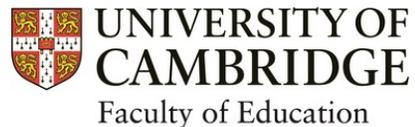


# AI, trust and the public

TRUSTWORTHY AND RESPONSIBLE AI

Cambridge, June 2023

Dr Richard Milne



THE  
KAVLI  
FOUNDATION

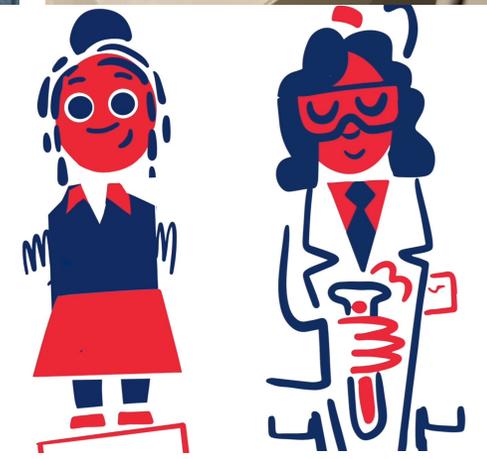


wellcome  
connecting  
science

# The Kavli Centre for Ethics, Science, and the Public



Aim is to experiment with ways of bringing **scientists and publics** together to explore, anticipate and act on the ethical challenges associated with the future of discovery science



# Bringing the public into the ethics of science and AI

**Early Human Embryo Research**  
2023, Life Sciences and Biotechnology

**Pilot Public Engagement Panel on the Use of Public Sector Data (Scottish Government)**  
2022, Data, AI and Robotics

**Public dialogue on genome editing and farmed animals**  
Genomics and genome editing, Life Sciences and Biotechnology

**Net Zero Strategy: Build Back Greener**  
October 2021

**The ethics of location data**  
2021, Data, AI and Robotics

**Future Flight Challenge**  
Other areas of science, technology and innovation

**POLICY FORUM**

**BIOTECHNOLOGY**

## Global citizen deliberation on genome editing

Global governance can be informed by a deliberative assembly composed of lay citizens

By **John S. Dryzek, Dianne Nicol, Simon Niemeyer, Sonya Pemberton, Nicole Bächtiger, Philip Batterham, Bjørn Bedsted, Simon Burall, Michael Burgess, Yurij Castelfranchi, Hervé Chneiweiss, George Church, Merlin Crossley, Jant Mahmud Farooque, Marit Hammond, Baogang He, Ricardo Mendonça, Jenni Anna Middleton, John E. J. Rasko, Ine Van Hoyweghen, Antoine Vergne**

## UK citizens' jury on genome editing

Handbook

13-16 September 2022  
Wellcome Genome Campus

## Artificial Intelligence: Real Public Engagement

RSA

Engaging citizens in the ethical use of AI for automated decision-making

Ada Lovelace Institute | The Alan Turing Institute

## How do people feel about AI?

A nationally representative survey of public attitudes to artificial intelligence in Britain

Independent report

## Public attitudes to data and AI: Tracker survey (Wave 2)

Published 2 November 2022

**Britainthinks**  
Insight & Strategy

## CDEI | AI Governance

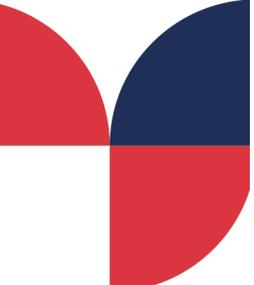
Full report  
April 2022

Oxford Commission on AI & Good Governance

## Global Attitudes Towards AI, Machine Learning & Automated Decision Making

Implications for Involving Artificial Intelligence in Public Service and Good Governance

Lib-Maria Neuber, Atsuki Kuroki, Philip N. Howard  
7 October 2020



# Bringing the public into questions of *trustworthy AI*

1. An ecosystem of trust
2. Trust: What, Who and Why
3. A culture of trust

Trustworthy AI should be:

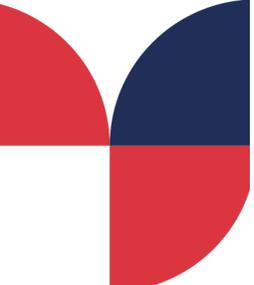
- *Lawful*
- *Ethical*
- Technically and socially *robust*



## Trust in socio-technical systems

While “Trust” is usually not a property ascribed to machines, this document aims to stress the importance of being able to trust not only in the fact that AI systems are legally compliant, ethically adherent and robust, but also that such trust can be ascribed to all people and processes involved in the AI system’s life cycle.





“Maintaining public trust over the safe and secure use of their data is paramount to the successful widespread deployment of AI and there is no better exemplar of this than personal health data.”



HOUSE OF LORDS

Select Committee on Artificial Intelligence

---

Report of Session 2017–19

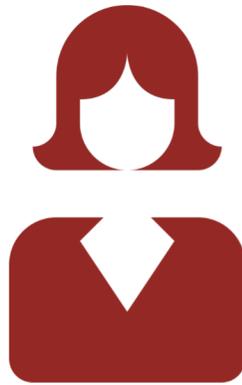
# **AI in the UK: ready, willing and able?**



“...currently AI ethics overloads the notion of trust and trustworthiness and turns it into an umbrella term for an inconclusive list of things deemed “good””

(Reinhardt 2022)

# Trust and trustworthiness



Trust usually related to specific things as a 3 part relation: A trusts B to do X (/with valued item C/ in domain D)

More than 'mere' reliance (acting on the supposition that something or someone will behave as we expect it to)

Trust and trustworthiness

Current controversy

 **Trust and the Goldacre Review: why trusted research environments are not about trust**  
Mackenzie Graham , Richard Milne,<sup>2,3</sup> Paige Fitzsimmons,<sup>4</sup> Mark Sheehan<sup>4,5</sup>

Original research

 **Before and beyond trust: reliance in medical AI**  
Charalampia (Xaroula) Kerasidou , Angeliki Kerasidou , Monika Buscher , Stephen Wilkinson 

# An 'ecosystem' of trust in AI

- Trust essential feature of social and economic interaction
- Means of dealing with uncertainty, vulnerability and social complexity

“Citizens **fear being left powerless** in defending their rights and safety when facing the information asymmetries of algorithmic decision-making, and **companies are concerned** by legal uncertainty. While AI can help protect citizens' security and enable them to enjoy their fundamental rights, citizens also **worry that AI can have unintended effects or even be used for malicious purposes**... lack of trust is a main factor holding back a broader uptake of AI.” (EU white paper on AI, 2020)

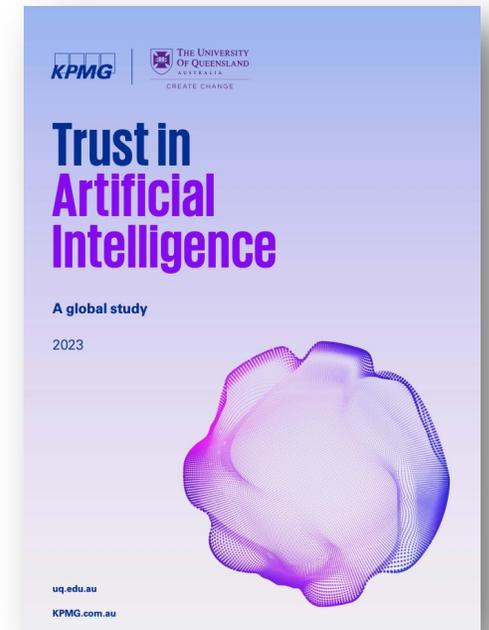


# Who is trusted (and by whom)?



# Trust in AI: Whom do people trust?

- Queensland/KPMG study 2023 on trust in organisations responsible for developing and implementing AI across 17 countries
  - 50% high or complete confidence in universities and research institutions
  - 38% technology companies
  - 33% national government
- When people are confident in entities to develop and govern AI, they are more likely to trust in AI systems
- Trust is a key driver of AI acceptance



# Trust in data

- Health systems and academic researchers are trusted more with health and genomic data than private sector researchers and governments
- Levels of trust vary internationally
- Trust is a strong predictor of someone's willingness for data about them to be used (or to make data about themselves available)

Trust in organisations	(Central) Government	NHS	Social media companies	Big technology companies
To act in your best interests	39%	89%	36%	60%
Keep data about you safe	47%	75%	33%	49%

Please cite this article in press as: Middleton et al., Global Public Perceptions of Genomic Data Sharing: What Shapes the Willingness to Donate DNA and Health Data?, *The American Journal of Human Genetics* (2020), <https://doi.org/10.1016/j.ajhg.2020.08.023>

**ARTICLE**

**Global Public Perceptions of Genomic Data Sharing: What Shapes the Willingness to Donate DNA and Health Data?**

Anna Middleton,<sup>1,2,\*</sup> Richard Milne,<sup>1,3</sup> Mohamed A. Almari,<sup>4</sup> Shamim Anwer,<sup>3</sup> Jerome Atutornu,<sup>1</sup> Elena E. Baranova,<sup>5</sup> Paul Bevan,<sup>6</sup> Maria Cerezo,<sup>7</sup> Tali Cong,<sup>8</sup> Christine Critchley,<sup>3,10</sup> Josephine Fernow,<sup>11</sup> Peter Goodhand,<sup>12</sup> Quratulain Hasan,<sup>3,14</sup> Aiko Hibino,<sup>15</sup> Gry Houeland,<sup>13</sup> Heidi C. Howard,<sup>13,16</sup> S. Zakir Hussain,<sup>14</sup> Charlotta Ingvaldsdottir Matlgren,<sup>16,17</sup> Vera L. Izhevskaya,<sup>19</sup> Aleksandra Jedrzejak,<sup>19</sup> Cao Jinhong,<sup>20</sup> Megumi Kimura,<sup>21</sup> Erika Kleiderman,<sup>22</sup> Brandi Leach,<sup>23</sup> Keying Liu,<sup>24,25</sup> Deborah Mascalzoni,<sup>26,31</sup> Alvaro Mendes,<sup>27</sup> Jusaku Minari,<sup>28</sup> Nan Wang,<sup>8</sup> Dianne Nicol,<sup>10</sup> Emilia Niemiec,<sup>11</sup> Christine Patch,<sup>1,29</sup> Jack Pollard,<sup>23</sup> Barbara Prainsack,<sup>30,31</sup> Marie Riviere,<sup>32</sup> Lauren Roberts,<sup>1</sup> Jonathan Roberts,<sup>1</sup> Virginia Romano,<sup>1,29</sup> Haytham A. Sheerah,<sup>26</sup> James Smith,<sup>4</sup> Alexandra Soulier,<sup>11</sup> Claire Steed,<sup>3</sup> Vigdis Stefansdottir,<sup>33</sup> Comelia Tandir,<sup>11</sup> Adrian Thorogood,<sup>22</sup> Torsten H. Voigt,<sup>34</sup> Anne V. West,<sup>35</sup> Go Yoshizawa,<sup>36</sup> and Katherine I. Morley<sup>35,37,38</sup>

**Summary**

Analyzing genomic data across populations is central to understanding the role of genetic factors in health and disease. Successful data sharing relies on public support, which requires attention to whether people around the world are willing to donate their data that are then subsequently shared with others for research. However, studies of such public perceptions are geographically limited and do not enable comparison. This paper presents results from a very large public survey on attitudes toward genomic data sharing. Data from 36,268 individuals across 22 countries (gathered in 15 languages) are presented. In general, publics across the world do not appear to

Human Genetics  
<https://doi.org/10.1007/s00439-019-02062-0>

**ORIGINAL INVESTIGATION**

**Trust in genomic data sharing among members of the general public in the UK, USA, Canada and Australia**

Richard Milne<sup>1,2</sup>, Katherine I. Morley<sup>3,4,5</sup>, Heidi Howard<sup>6</sup>, Emilia Niemiec<sup>6</sup>, Dianne Nicol<sup>7</sup>, Christine Critchley<sup>7,8</sup>, Barbara Prainsack<sup>9,10</sup>, Danya Vears<sup>11,12,13,14</sup>, James Smith<sup>15</sup>, Claire Steed<sup>15</sup>, Paul Bevan<sup>15</sup>, Jerome Atutornu<sup>1,16</sup>, Lauren Farley<sup>1</sup>, Peter Goodhand<sup>17</sup>, Adrian Thorogood<sup>18</sup>, Erika Kleiderman<sup>18</sup>, Anna Middleton<sup>1,19</sup>, on behalf of the Participant Values Work Stream of the Global Alliance for Genomics and Health

Received: 8 July 2019 / Accepted: 9 September 2019  
 © The Author(s) 2019

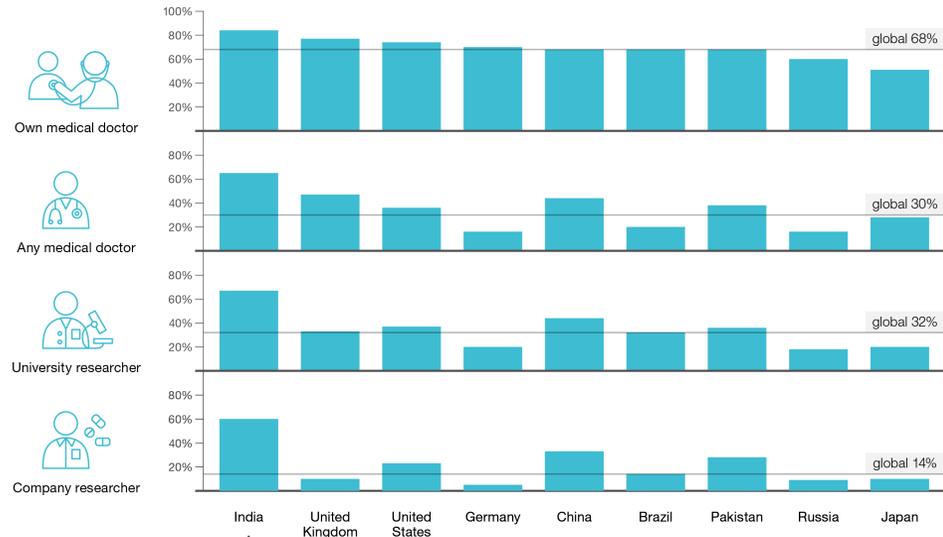
**Abstract**

Trust may be important in shaping public attitudes to genetics and intentions to participate in genomics research and big data initiatives. As such, we examined trust in data sharing among the general public. A cross-sectional online survey collected responses from representative publics in the USA, Canada, UK and Australia (n = 8967). Participants were most likely to



# Genomic data

## How much trust people place in different organisations varies



How much people said that they trust differed across countries.

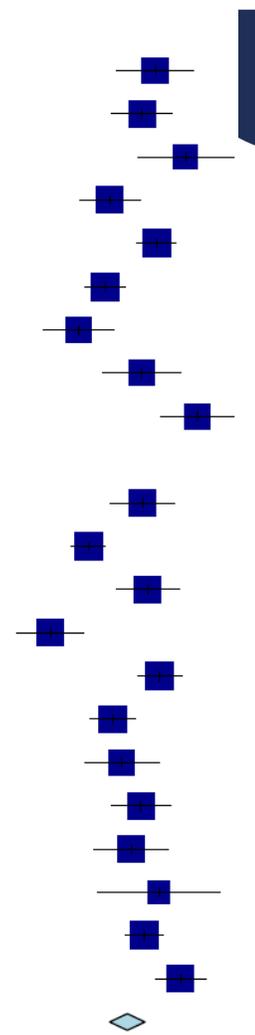
Company researchers were trusted less than university researchers, but the gap was bigger for respondents from the United Kingdom and Brazil than the United States or Pakistan.

Respondents from India were much more likely to trust a range of organisations than anywhere else.

Source	OR (95% CI)
<b>Trust = Yes</b>	
Argentina	4.81 [3.54; 6.54]
Australia	4.34 [3.39; 5.55]
Belgium	6.17 [4.18; 9.11]
Brazil	3.34 [2.62; 4.27]
Canada	4.87 [4.16; 5.71]
China	3.22 [2.74; 3.77]
Egypt	2.59 [1.95; 3.46]
France	4.32 [3.15; 5.91]
Germany	6.78 [5.05; 9.12]
India	0.73 [0.47; 1.14]
Italy	4.36 [3.37; 5.64]
Japan	2.80 [2.45; 3.21]
Mexico	4.53 [3.51; 5.85]
Pakistan	2.06 [1.58; 2.70]
Poland	5.00 [4.18; 5.98]
Portugal	3.41 [2.84; 4.10]
Russian Federation	3.68 [2.72; 4.97]
Spain	4.29 [3.37; 5.47]
Sweden	3.97 [2.94; 5.36]
Switzerland	4.96 [3.03; 8.14]
United Kingdom	4.43 [3.80; 5.15]
United States	5.92 [4.82; 7.27]
Total	3.85 [3.34; 4.44]

Heterogeneity:  $\chi^2_{21} = 174.02 (P < .01), I^2 = 88\%$

Willingness to donate: OR (95% CI)



Middleton et al. 2020 <https://doi.org/10.1016/j.ajhg.2020.08.023>

[https://www.ga4gh.org/news\\_item/public-attitudes-for-genomic-policy-brief-trust-and-trustworthiness/](https://www.ga4gh.org/news_item/public-attitudes-for-genomic-policy-brief-trust-and-trustworthiness/)

# People vary in their propensity to trust

Younger generations, the university educated, and men from emerging economies are more likely to trust AI systems (Gillespie et al. 2023)

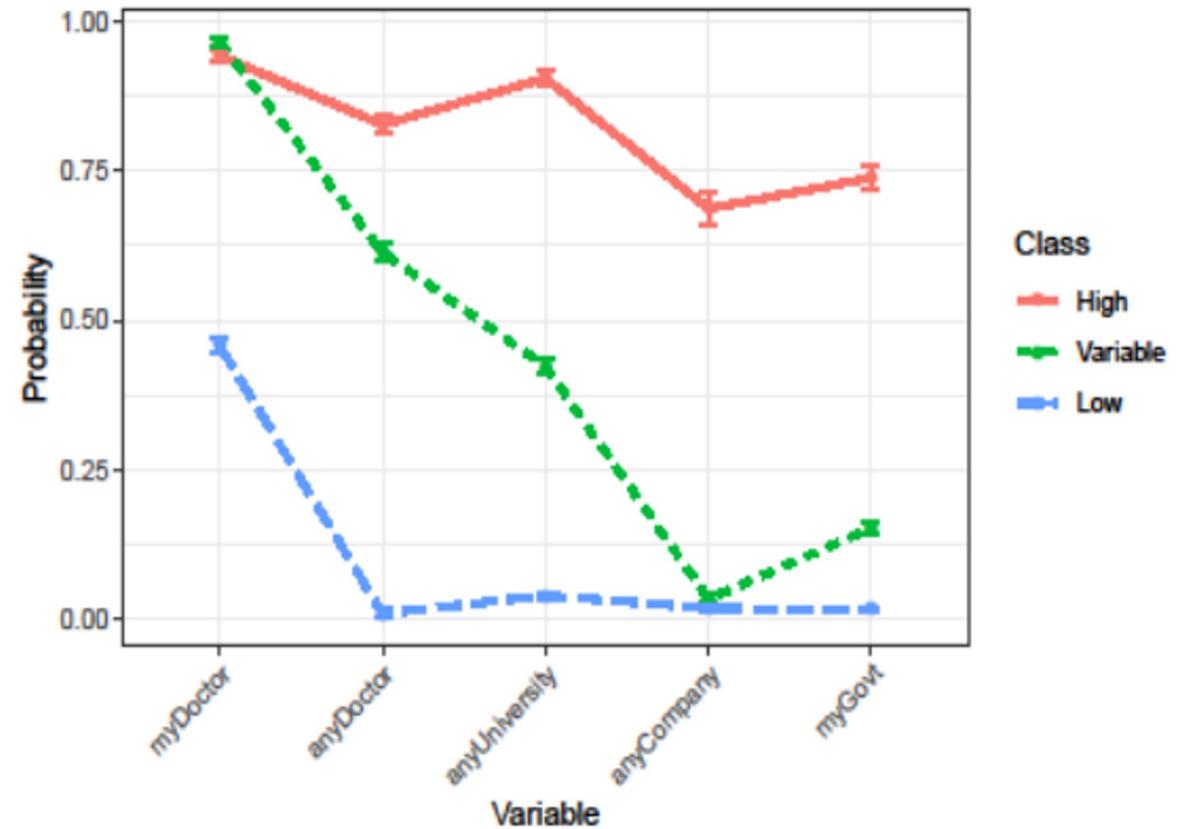
Human Genetics  
<https://doi.org/10.1007/s00439-019-02062-0>

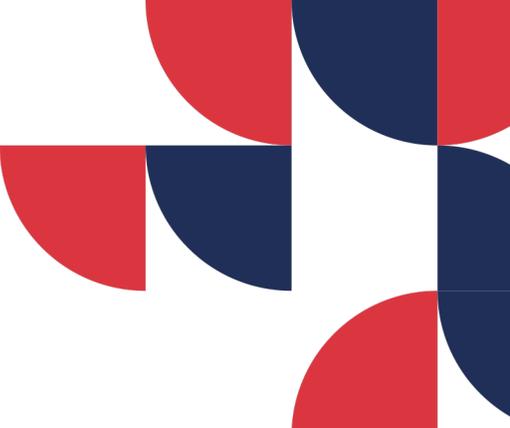
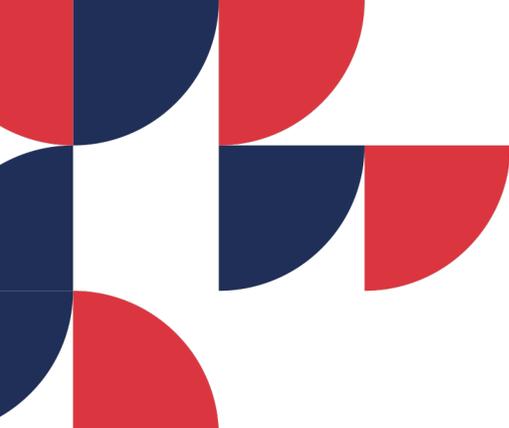
ORIGINAL INVESTIGATION

Trust in genomic data sharing among members of the general public in the UK, USA, Canada and Australia

Richard Milne<sup>1,2</sup> · Katherine I. Morley<sup>3,4,5</sup> · Heidi Howard<sup>6</sup> · Emilia Niemiec<sup>6</sup> · Dianne Nico<sup>7</sup> · Christine Critchley<sup>7,8</sup> · Barbara Prainsack<sup>9,10</sup> · Danya Vears<sup>11,12,13,14</sup> · James Smith<sup>15</sup> · Claire Steed<sup>15</sup> · Paul Bevan<sup>15</sup> · Jerome Atutornu<sup>1,16</sup> · Lauren Farley<sup>1</sup> · Peter Goodhand<sup>17</sup> · Adrian Thorogood<sup>18</sup> · Erika Kleiderman<sup>18</sup> · Anna Middleton<sup>1,19</sup>  on behalf of the Participant Values Work Stream of the Global Alliance for Genomics and Health







# Judging trustworthiness

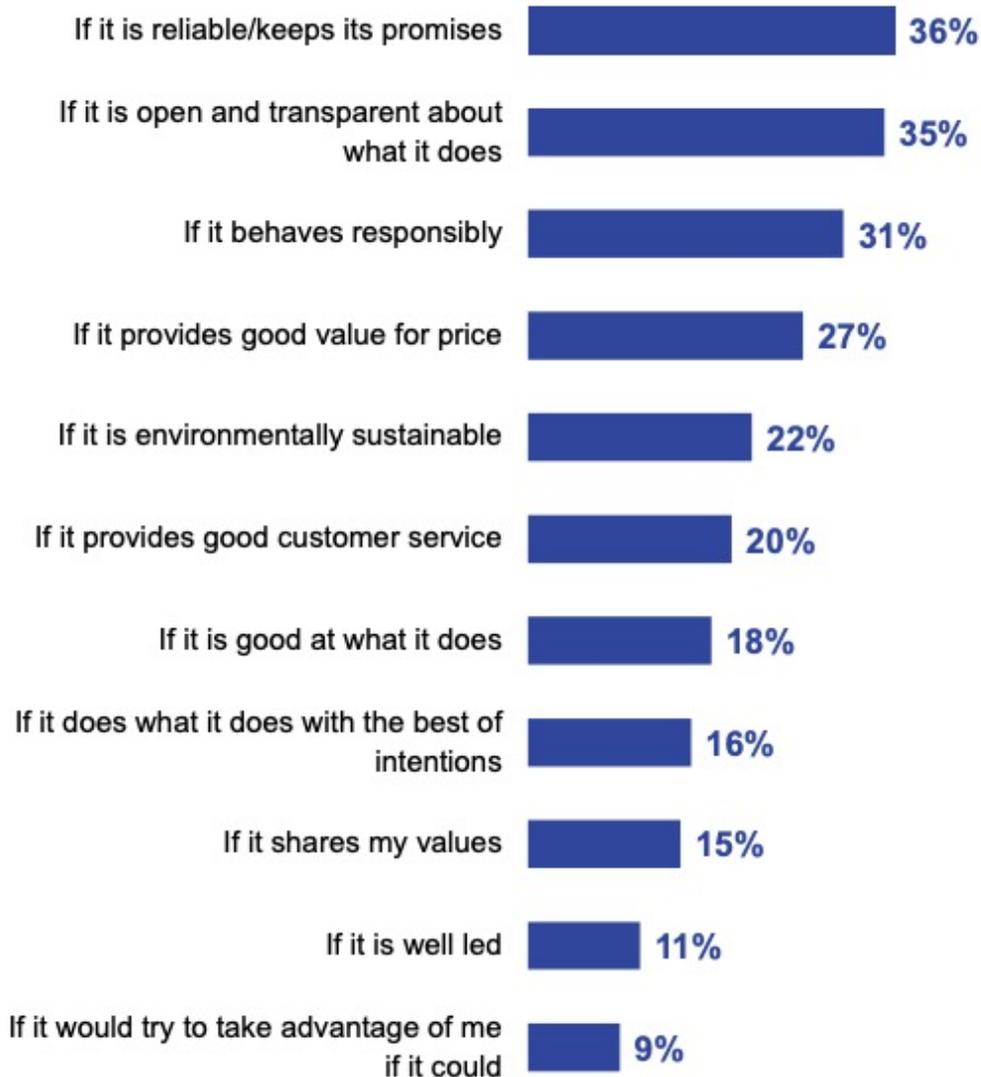
“trust is valuable only when directed to agents and activities that are trustworthy.” (O’Neill 2018)



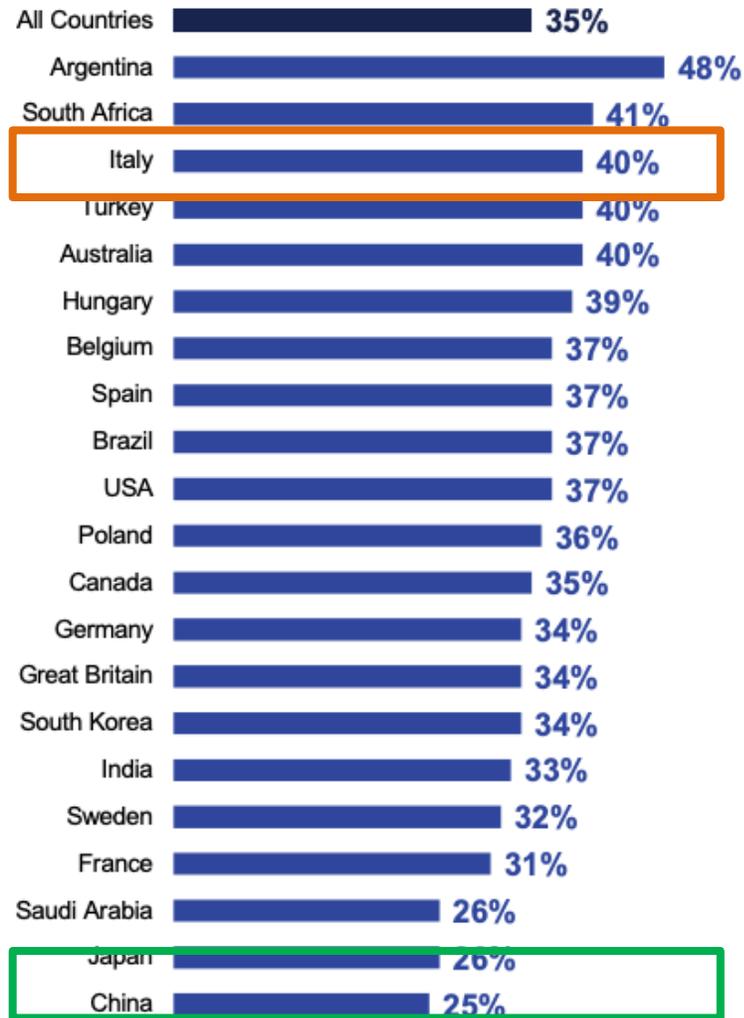
# Global Country Average

Q.  
Which two or three of the following attributes, if any, are most important to you when deciding whether or not to trust an organisation or institution?

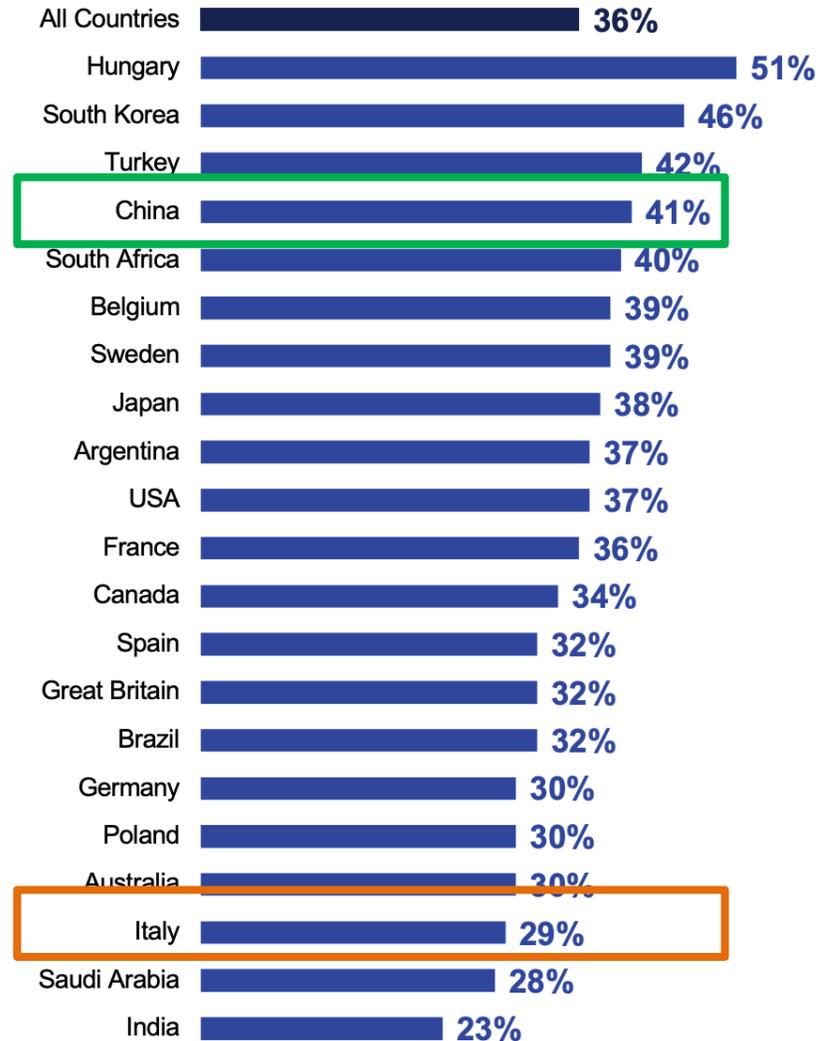
*Ipsos Global Trustworthiness Monitor 2022 - 16017 adults aged 16/18-74 in 21 countries (ca. 500 or 1000 per country), interviewed online 26 August – 9 September 2022  
Online samples in Brazil, mainland China, India, Saudi Arabia, South Africa and Turkey tend to be more urban, educated, and/or affluent than the general population.  
The "Global Country Average" reflects the average result for all the countries where the survey was conducted. It has not been adjusted to the population size of each country or market and is not intended to suggest a total result.*



## If it is open and transparent about what it does

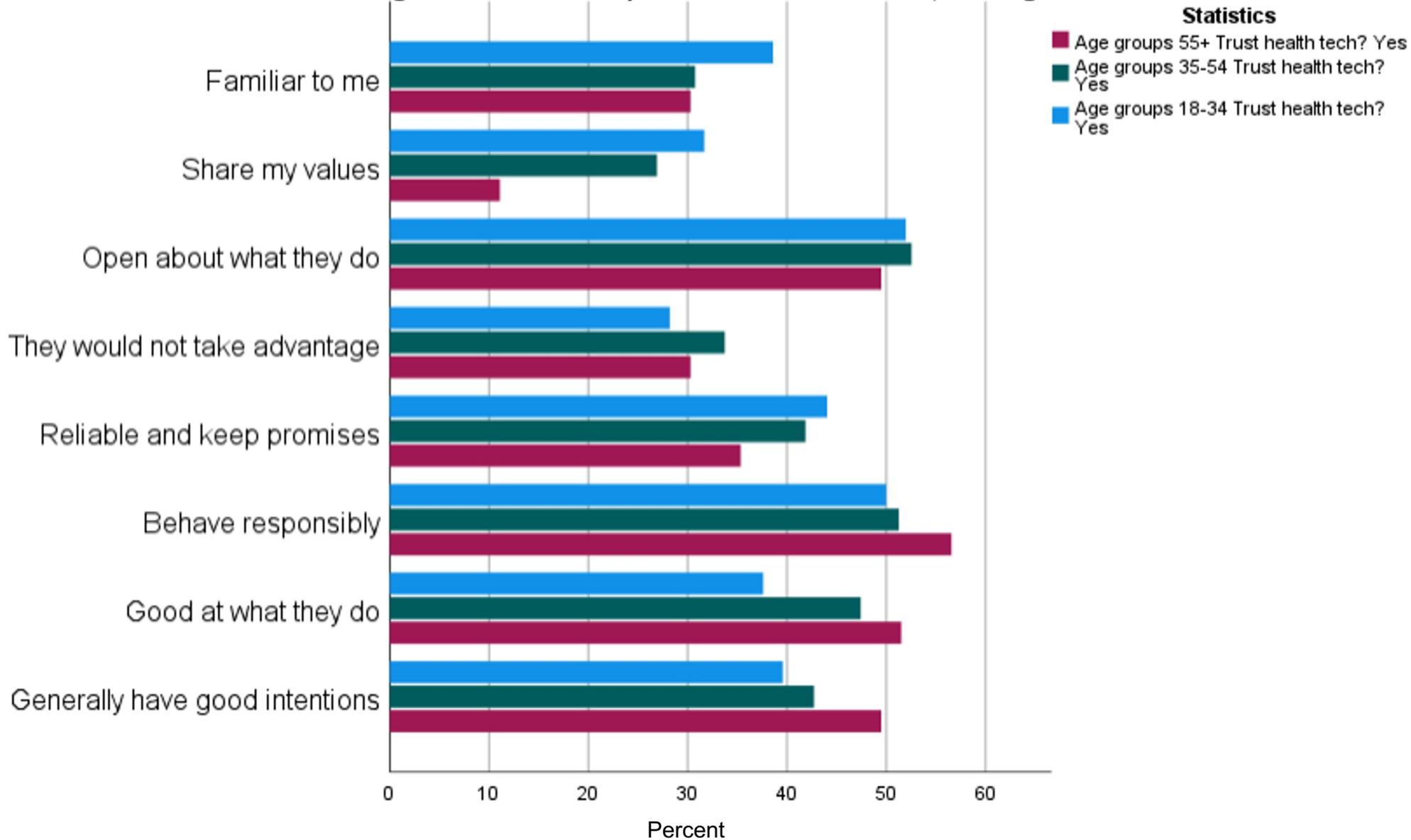


## If it is reliable/keeps its promises



Ipsos  
Trustworthiness  
Monitor 2022

## Reasons for trusting health tech companies with health data, among those who do



Reasons for trusting health tech companies with health data, among those who do

Statistics

**Reasons why one would trust health technology companies with health data, among those who do not**

Row



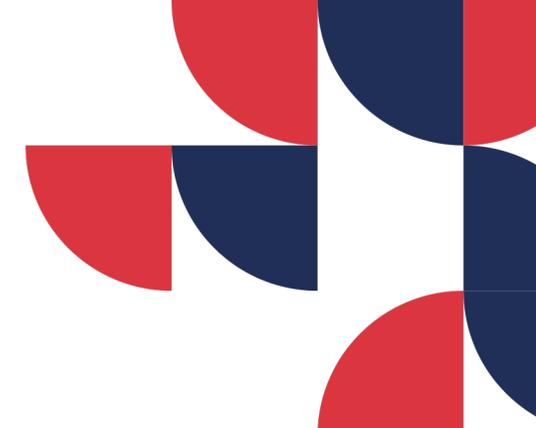
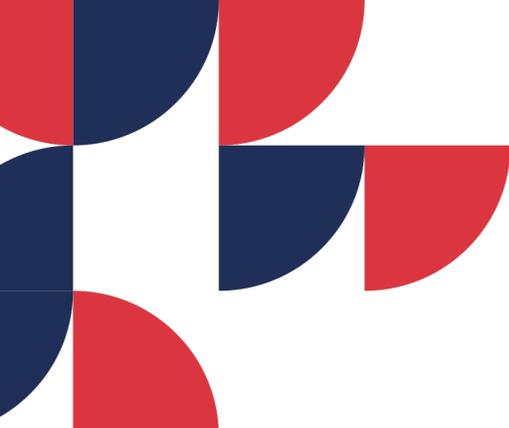
# The qualities of trustworthy AI

- Principles of transparency, accountability, fairness accord with the value of responsibility, openness and keeping promises for trustworthiness
- Emphasised in both quant (e.g. Gillespie et al. 2023; Turing/Ada Lovelace 2023) and qualitative work (e.g. CDEI 2022, AAAS 2021)
- But how they play out may vary by domain of application and by socio-cultural context

Word cloud of responses on what enhances trust in AI systems



Gillespie et al. 2023



# Cultures of Trust





# Trustworthy and Responsive AI

- In some accounts, trustworthiness involves being appropriately responsive to the reason to do what you are being depended on to do
  - Trustworthiness is expressed in action when activated by being counted on. To be trustworthy with respect to A in D thus requires that B be capable of recognizing that A is counting on her and, roughly, what they are counting on her for. (Jones 2012)
- Requires a level of awareness and reflexivity
  - to be fully trustworthy, we need to have a clear vision of what it is that others are trusting us to do, what we are able to offer them, and how social relations may shape our relationship of trust (Potter 2002:27-28)
  - Challenge for AI tools (cf CDEI 2022 work on AI Governance)
  - But also for the AI ‘ecosystem’
- And effectively signalling this to the trustee

# What is it we are being counted upon to do? Building a trust culture

- Establishing effective trust relations is central to aligning AI with societal values and goals
  - We know a lot about who is trusted (or not) to develop AI
  - We know something (conceptually and empirically) about what it might be to be trustworthy to the public
  - We know little about trust as a responsive system
- Focus on *closer* and *responsive* trust relationships between the AI ecosystem and the public
  - **Terms and limits of trust** – align and evolve expectations about when and why trust is needed (including when it's not) and what trusting and being trusted involves
  - **Signals of trust** – understand which *shared* cues we can use to identify both when an individual, organization or system can be expected to respond (be trustworthy) but also so that individuals/organisations/systems know when they are being trusted (rather than relied upon)
  - **Acting on trust** – what kinds of organizations, systems and governance are able to respond and adapt appropriately when they are being trusted
  - Challenges that require **iterative** engagement, communication and deliberation



# Acknowledgements

Anna Middleton

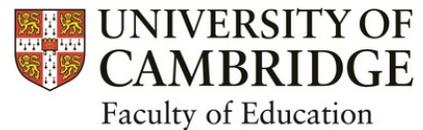
Alessia Costa

Tuba Bircan

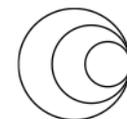
Catherine Galloway

Daniela Boraschi

The Engagement and Society team, Wellcome  
Connecting Science



THE  
KAVLI  
FOUNDATION



**wellcome**  
**connecting**  
**science**