



**Trinity College Dublin**  
Coláiste na Tríonóide, Baile Átha Cliath  
The University of Dublin

# What is the Internet Doing to Me: Ethics on the Internet Emerging Practice with AI and Data

Dave Lewis, [dave.lewis@scss.tcd.ie](mailto:dave.lewis@scss.tcd.ie)

Thanks to: Wessel Reijers, Arturo Calvo, Killian Levacher

## **Student Online Teaching Advice Notice**

**The materials and content presented within this session are intended solely for use in a context of teaching and learning at Trinity.**

**Any session recorded for subsequent review is made available solely for the purpose of enhancing student learning.**

**Students should not edit or modify the recording in any way, nor disseminate it for use outside of a context of teaching and learning at Trinity.**

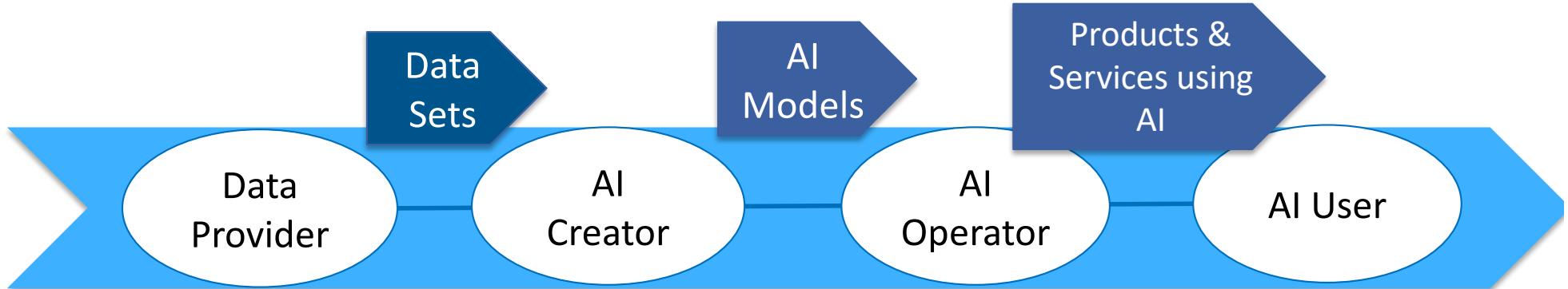
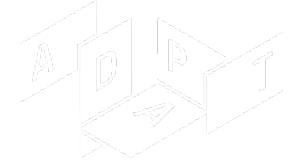
**Please be mindful of your physical environment and conscious of what may be captured by the device camera and microphone during videoconferencing calls.**

**Recorded materials will be handled in compliance with Trinity's statutory duties under the Universities Act, 1997 and in accordance with the University's policies and procedures.**

**Further information on data protection and best practice when using videoconferencing software is available at [https://www.tcd.ie/info\\_compliance/data-protection/](https://www.tcd.ie/info_compliance/data-protection/).**

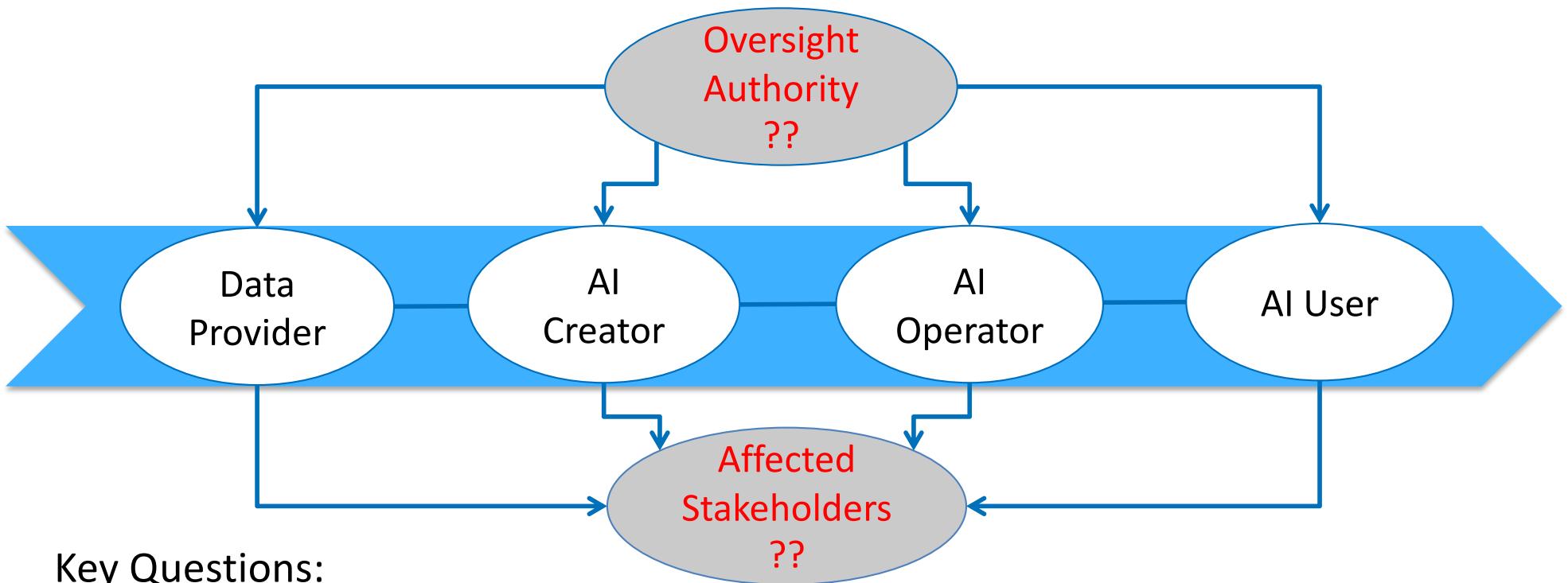
© Trinity College Dublin 2020

# AI Value Chain Roles for Governance



- Data Provider: Official data sources, Data Brokers, You!
- AI Creator: Uses data to build/train AI models
- AI Operator: Uses AI models, perhaps several, in a product or service
- AI User: Decision makers, consumers, You!
- Organizations can take on several roles at once
- AI driven decisions can affect many people beyond User:
  - Your Friends/Family/Community, Patients, Students, Job or Insurance Applicants, Displaced Workers

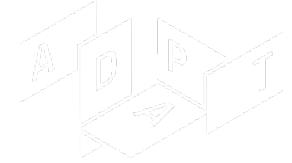
# Value Chain and Societal Governance



Key Questions:

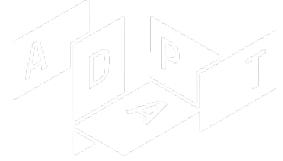
- Which (non-User) Stakeholders are affected?
- Who provides Oversight? Company, Sector, Government? “FDA for Algorithms”
- What Authority do they wield?
- Legitimacy of Oversight for affected Stakeholders?

# Does GDPR help in Governing AI/Data



- Informed consent, data minimization and storage limitations help
- Right to Portability: intended to give power to consumers, but where to port to?
- Pseudo anonymization: if personal information can be extract from data sets it is subject to GDPR
  - As AI get better, more data sets are subject to regulation
  - Industry balance with statistical techniques, e.g. differential privacy
  - Profiling covered: inference of new data that “evaluates personal aspects”
- Right to Explanation and automated decision making:
  - Human explanation and intervention in automated decisions
  - Explainability of machine learning can make this a challenge

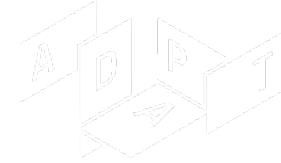
# Approaches to AI/Data Governance



- Self Governance by Organisations
  - Internal tech ethics boards – current examples lack transparency
- Self Regulation by Industries
  - Example: Partnership for AI: <https://www.partnershiponai.org/>
  - Lack of transparency and enforcement
- Government Regulation
  - EU white paper on AI: regulator “high risk” algorithms  
<https://ec.europa.eu/digital-single-market/en/news/white-paper-artificial-intelligence-public-consultation-towards-european-approach-excellence>
  - Labelling of AI projects akin to energy efficiency
    - Ethics much more complicated than energy consumption
- Hybrid:
  - Supplier declaration of conformance for data sets or trained models
  - External certification of declaration processes
- Machine Ethics
  - Stuart Russell: Human Compatible: AI and the Problem of Control



# Emerging International Standards for AI: ISO/IEC SC42



## AI Foundations

- Big Data Reference Framework
- Terms and Concepts
- Machine Learning Framework
- Use Cases

## Trustworthy AI

- Overview of Trustworthy AI
- Neural Net Robustness
- Risk Management
- Bias
- Quality
- Safety
- Explainability
- Ethical and Societal Issues

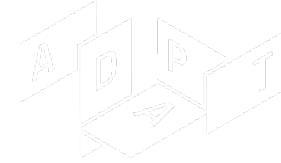
## AI Systems

- Governance
- Systems Engineering
- Management System Standard

<https://www.iso.org/committee/6794475.html>



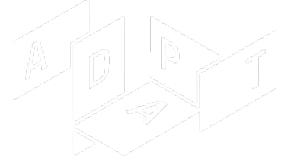
# Trustworthy AI Standards: Some Key Concepts



- **artificial intelligence:** <system> capability of an engineered system to acquire, process and apply knowledge and skills
- **trustworthiness:** ability to meet stakeholder's expectations in a verifiable way [JTC1 AG]
- **stakeholder:** any individual, group, or organization that can affect, be affected by, or perceive itself to be affected by a decision or activity [ISO/IEC 38500:2015]
- **Stakeholder types for Social Responsibility issues** [ISO 26000:2010]
  - Human Rights: everyone
  - Labour practices: workers
  - The environment: future generations
  - Fair operating practices: customers and providers
  - Consumers
  - Local Community involvement and development



# SC42: Social Responsibility for AI



- Ethical and Societal Issues:
  - ISO need international consensus BUT avoids importing specific value-sets
  - Needs principles, which ones?
- ISO already has non-ICT specific principles: ISO 26000 – Social Responsibility
- Identification and engagement with stakeholder is key

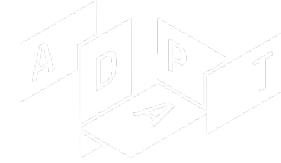
## Principles

- Accountability
- Transparency
- Ethical behavior
- Respect for stakeholder interests
- Respect for the rule of law
- Respect for international norms of behaviour
- Respect for human rights

## Core Subjects (stakeholders)

- Organizational Governance Mitigations (governance board, managers, shareholders)
- Human Rights (everyone)
- Labour Practices (workers)
- The Environment (future generations)
- Fair Operating Procedures (suppliers, customers, regulators)
- Consumer Issues (consumers)
- Community Involvement and Development (local communities)

# Role of Data in Trustworthy AI and Data Ethics

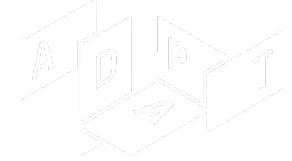


- Problem: for Data, **Possession** is 9.9/10<sup>ths</sup> of the Law
- Power of AI-driven digital engagement grows with the volume (and quality) of its training data
- **Controlling the Flow of Data** is the Key to Governing AI
- Edge Platforms emerging for **maintaining possession** of Personal Data:
  - Inrupt.com
  - hubofallthings.com

BUT

- These are Tech solutions to ethical, legal and societal problems?
- Are there real Market-driven Pathways to adoption?
- Have they Governance structures or Democratic oversight?

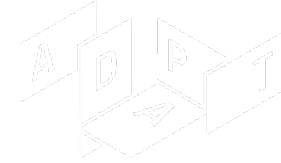
# Could new patterns of Data Stewardship Help?



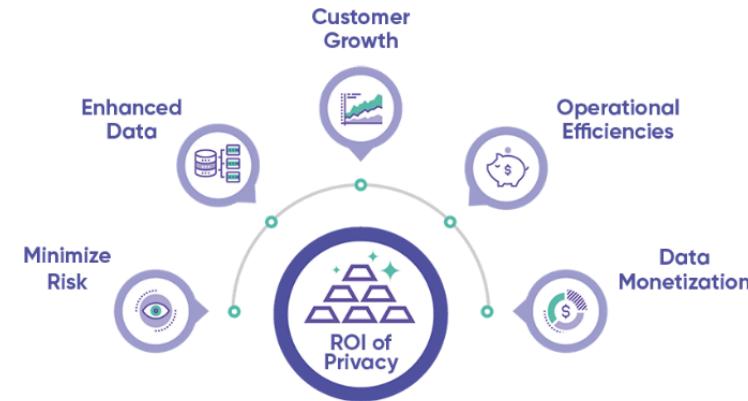
Organizations could already transfer governance responsibility to more representative groups:

- Data Unions – Data as Labour –  
<https://blog.singularitynet.io>
- Data Trusts
- Data Co-ops

# Example: Data Trusts

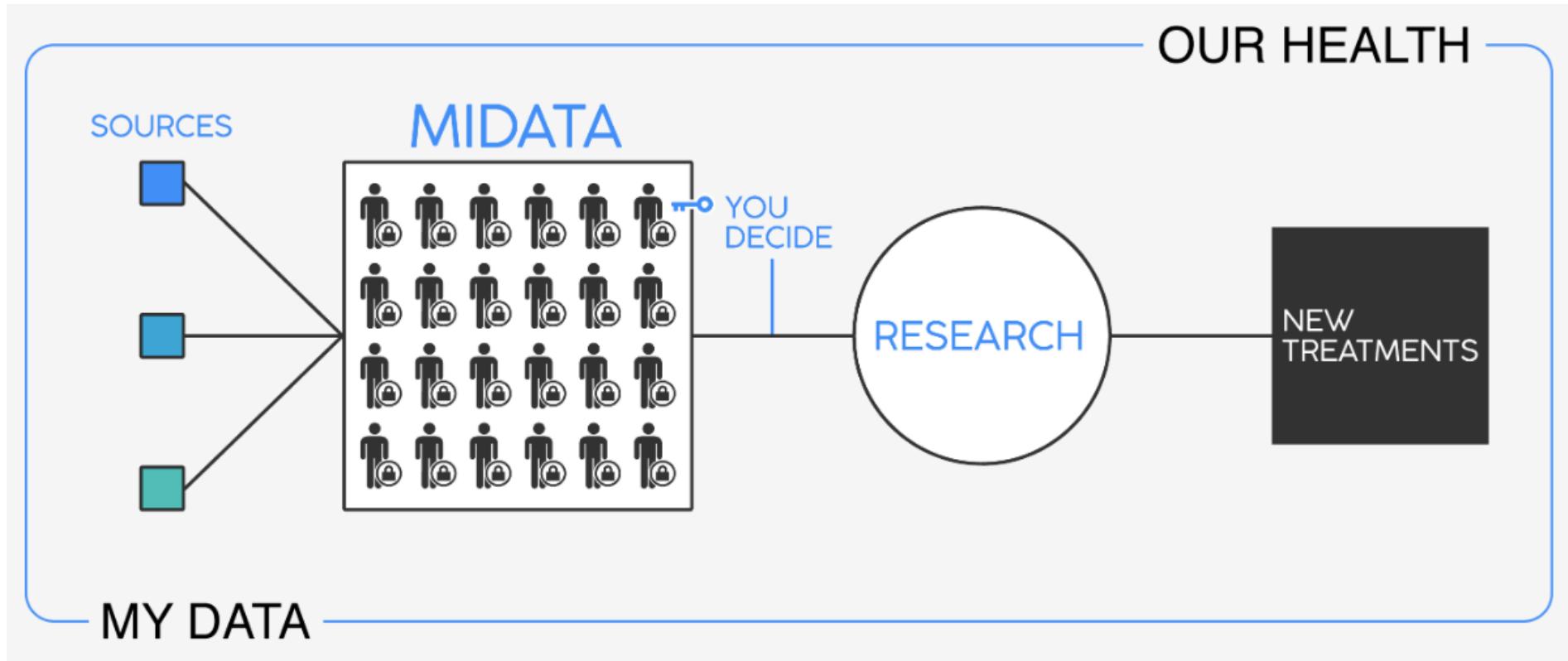
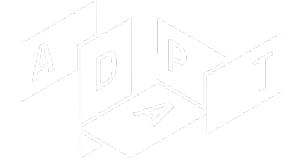


- Truata.com: Anonymised Data Analytics Services



- Offers large clients secure, anonymised analytics services of their own data
- Outsources data protection risks without loss of benefits from data analytics
- Part of business model is a Data Trust - constituted separately to the business/profit driven part of the company
- Data Trust gives clients (and their customers) confidence that the rules can't change for business reasons
- Other examples: <https://theodi.org/article/odi-data-trusts-report/>

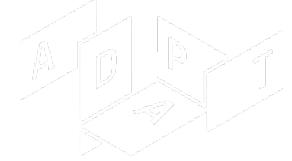
# Example: MIDATA Medical Data Coop



- Control of Data from hospitals and medical studies handed to MIDATA
- Operates as a **cooperative** in the interest of its members medical data subjects
- Management appointed and operated under **democratic principles**

<https://www.midata.coop>

# Are more fundamental data rights required

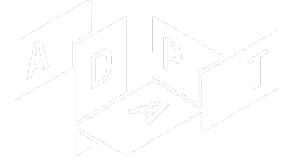


Example: A Rights-Based Approach to Trustworthy AI in Social Media, Dave Lewis, Joss Moorkens, Sept 2020 <https://doi.org/10.1177/2056305120954672>

Example: Data Ownership Right for Data Sharing Contracts in European Data Economy:

1. Non-exclusive Data Ownership Right
2. Secured through evidence of production/contribution – trace logs
3. Right is coupled with an obligation to share data under fair, reasonable and non-discriminatory terms  
<https://www.twobirds.com/en/news/articles/2017/global/data-ownership-in-the-context-of-the-european-data-economy>
4. Further Step: Pool Ownership a lever to negotiate and enforce safeguards

# Why Research Trustworthy AI for Digital Engagement



- Careers that make a difference
  - <https://80000hours.org/problem-profiles/positively-shaping-artificial-intelligence/>
- Public policy and regulation
  - [https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust\\_en](https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en)
- A new factor in Environmental, Social and Governance (ESG) investing
  - <https://www.fool.com/investing/what-is-esg-investing.aspx>
- New sphere of economic activity
  - <https://blog.singularitynet.io>
- New spheres for community and democratic engagement
  - <https://theodi.org/article/odi-data-trusts-report/>
  - <https://www.midata.coop>

# Summary

- As digital tech becomes more powerful and ubiquitous, risks of individual and societal impact and harm grows
- Tech Ethics becoming a priority for governments and companies, e.g. for AI, Big Data, Robotics, IoT etc
- Modern innovation techniques feeding AI and Big Data applications need appropriate forms of ethical consideration – agile, accessible
- Data is Key but new forms of Governance and Oversight yet to emerge
- How to address ethics as a practitioner?

