to undertake.

Transparency registers

Governments need to ensure commitment to creating meaningful understanding of public-sector algorithmic practice.¹¹ To be meaningful, transparency mechanisms must cover a full evaluation of the sociotechnical system around an algorithm, including explicit articulation of values being propagated through the systems. Models to consider include Helsinki and Amsterdam's algorithmic registers. Any transparency mechanism should be designed to be accessible to different user groups (publics, regulators, researchers, for example), and transparency will require the ecosystem to be able to be receptive to that information.

6. Regulation of artificial intelligence

a) Core Principles

Effective governance ecosystems for emerging technologies

Regulating AI and algorithms requires effective governance ecosystems that are tailored to the unique challenges and risks that emerge around novel data-driven technologies. They will require new approaches and frameworks to be effective, rather than direct translations of other governance ecosystems, such as aviation safety, finance, food or hygiene. Building effective, trustworthy governance ecosystems for data and AI is paramount to ensuring their benefits are maximised, their risks minimised and public trust in them is earned.

Key aspects of an effective governance ecosystem include:

¹¹ See our article on Meaningful transparency and (in)visible algorithms for a discussion of the challenges in delivering on meaningful transparency, in particular in considering the varying goals for transparency. Ada Lovelace Institute (2020). Meaningful transparency and (in)visible algorithms. Available at:

https://www.adalovelaceinstitute.org/blog/meaningful-transparency-and-invisible-algorithms

1) Cross-sector and cross-jurisdiction cooperation at both national and international levels (and/or potentially creating new national regulators to address regulatory gaps)

Where regulatory action is initiated against major/global players, new measures should be considered as part of an international regulators' fora, that will provide the possibility to create ad hoc enforcement task forces across sectors and geographic jurisdictions, and to institutionalise such bodies, where necessary. The possibility of creating multi-sectoral and multi-geographic oversight and enforcement bodies focusing only on the biggest players in the global data, Al and digital economy should be actively considered.

2) Coherent and holistic enforcement

In order to achieve coherent and holistic enforcement it is necessary to create formal channels of communication between enforcement bodies, to be able to share sensitive information that might be needed in investigations. Currently, many enforcement authorities cannot share important information they have obtained in the course of their procedures with enforcement authorities that have a different area of competence or operate in a different jurisdiction. As data and all-purpose technologies such ChatGPT are currently used by large technology companies, any single enforcement body will not be able to see the full picture of risks and harms, leading to suboptimal enforcement.

b) Key Commitment/ Pledges/ Actions

Regulatory capacity and coordination

The Ada Lovelace Institute's Regulate to Innovate report details key regulatory capacity and coordination functions for AI regulation and makes a series of recommendations.¹² These recommendations translate into commitments that governments could support:

- 1. expanded funding for regulators to help them deal with analytical and enforcement challenges posed by AI systems
- 2. expanded funding and support for regulatory experimentation and the development of anticipatory and participatory capacity within individual regulators

¹² Ada Lovelace Institute (2021). Regulate to innovate. Available at <u>https://www.adalovelaceinstitute.org/report/regulate-innovate/</u>

- 3. the development of formal structures for capacity sharing, coordination and intelligence sharing between regulators dealing with AI systems
- 4. consideration of what additional powers regulators may need to enable them to make use of a greater variety of regulatory mechanisms.

Developing expertise is a challenge as the cross-cutting nature of digital, data and AI technologies can be overwhelming. Ensuring the necessary skills, knowledge and experience are embedded in an organisation is necessary and will require significant investment through expanded funding to attract and retain the relevant 'talent.'

Although the remit of regulating Al systems currently falls within the range of several different regulators (data protection and consumer protection, competition, financial, media regulators), there may be regulatory gaps in which certain aspects of Al systems that must be regulated are not clearly assigned to a specific regulator. It is essential that regulators are given the capacity, resources and powers to identify these gaps and coordinate their regulation of Al systems in a coherent way.

Data sharing among regulators

Shared regulatory capacity could be strengthened by supporting capacity for shared investigations, including enabling data-sharing between regulators and developing shared technical infrastructure, including data science support and investigatory tools for auditing (e.g. a user-friendly platform to enable API access to data held by subjects of investigations). In situations where regulators need to access the same business's data to answer different questions, shared infrastructure will both reduce the costs incurred by each regulator and reduce the overall regulatory burden for businesses as they can deal with a single standardised process across different regulatory access requests.

8. Other areas (please specify):

Rebalancing digital power for the benefit of people and society

a) Core Principles (2500 characters)

Affirmative future vision for data

Today's digital economy is built on deep-rooted exploitative and extractive data practices which have resulted in the accrual of vast amounts of power to a handful of large platforms.

Power imbalances have prevented benefits accruing to people, who are largely unable to control how their data is collected and used, and are increasingly disempowered from engaging in, seeking redress or contesting data-driven decisions that affect their lives.

There is an urgent need for a comprehensive and transformative vision for data that can serve as a 'North Star', directing our efforts and encouraging us to think bigger and move further.

The Ada Lovelace Institute's Rethinking data and rebalancing digital power report offers four cross-cutting interventions (described below) concerning infrastructure, data governance, institutions and participation that can help reduce that power imbalance in the current digital ecosystem and move towards a people-focused vision for data. These interventions are not sufficient to solve all problems, but they represent a realistic first step towards a systemic change.

b) Key Commitment/ Pledges/ Actions

To achieve a people focused vision for digital ecosystems, we propose four cross-cutting interventions as objectives for policy and institutional change:

1. Transforming infrastructure into open and interoperable ecosystems. Today's digital markets are highly consolidated, leaving individuals with little power to switch services. Mandatory interoperability measures would require that products and services from competing companies can work together (such as email and telecommunications). This has the potential to transform digital infrastructure, to

enable innovative services and new experiences for users, preventing 'locked in' and creating greater user choice.

2. Reclaiming control of data from dominant companies through 'access to data mandates' which would allow regulators and researchers to gain access to data in order to study and examine the consequences and harms from algorithmic systems. Access to data mandates can be part of a new 'ecosystem of inspection', which includes algorithmic impact assessments and transparency registers for AI systems, especially systems implemented in the public sector. Independent scrutiny over opaque systems is essential for discovering risks and regulating effectively.

3. Rebalancing power in digital markets with new (non-commercial)

institutions. Large platforms and data brokers currently collect and store large pools of data, which they are incentivised to use for corporate rather than societal benefit. Non-commercial institutions can re-shape the market by introducing different incentives, forcing commercial competitors to compete on quality as well as viewer share. These non-commercial organisations could take the form of public sector organisations, or of alternative data governance models such as data trusts, data cooperatives and data unions that aim to give individuals and communities greater control over data.

4. Ensuring public participation as an essential component of technology

policymaking to make sure that the values, experiences and perspectives of those affected by data-driven technologies are represented and accounted for by policymakers. 'Public participation' includes a range of approaches including citizens' juries, regional deliberation exercises and the co-design of policies, and could provide a more robust evidence base for developing technology policies and practices by building a better understanding of people's lived experiences and values.