



Engaging Content
Engaging People

TEU0031: What's the Internet Doing to Me?

Data Ethics

Dave Lewis, dave.lewis@scss.tcd.ie

@ethicscanvas

ethicscanvas.org

hello@ethicscanvas.org

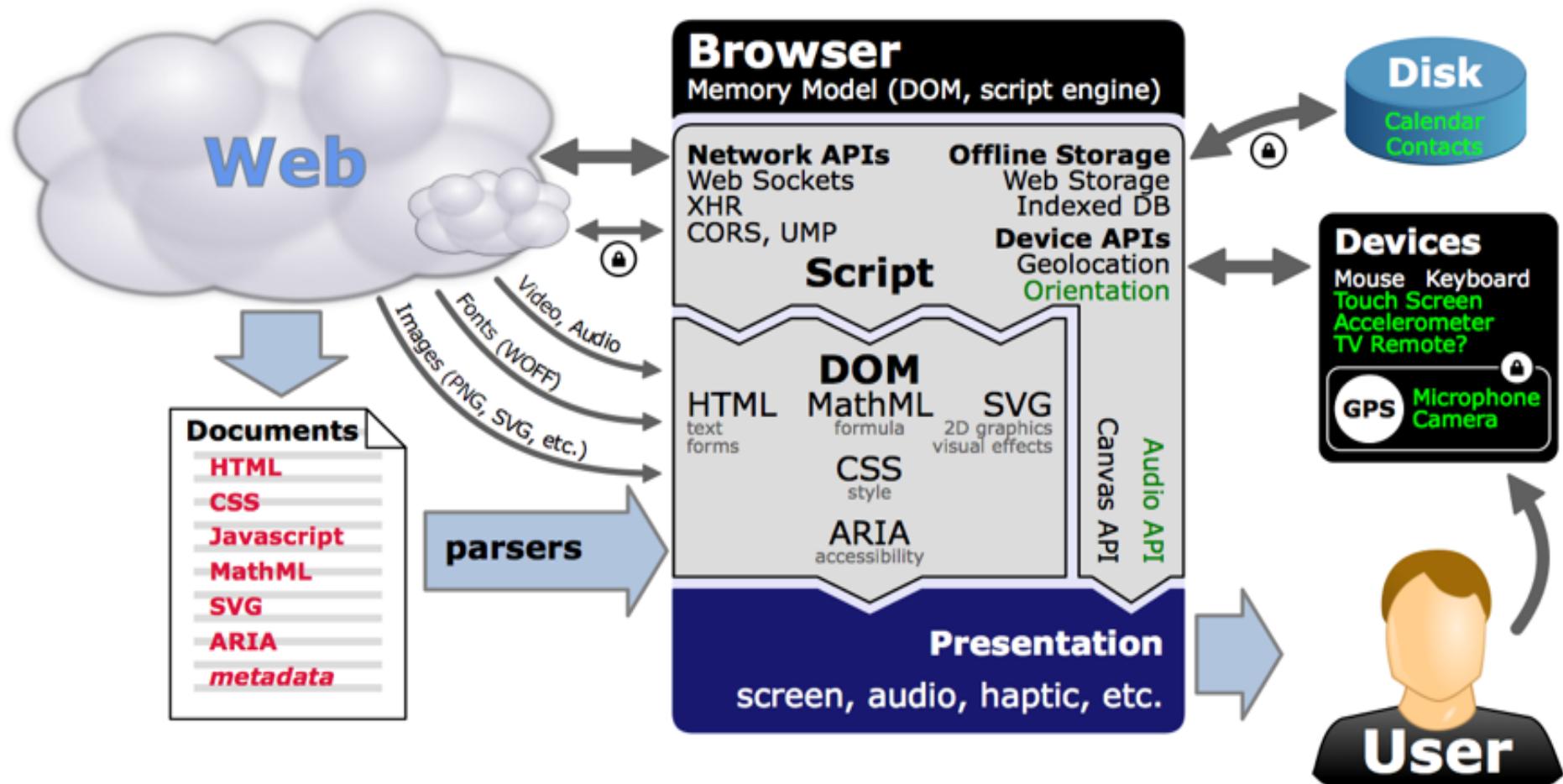


Why Ethics?

- Because new Internet technologies have a **profound impact** on the way **we live**, on the **relationships we have**, on the **societal & political processes we engage in**.
- For Internet tech innovators?
 - First: because it is good for the image of your business (instrumental goal)
 - Second: because it actually improves the service you provide! (substantive goal)
 - Third: because it is the *good* thing to do, it contributes to your idea of a better society and being a good person (normative goal)



Web as an Application Platform – HTML5



Internet: Connecting AI, Big Data and You



- Companies harvest and utilise personal data on a **massive scale**
- Growing concerns about the **collection, linking, use and leakage** of personal data from **mobile devices, bio-sensors, cameras, GPS trackers and social media**.
- Machine Learning deliver new levels of **insights and predictions** about an individual's behaviour and also feed increasingly **personalised AI-driven interactive digital experiences**.
-

Societal and Ethical Challenges of AI and Big Data



- Rate of innovation increasingly **outpaces** public policy debate and the development of new regulation
- Individuals and groups **struggle to understand** the impact of personal information processing
- Companies, especially SMEs, often lack the knowledge and expertise needed to address these **complex legal and ethical issues.**
-



Example: Gender in Google Translate

- Some languages, like Turkish, don't have gender specific pronouns
- Google translate has to guess the gender when translating in English
- Statements allocating gender to role reveal gender bias
- What is the source of this?
- Is it a problem?

Sample Google Translate output:

he is a soldier
she's a teacher
he is a doctor
she is a nurse

<https://qz.com/1141122/google-translates-gender-bias-pairs-he-with-hardworking-and-she-with-lazy-and-other-examples/>



Algorithmic Power on Behaviour

- “Race to the Bottom ... of the Brain Stem”
Tristian Harris
- 70% of YouTube views are based on algorithmic recommendations
- Business model maximises video views to maximise ad views
- Outrage/fear/anger the most reliable reactions that drive us to keep watching
- -> Recommender algorithm inevitably drive us to content that builds outrage to keep us watching

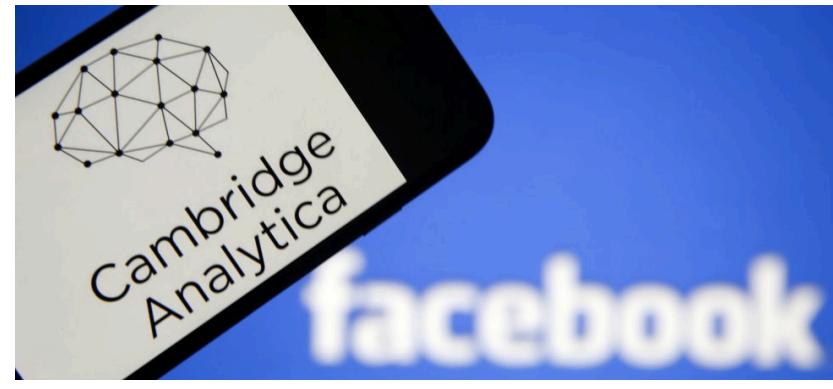


- Evidence to US Congress: <https://www.youtube.com/watch?v=WQMuxNiYoz4>
- Agenda:
<https://humanetech.com/wp-content/uploads/2019/06/Technology-is-Downgrading-Humanity-Let's-Reverse-That-Trend-Now-1.pdf>

Example: Cambridge Analytica

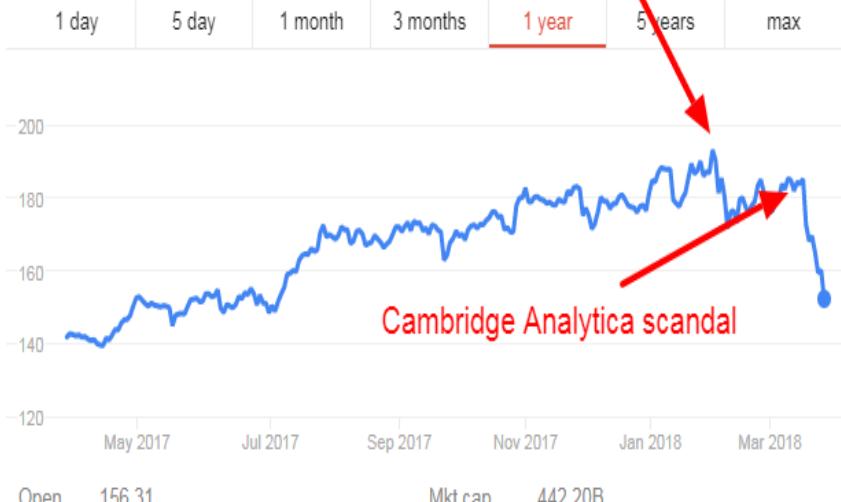
- Academic research into Psychographics (U. Cambridge) revealed the link between philological profiles and Facebook profiles
- Correlated major psychological types to elements in the social graph: Openness, Conscientiousness, Extroversion, Agreeableness and Neuroticism
- Cambridge Analytica applied psychographics to help target political ads in 2016 US elections....

<https://www.theguardian.com/news/2018/mar/17/data-war-whistleblower-christopher-wylie-facebook-nix-bannon-trump>



Facebook, Inc. Common Stock
NASDAQ: FB - Mar 28, 6:15 AM EDT

152.22 USD **↓7.84 (4.90%)** Facebook's share price peak
After-hours: 151.38 **↑0.55%**



Data Risks: Algorithmic selection of digital content



- Manipulation of individuals or groups,
- Diminishing variety that creates biased views and distortion of reality,
- Constraints on communication and freedom of expression,
- Threats to privacy and data protection rights,
- Social discrimination,
- Violation of intellectual property rights,
- Impact on the human brain and cognitive capacity and
- Algorithmic power over human behavior and development.

Latzer, M., Hollnbuchner, K., Just, N., & Saurwein, F. (2016). The economics of algorithmic selection on the Internet. *Handbook on the Economics of the Internet*, (October 2014), pp 395–425. Retrieved from <https://doi.org/10.4337/9780857939852.00028>



Ethics in AI/Big Data: Governance Policy

Big Data Governance is a relatively established field

- International Standards
- Data Protection Legislation – e.g. GDPR

AI Governance spurred by convergence of factors

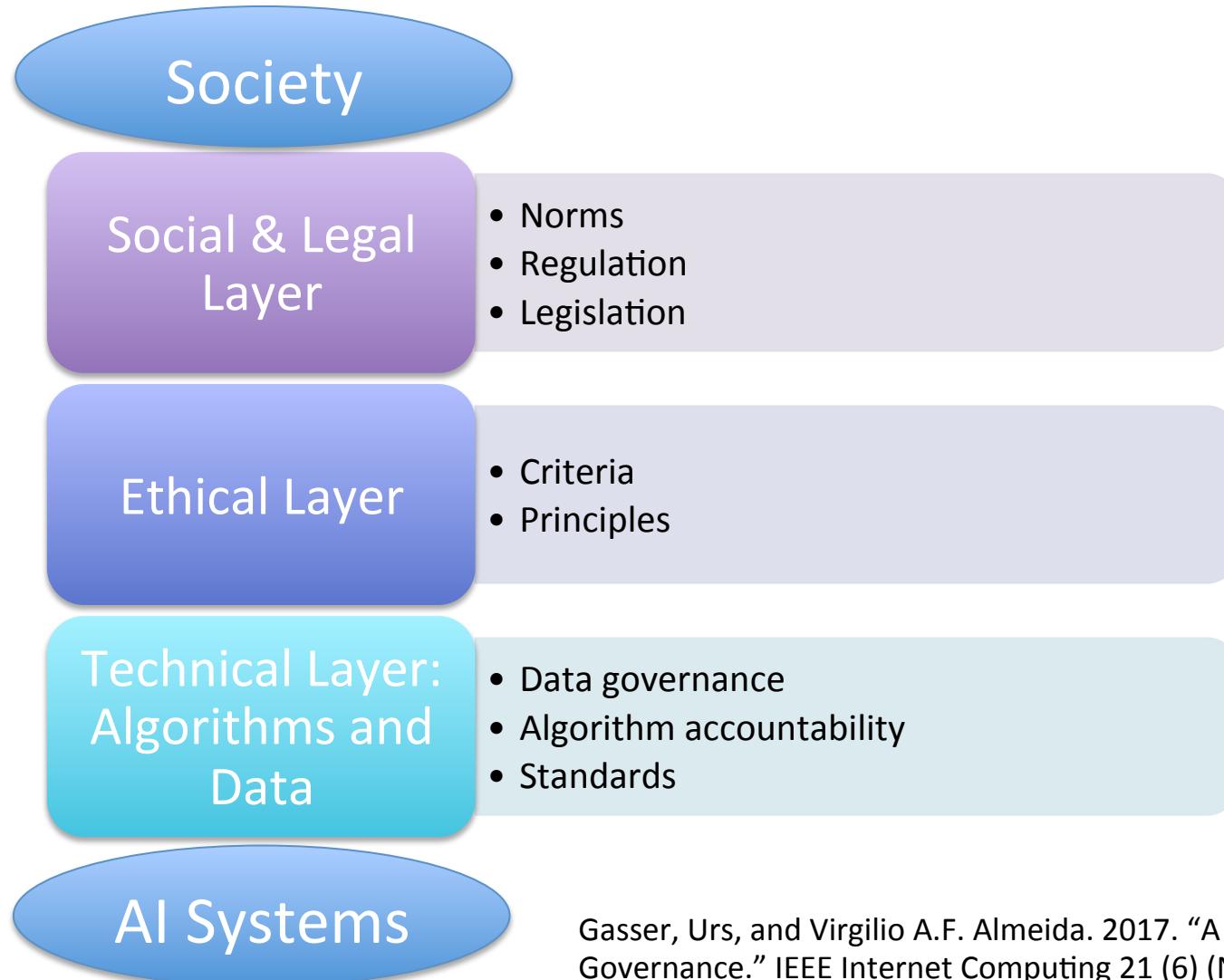
- Successes of machine learning in some human level tasks
- Growing AI adoption in industry
- Examples of AI demonstrating bias or life threatening errors
- Predictions of human worker displacement

Last 3 years,

- Initiatives emerge to address AI Governance
- Industry, national governments, international bodies, standards
- “Ethical” and “Trustworthy” characteristics to the fore
- Tension with perception of AI as a strategic asset by world powers



AI Governance: Layered Model



Gasser, Urs, and Virgilio A.F. Almeida. 2017. "A Layered Model for AI Governance." *IEEE Internet Computing* 21 (6) (November): 58–62.



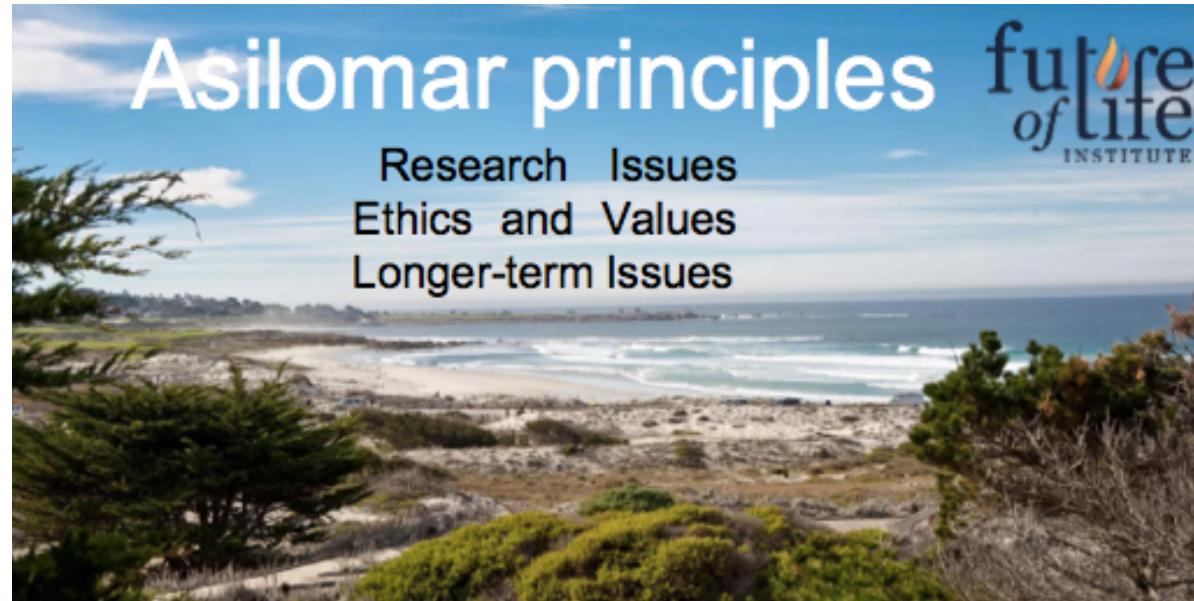
Asilomar Principles

Ethical Principles

- Safety
- Failure Transparency
- Judicial Transparency
- Responsibility
- Value Alignment
- Human Values
- Personal Privacy
- Liberty and Privacy
- Shared Benefit
- Share Prosperity
- Human Control
- Non-subversion
- AI Arms Race

Researcher-led Principles - 2017:

- Research,
- Ethical,
- Longer Term



<https://futureoflife.org/ai-principles/>

IEEE Ethically Aligned Design - 2019

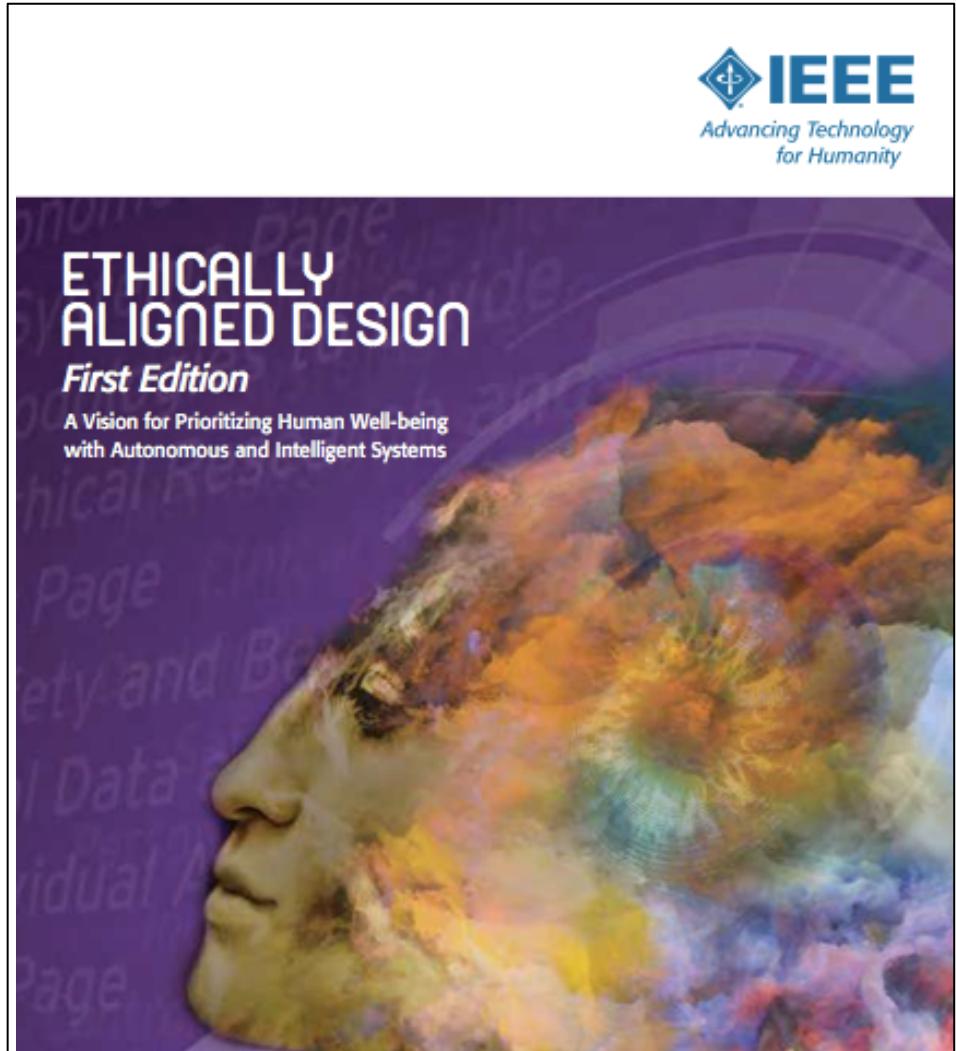


Principles:

- Human Rights
- Well-being
- Data Agency
- Effectiveness
- Transparency
- Accountability
- Awareness of Misuse
- Competence

Comprehensive survey including:

- Classical Ethics
- Affective computing
- Design Methods
- Sustainable development
- Embedding Values
- Policy and Law



<https://ethicsinaction.ieee.org/>

EU Ethics Guidelines for Trustworthy AI - 2019



Ethical Principles mapped from EU Charter of Fundamental Right

International AI Policy Differentiator for EU

Ethical AI, alongside Lawful AI and Robust AI

Requirements

- Human Agency and Oversight
- Technical Robustness and Safety
- Privacy and Data Governance
- Transparency
- Diversity, Non-Discrimination and Fairness
- Societal and Environmental Well Being
- Accountability



EU Ethics Guidelines for Trustworthy AI



Methods

- Technical:
 - Architecture,
 - Ethics/privacy-by-design,
 - Explanation,
 - Testing/validation,
 - QoS Indicators
- Non Technical:
 - Regulation
 - Code of Conduct
 - Standardisation
 - Certification
 - Accountability via Governance Frameworks
 - Education & Awareness
 - Stakeholder Participation
 - Diverse Design Teams



Assessment Pilot Commencing

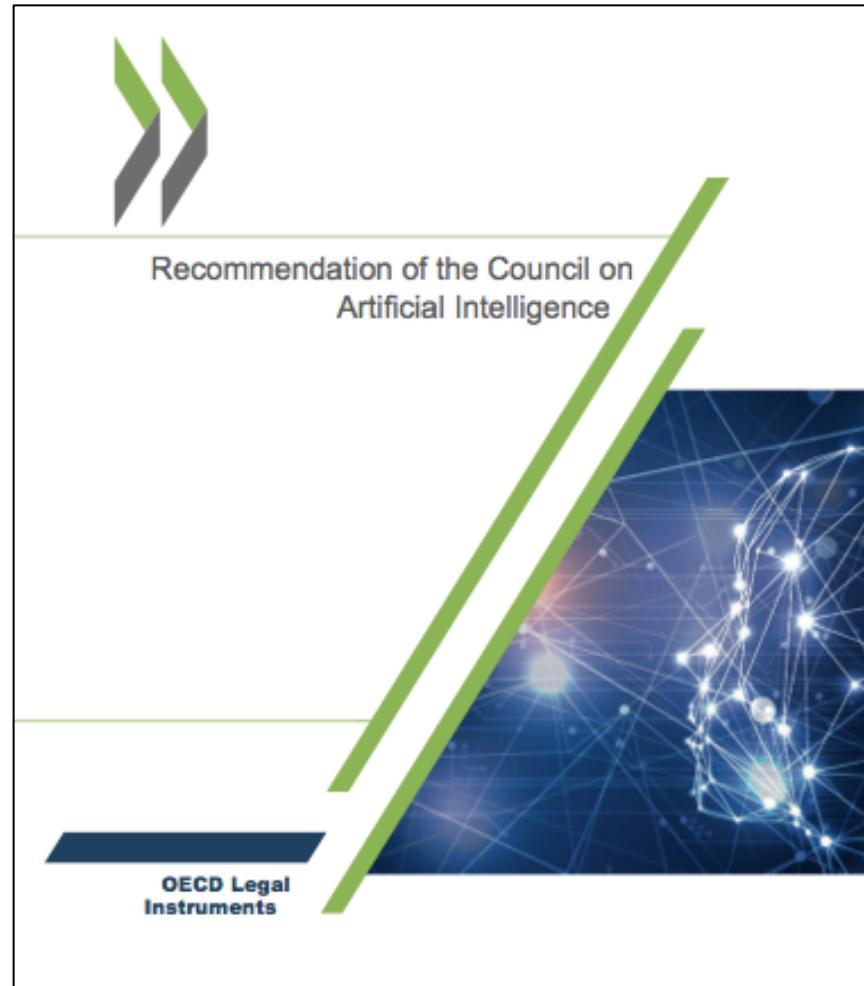
OECD AI Recommendations - 2019

Principles

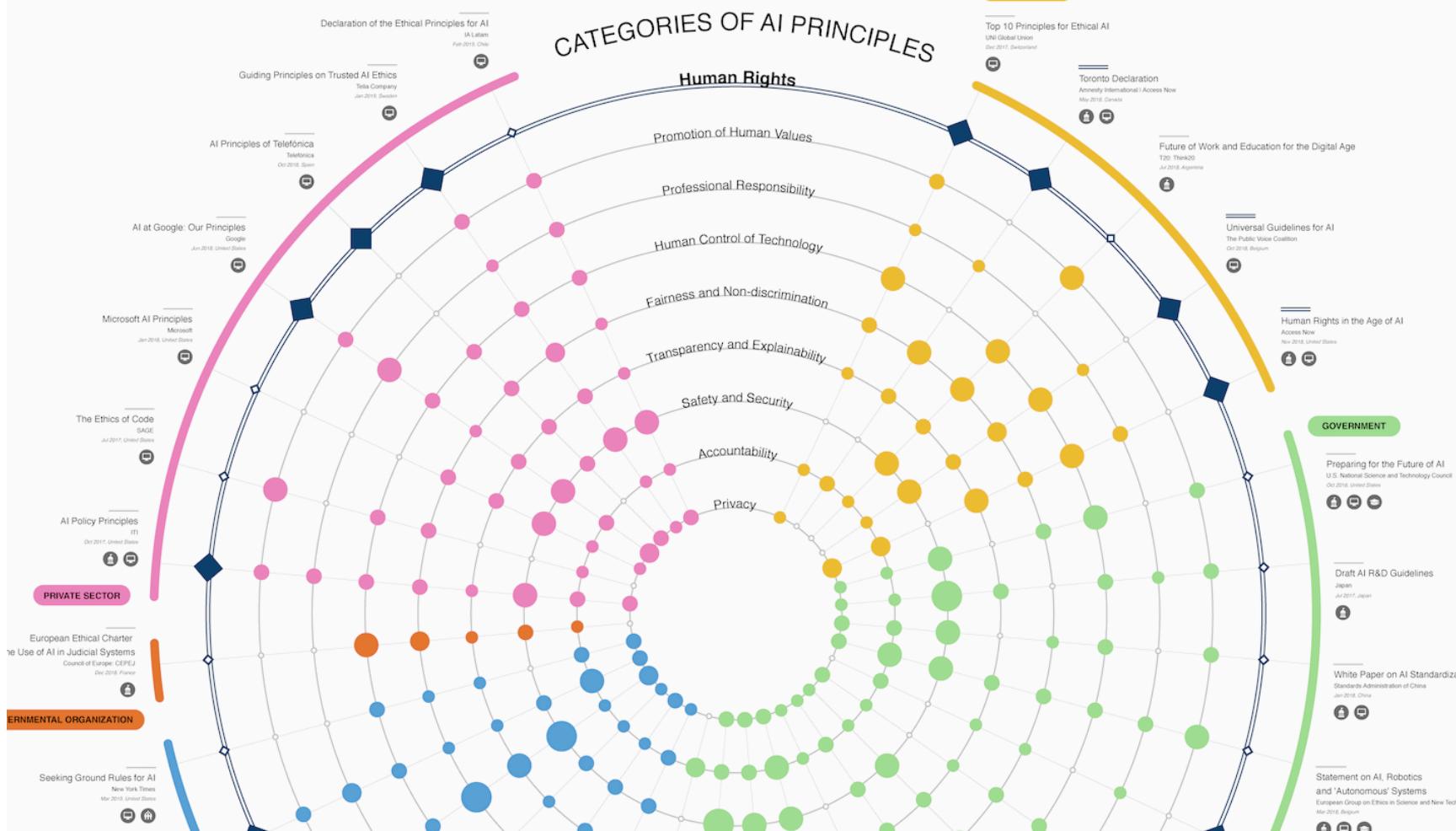
- Inclusive Growth, Sustainable Development and Well-being
- Human-centred values and fairness
- Transparency and Explainability
- Robustness, Security and Safety
- Accountability

Recommendations

- Investing in AI R&D
- Fostering a digital ecosystem for AI
- Shaping an enabling policy environment for AI
- Building human capacity and preparing for labour market transformation
- International cooperation for Trustworthy AI

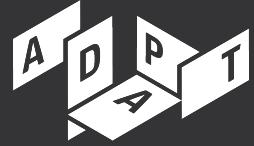


Reports in abundance



<https://ai-hr.cyber.harvard.edu/primp-viz.html>

AI/Data Ethics – where next?



- Many countries and transnational bodies working on principles or guidelines
- Some taking a wait-and-see approach
 - BUT incoming President of European Commission, Ursula von der Leyen, want to legislate for EU regulation ‘within 100 days’
- Danger of becoming a subject for international geopolitics
- AI industry has failed sofar to self-regulate – companies trying their own, e.g. FaceBook ‘Supreme Court’
- Little guidance for industry practitioners

Trustworthy AI Standards: ISO/IEC SC42

ISO/IES SC42 Artificial Intelligence

- subsumed Big Data

Current activities:

- Overview of Trustworthy AI
- Risk Management
- IT Governance
- Neural Net Robustness
- Bias in AI
- AI Ethical and Societal
- Social Responsibility



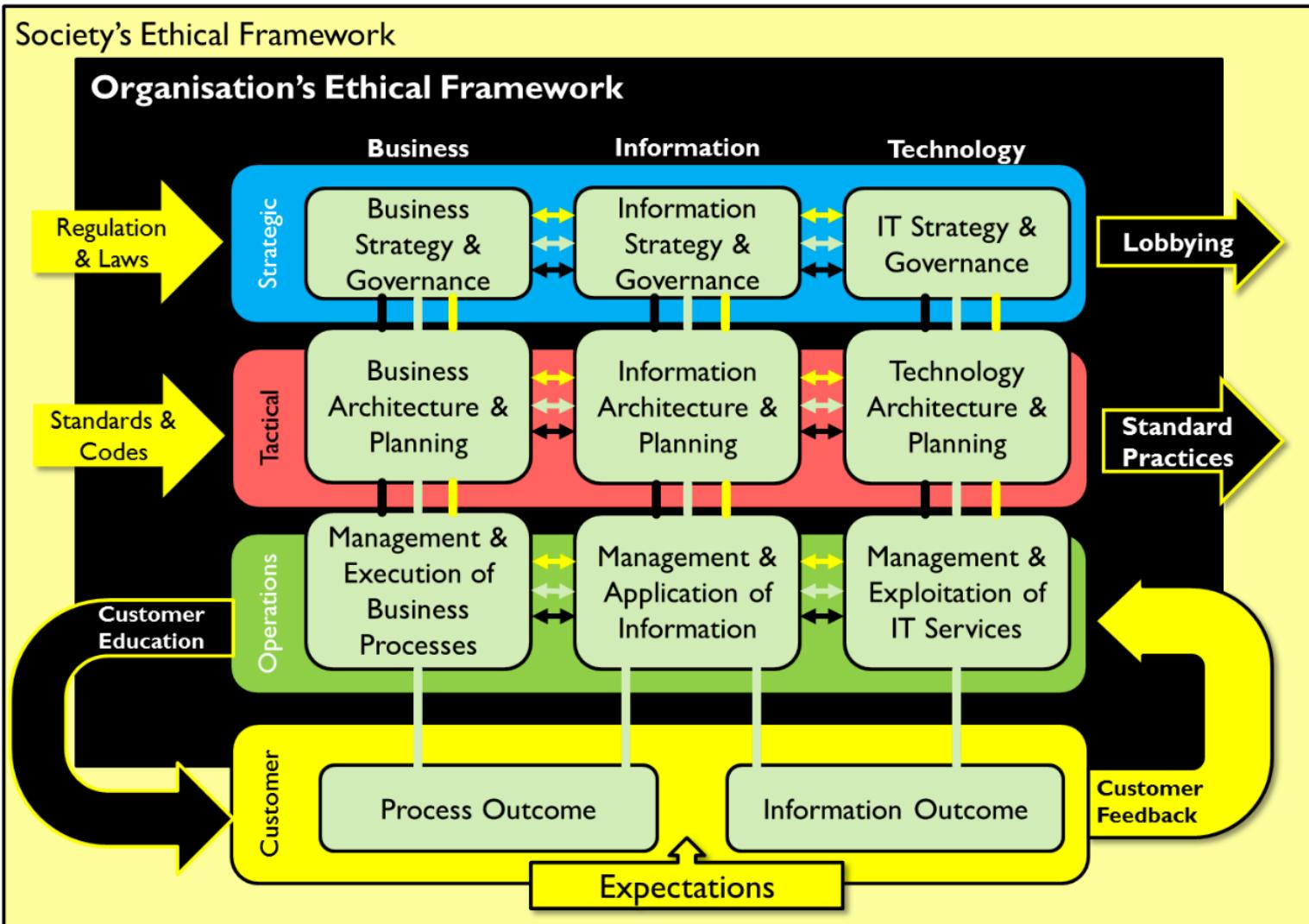
Data Ethics: Organizational Viewpoint



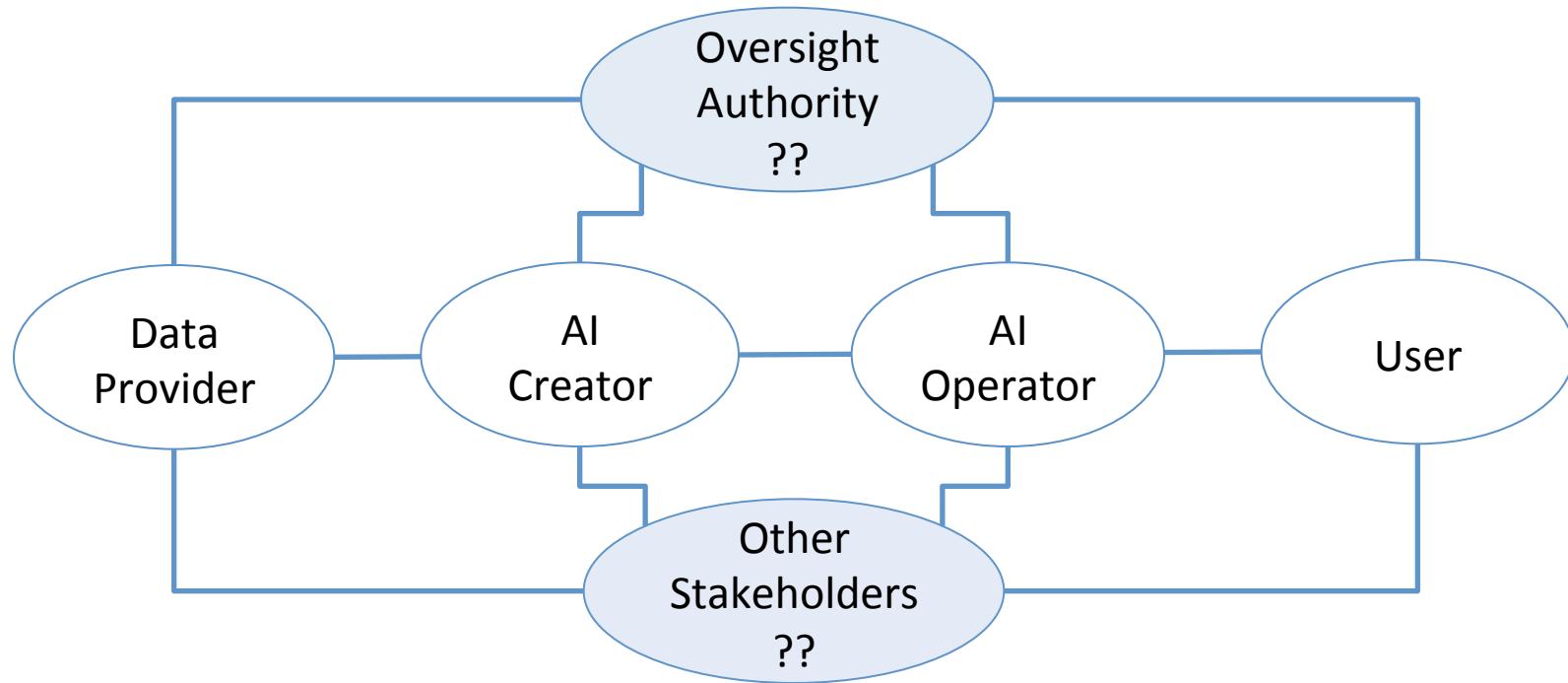
**ETHICAL
DATA AND
INFORMATION
MANAGEMENT**

CONCEPTS, TOOLS AND METHODS

KATHERINE O'KEEFE DARAGH O'BRIEN



AI Governance: Societal Challenges and Gaps



- Nature of Oversight Authority
- Identification of and Prioritising between affected stakeholder
- Collective impacts, e.g. on social cohesion
- Governmental use of AI
- Malevolent AI use



Challenges in Regulating AI

- **Definition:** Difficult to reach stable consensus on what defines AI
- **Discreetness:** Growing access to AI skills and computing power, it can be developed out of sight
- **Diffuseness:** AI used in a diffuse set of locations and jurisdictions
- **Discreteness:** Impact of an AI component only apparent when assembled into a system
- **Opacity:** Modern machine learning yields results without clear explanations
- **Forseeability:** AI-driven autonomous system can behave in unforseeable ways – ‘liability gap’
- **Control:** AI can work in ways out of control of those responsible for them
- **Pacing:** AI tech develops faster than societies ability to regulate it
- **Securitisation:** International competition as AI perceived as a strategic economic/military resource
- **Innovation:** Perceived impediment to AI-based innovation and its economic and social benefits
- **Asymmetry:** Power concentration in digital platforms benefiting from network effects

Role of Data in AI Governance



- Problem: for Data, **Possession** is 9.9 tenths of the Law
- Power of AI grows with the volume (and quality) of its training data
- Controlling the flow of Data is the Key to Governing AI
- Platforms emerging for maintaining possession Personal Data:
 - Inrupt
 - Hub-of-all-things



Role of Data in AI Governance

BUT

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

INSTEAD

- Problem: for Data, Possession is 9.9 tenths of **the Law**



Could a form of Data Ownership Help?

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

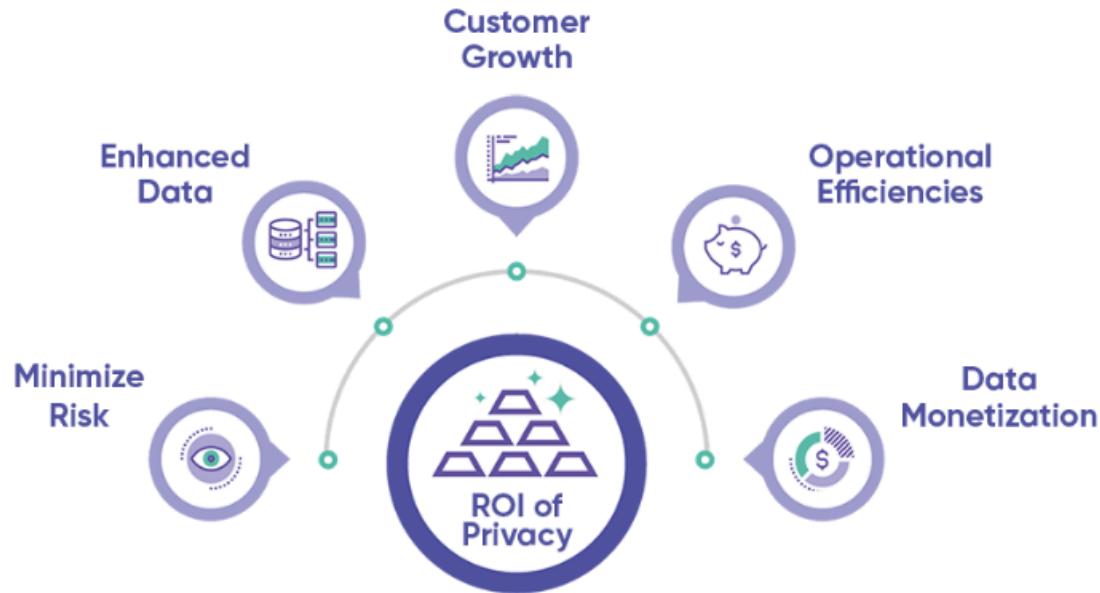
<https://www.twobirds.com/en/news/articles/2017/global/data-ownership-in-the-context-of-the-european-data-economy>

Organizations could transfer governance responsibility to more representative groups:

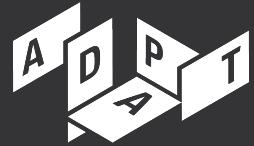
- Data Trusts
- Data Co-ops
- Data Unions – Data as Labour - <https://blog.singularitynet.io>

Example: Data Trusts

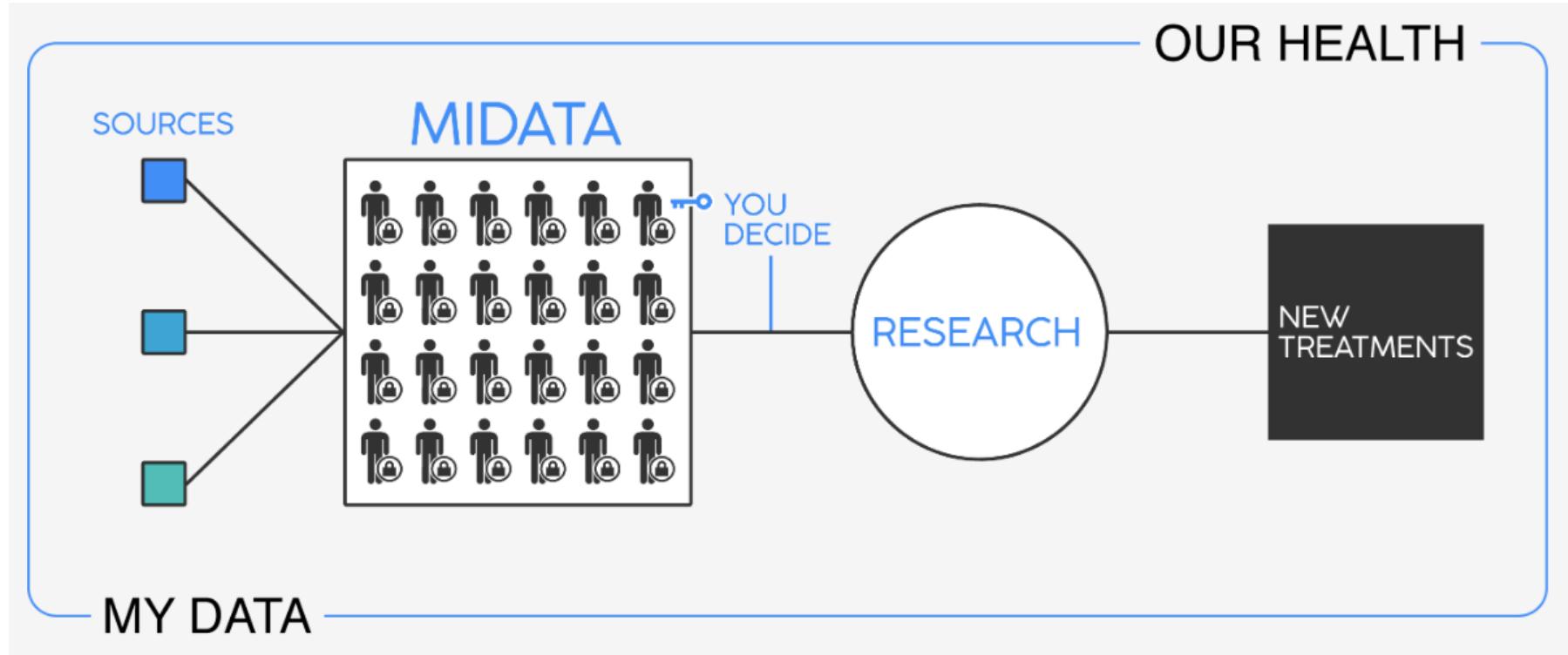
- Truata.com: Anonymised Data Analytics Services



- Offers large clients secure, anonymised analytics services of their own data
- Outsources data protection risks without losing all benefits of data analytics
- Part of business model is a data trust which is constituted separate to the business/profit driven part of the company
- Trust gives clients (and their customers) confidence that the rules can't change for business reasons



Example: MIDATA Medical Data Coop



<https://www.midata.coop>



Summary

- As Internet becomes more powerful and ubiquitous, risks of individual and societal impact and harm grows
- Becoming a priority for governments and companies for Digital Content, AI, Big Data, IoT
- Governments world-wide grappling with how to address these problems – regulation poses major challenges
- Data is Key but new forms of Governance and Oversight yet to emerge



Engaging Content
Engaging People

Thanks for your attention!

Any questions or feedback?

@ethicscanvas

ethicscanvas.org

hello@ethicscanvas.org

