

The Ada Lovelace Institute  
Ensuring data  
and AI work for  
people and society

# Foreword

The **Ada Lovelace Institute** (Ada) has been established by the Nuffield Foundation with a £5m commitment, and with the support of founding partners.

The Nuffield Foundation is an independent charitable trust with a mission to advance educational opportunity and social well-being across the UK. It has long been at the forefront of addressing the ethical questions raised by scientific advancements. In 1991, the Foundation established the Nuffield Council on Bioethics, which has been influential in establishing ethical frameworks for policy and regulation relating to innovations in biology and medicine.

Ada has been created with the support of founding partners: The Alan Turing Institute, the Royal Society, the British Academy, the Royal Statistical Society, Wellcome Trust, Luminare, techUK and the Nuffield Council on Bioethics. Ada is named after the mathematician Ada Lovelace (1815–1852), who wrote in visionary terms about computing and its application in the wider world.

Ada Lovelace is famous for some of the earliest ever written reflections on artificial intelligence, arguing of the computer in a statement that was famously framed as the 'Lovelace objection' that: *'The Analytical*

*Engine has no pretensions whatever to originate anything. It can do whatever we know how to order it to perform.'*

Much later, this influenced Alan Turing to intervene with the notion of a 'Turing Test' – the test of a machine's ability to exhibit intelligent behaviour indistinguishable from that of a human.

This debate between perspectives remains a live part of the discussion about the ethical and social implications of artificial intelligence – the extent to which technology is able to, or should, replace or work in partnership with, human intelligence.

We are very grateful to the Trustees and staff of the Nuffield Foundation, particularly Imogen Parker who has led the development of the Institute, for their enthusiasm and support, and to the colleagues from our partner organisations who formed the initial Advisory Board of the embryonic Institute: Dame Colette Bowe, Julie Maxton, Hetan Shah, Genevra Richardson, Martin Tisne, Dan O'Connor, Pete Mills, Helen Margetts and Antony Walker. We would also like to acknowledge the large number of individuals working in this field for their time and generosity in contributing to this project.

**Alan Wilson**  
**Executive Chair, Ada Lovelace Institute**

A new, a vast, and a powerful language is developed for the future use of analysis, in which to yield its truths so that these may become of more speedy and accurate practical application for the purposes of mankind than the means hitherto in our possession have rendered possible.

**Ada Lovelace**

# Contents

2

The Ada Lovelace Institute

3

The context: the data-driven AI revolution

4

The agenda: impacts and ethics

6

Ways of working

8

Work with us

# The Ada Lovelace Institute

The Ada Lovelace Institute is an independent research and deliberative body with a mission to ensure data and AI work for people and society. Ada will promote informed public understanding of the impact of AI and data-driven technologies

on different groups in society. It will guide ethical practice in the development and deployment of these technologies, and will undertake research and long-term thinking to lay the foundations for a data-driven society with well-being at its core.

## Ada has three core aims:

**1**

Build evidence and foster rigorous research and debate on how data and AI affect people and society.

**2**

Convene diverse voices to create a shared understanding of the ethical issues arising from data and AI.

**3**

Define and inform good practice in the design and deployment of data and AI.

## In undertaking its work, Ada will:

- Be outward-facing and collaborative, acting independently of vested interests, and transparent about relationships and funding.
- Be at the forefront of developing change needed to improve people's lives, through a focus on: rights-based approaches; establishing norms; influencing professional practice; technological innovation; regulation and the law; and public dialogue.
- Recognise the potential value of data, algorithms, and AI for individual and social well-being, taking account of human capacity to adapt and respond to new technological challenges.
- Combine reflective deliberation and rigorous research with the need to respond to a rapidly evolving social, technological and economic context.
- Work with other organisations to situate its work in a global context.



# The context: the data-driven AI revolution

We are living through the fourth industrial revolution, driven by computing power and information flows across the internet. This generates and stores huge volumes of data, both on and for, individuals and organisations. The analytics' tools of data science and AI transform both the way we live and business processes in all kinds of organisations – public and private. Some organisations are leaders in these developments, some have as yet barely engaged. Ultimately, these transformations will be all-embracing. No social policy or analysis of the conditions for future social well-being can be put forward without now acknowledging the fundamental and inevitably disruptive impacts that AI will bring.

The tools of AI generate new data and deploy existing data in a number of different ways. The data provides the basis for new analytics and computer modelling. 'Machines' – computers or networks of computers linked by

the internet – can read, see, hear, translate languages and can use data to make decisions, often better than humans. These decisions are made on rules-based processes that can be automated.

When human responsibility for decisions is still needed, AI can be thought of as 'augmented intelligence'. Machines can't yet pass the Turing test and 'think' but they can 'learn'. Health care and many other areas of service delivery will be revolutionised by learning machines: receiving and linking data, and providing inputs for machine learning algorithms. For example, machines can generate precision medicine and rapid medical diagnosis. Treatment plans can then be evaluated and 'fed back' through the system thus creating the learning machine. This kind of application of learning machines is likely to become more prevalent in a range of different contexts.



# The agenda: impacts and ethics

Ada's first objective is to chart the impacts of these technological developments on people and society. Since these are rooted in the data-driven AI economy, analysis will begin there. We will seek to understand the full capabilities of the technologies – for the present and, through horizon scanning, for their future potential. There are benefits to individuals directly – for example through phone apps such as City Mapper or direct connections to taxi services such as Uber – and in the delivery of a wide range of services, for example in health, education, justice, financial services, retail

and the sharing economy, and security. Combined, these benefits create smarter urban (and rural) environments.

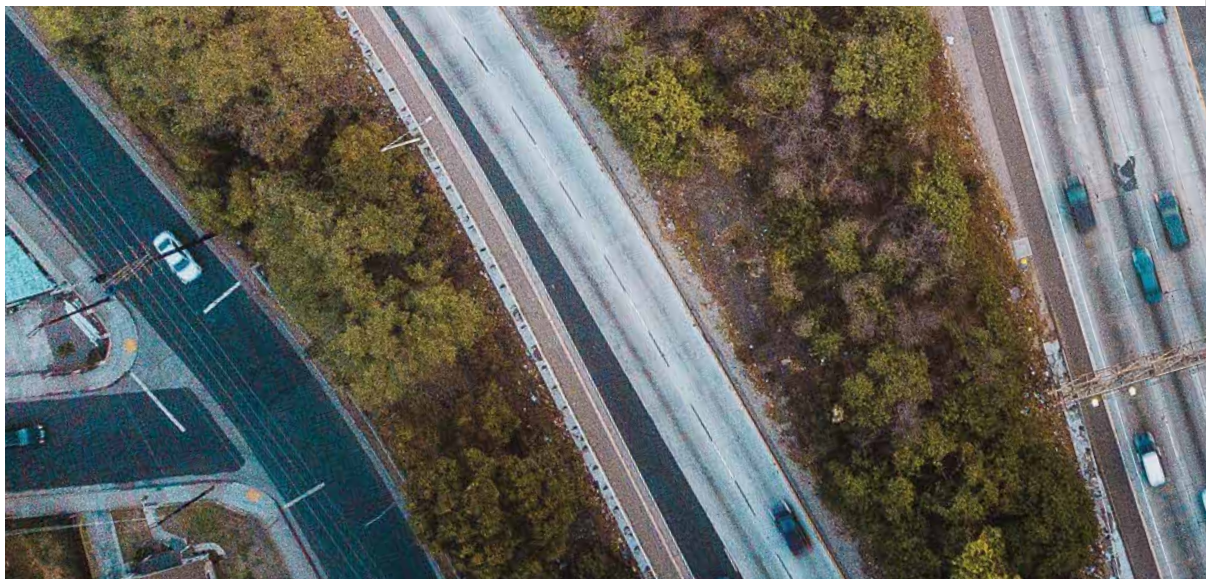
Individual benefits sit alongside social consequences. The impact of digital reforms in the justice system, for example, raises concerns about the rights of the individual in relation to the state; and Uber, though beneficial to customers, has led to controversy relating to the workforce. Individual autonomy is simultaneously enhanced and challenged.

---

## Ethical challenges

These developments therefore inherently generate ethical challenges. The data and the algorithms have to be accurate and reliable and we have to be aware of the consequences of failings in this respect. Many benefits can only be delivered effectively through use of personal data, with analytical power enhanced by linking of data from a variety of sources. This raises issues of privacy and consent. Machine learning-constructed decisions, especially those based on 'deep learning' algorithms, are typically not transparent;

or, because of their training data, can be biased. Are the outcomes fair, and how can we test that? Are lines of what is acceptable being crossed? For example, while marketing and electioneering were relatively open, their methods, however contested, were not seen as a threat to the democratic process itself. But some contemporary methods are now clearly unacceptable – because of their deliberately covert nature, directing 'fake' news and information at targeted audiences to undermine the basis of shared public debate.



---

### Understanding the impacts of innovation

The speed at which these technologies change society, and the dynamism of the markets they fuel, mean that neither the research communities, nor the designers or developers have the space clearly to articulate the conditions for a successful AI-driven society. There is no shared discourse on how we can begin to measure or even account for the social value of

data. In bringing together the disciplines of social science and the humanities with those of data and medical science and technical innovators, Ada's work will offer a deeper understanding of the impacts of technological innovation as they play out across the very different strata and cultures of our increasingly complex, connected, and fragmented society.

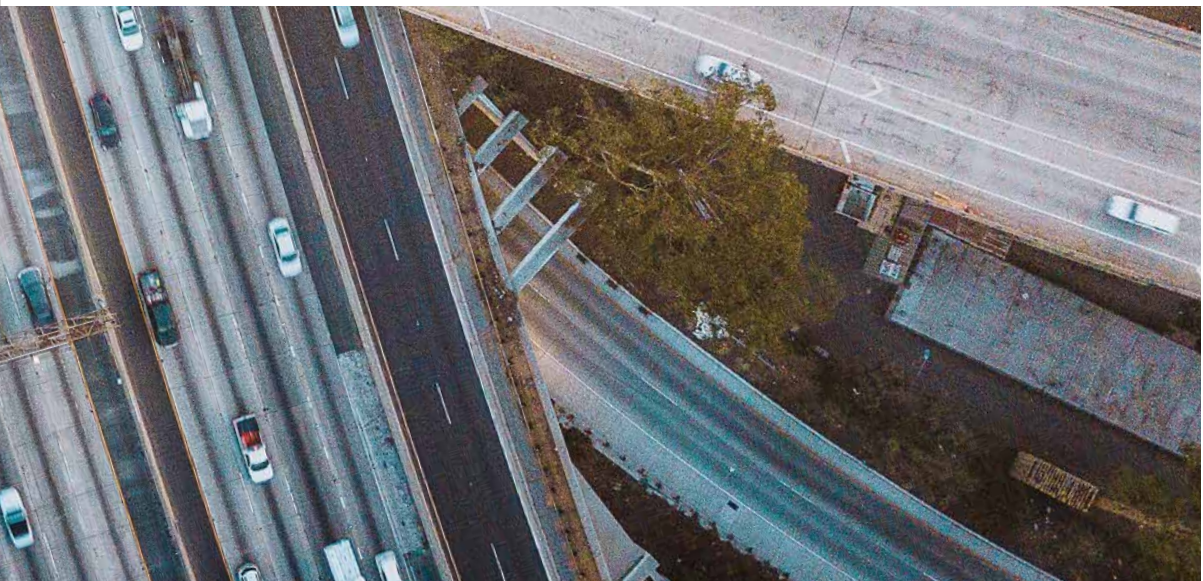
---

### Articulating the costs and benefits

We need to identify which of these issues are capable of technical resolution – transparency for example – and which raise social and political issues that are more challenging. Indeed, insofar as politics is about the relativities of power, we need to explore the networks of control and influence in the collection and use of data, and in the exploitation of data in AI. We need to be able to articulate the costs and benefits of these developments for different groups – the distributive consequences – notably taking into account issues of gender and diversity. Concepts such as explainability, privacy, consent, bias,

fairness and accountability all potentially have different meanings for different groups of people and different cultures; they all have value connotations.

Understanding and contributing to the ethics of data and AI involves unpicking these issues and sharpening our presentation of the vocabulary in multiple dimensions. Whose values? Can we set rules through establishing certain norms and principles, or are we forced into thinking through trade-offs in formulating best practice? All of these questions are priorities for Ada – through research and informed debate.



# Ways of working

The agenda outlined here is substantial in all its sensitivities and nuances: charting the impact of data and AI on people and society; research and debate on ethics; and informing best practice. Ada, both of necessity and as a matter of principle, will work collaboratively – with the range of relevant research communities along the spectrum from data science to social science; with organisations deploying data and AI in the public and private sectors; and with different publics and their representatives in an increasingly diverse society.

The range of research topics is extensive and essentially multidisciplinary. As noted,

we will seek to distinguish the challenges that are essentially technical from those that need social science and humanities lenses. The complexity of the research agenda implies that progress will be most rapid through the exploration of use cases. To develop our understanding of the almost-ubiquitous impact of data and AI on people and society, we will review aspects of how people live – housing, work and incomes, education, health, retail and a wide range of services – along with the organisations delivering the data-driven AI innovations across these dimensions. At a broader scale, we will seek to articulate societal implications.



## A collaborative approach

We are conscious that Ada will function within a complex landscape of exploration and research in these areas, including the Centre for Data Ethics and Innovation established by the Department for Digital, Culture, Media and Sport. We believe our focus of ‘people and society’ in the context

of the impact and ethics of data and AI will be distinctive, and we will ensure that we collaborate widely, communicating and learning in both directions. We are independent, connected to the research frontiers, and hope to provide thought leadership for the longer term.





These are large challenges that no one organisation or approach can solve alone. Ada has been established to research, to convene, and to collaborate.

**To take practical steps in establishing its work programme Ada will work in the following ways:**

- Undertake, commission and catalyse **research projects and methodological innovation**. We will produce working papers and contribute to journals and the media debate. We will work with different disciplines, and seek expertise in their related methods and domains, to build a shared vocabulary of issues.
- Provide **synthesis, analysis and communications** to inform and engage in public debate on data use and its impacts, and to make the uses of AI and its impacts more visible. We will hold events and support fellowships to bridge the gap between academic and technical expertise, public values and practical action. We will initiate working groups around ethical 'test cases'.
- Bring together the perspectives of **industry, government and civil society** to understand the complex systems interacting with data and AI and identify incentives and barriers to ethical practice.

We will work towards the articulation of best practice in all sectors through our research on a variety of use cases.

We will succeed if we can throw light on some of the key questions generated by our preliminary analysis:

1. How should society equitably distribute the power and benefits of data and AI while mitigating harm?
2. In whose interests, and to what purpose, should information be accessed or used?
3. What rights and protections should citizens have and how can they be realised?
4. How should we protect vulnerable, marginalised or disadvantaged groups in society?
5. How can we ensure those who have power to shape society act with public legitimacy?

As these questions show, our agenda is not only substantial but also urgent. The issues raised on impact, ethics and best practice are alive now, and connecting research communities to the public interest is a high priority.

## Work with us

We are keen to hear from organisations and individuals who would like to engage with the work of the Ada Lovelace Institute.

You can contact us at [dataethics@nuffieldfoundation.org](mailto:dataethics@nuffieldfoundation.org), and follow us on twitter [@AdaLovelaceInst](https://twitter.com/AdaLovelaceInst).

For more information on the development of Ada so far, visit our website and sign up to our mailing list at [www.adalovelaceinstitute.org](http://www.adalovelaceinstitute.org).

In addition to establishing the Ada Lovelace Institute, the Nuffield Foundation will continue to fund research, analysis and pilots in the field of digital society and data ethics. For more information about research funding from the Nuffield Foundation, go to [www.nuffieldfoundation.org/research](http://www.nuffieldfoundation.org/research).

---

## About the Nuffield Foundation

The Nuffield Foundation funds research, analysis, and student programmes that advance educational opportunity and social well-being across the United Kingdom.


We want to improve people's lives, and their ability to participate in society, by understanding the social and economic factors that affect their chances in life. The research we fund aims to improve the design and operation of social policy, particularly in Education, Welfare, and Justice.

We have recently established the Nuffield Family Justice Observatory to support the best possible decisions for children by improving the use of data and research evidence in the family justice system in England and Wales.

Our student programmes – Nuffield Research Placements and Q-Step – provide opportunities for young people, particularly those from disadvantaged backgrounds, to develop their skills and confidence in quantitative and scientific methods.

We are the founder and co-funder of the Nuffield Council on Bioethics, which examines and reports on ethical issues in biology and medicine. We have recently established the Ada Lovelace Institute, an independent research and deliberative body with a mission to ensure data and AI work for people and society.

We are a financially and politically independent charitable trust established in 1943 by William Morris, Lord Nuffield, the founder of Morris Motors.



Copyright © Nuffield Foundation 2018  
28 Bedford Square, London WC1B 3JS

T: 020 7631 0566  
Registered charity 206601

[www.nuffieldfoundation.org](http://www.nuffieldfoundation.org)  
[www.adalovelaceinstitute.org](http://www.adalovelaceinstitute.org)

@NuffieldFound | @AdaLovelaceInst

Designed by Soapbox  
[www.soapbox.co.uk](http://www.soapbox.co.uk)