ETHICAL DATA SCIENCE

GDPR, EU AI ACT, & RESPONSIBLE PRACTICE

Cecilia Dones :: June 2024

THIS TALK IS IDEAL FOR YOU IF YOU TOUCH ANY OF THESE AREAS...

As **Data Leader**, do you address:

Strategy & Culture: Conduct impact assessments to align data strategy with regulations, fostering responsible AI innovation.

Ethical Ecosystem: Champion resources & training for ethical AI development, fostering open discussions on bias.

Regulation Navigation: Proactively track evolving regulations and liaise with legal teams for long-term ethical & legal compliance.

As **Data Practitioner**, do you address:

Implementation & Advocacy: Integrate ethical considerations into coding & analysis, utilizing bias mitigation techniques.

Data Quality & Inclusivity: Advocate for high-quality, diverse data sets and responsible data collection practices.

Knowledge Gap Bridge: Communicate technical work clearly and advocate for responsible AI within the organization.

WHAT YOU WILL TAKE WITH YOU TODAY

Regulatory Updates

Understand status of:

- GDPR
- EU AI Act
- Current US Legislation

Best Practices

Established Tools for application:

- IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems
- Open Data Institute's "Ethical Data Science Maturity Model"

Continuous Learning

How To Stay Up-To-Date:

- Alan Turing Institute
- Montreal AI Ethics Institute
- All Tech Is Human

Thought Leaders

AGENDA

Introductions

- Cecilia Dones,
  3 Standard Deviations,
  Founder

Why?

Our Impact
- Industry
- Firm

Your Impact

What?

Current Regulations
- GDPR, CCPA
- EU AI Act

Philosophy

How?

Frameworks
- IEEE
- NIST
- ODI
- IAPP

Thought Leaders

Q & A

- Contact Me

BUT FIRST, A LITTLE BIT ABOUT YOU

1

How confident are you in navigating the AI Ethics space?

(1 – Not Confident 5 – Confident)

2

Who should be responsible for AI Ethics?

(1 – Government, 2 – Firm, 3 – Self)

3

What are you seeking from this talk?

CECILIA DONES :: SHORT INTRODUCTION

Founder, 3 Standard Deviations

CHASE

verizon

L'ORÉAL

reckitt

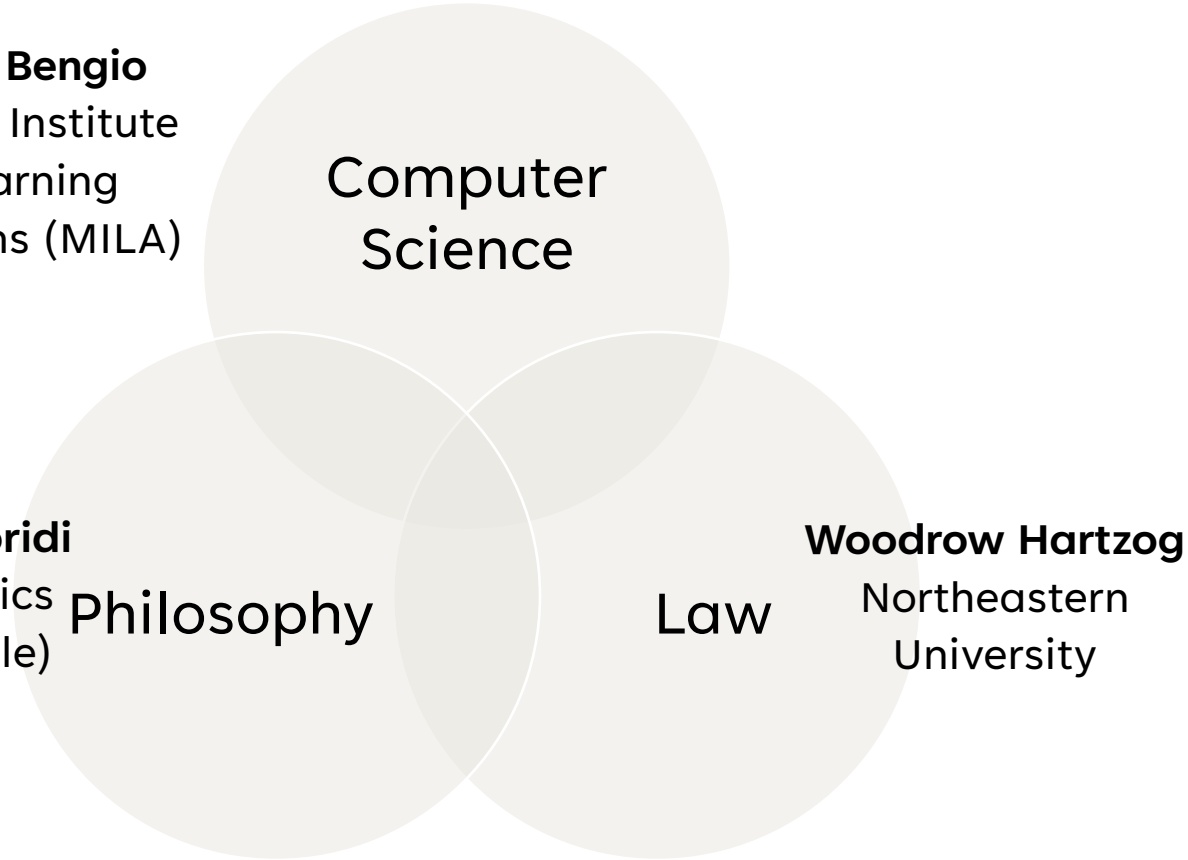
Moët Hennessy

Columbia Business School

Wharton University of Pennsylvania

MIT MANAGEMENT SLOAN SCHOOL

Yoshua Bengio
Montreal Institute for Learning Algorithms (MILA)



Luciano Floridi
Digital Ethics Center (Yale)

Woodrow Hartzog
Northeastern University

AGENDA

Introductions

- Cecilia Dones,
  3 Standard Deviations,
  Founder

Why?

Our Impact
- Industry
- Firms

Your Impact

What?

Current Regulations
- GDPR, CCPA
- EU AI Act

Philosophy

How?

Frameworks
- IEEE
- NIST
- ODI
- IAPP

Thought Leaders

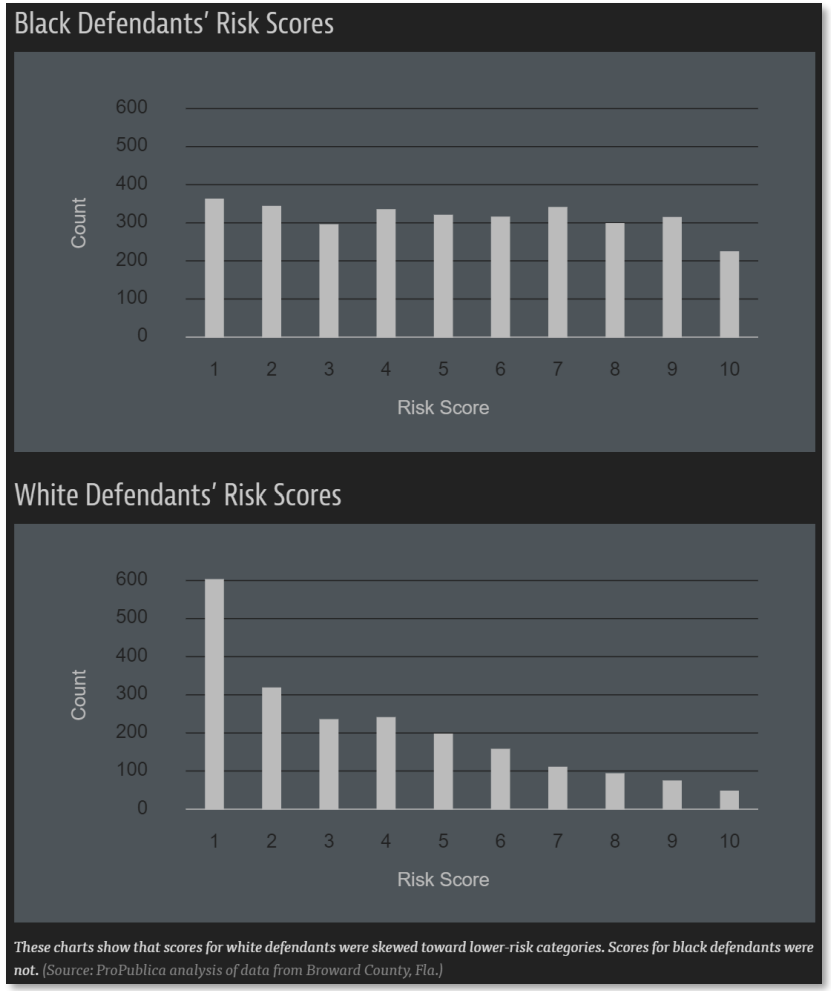
Q & A

- Contact Me

OUR INDUSTRY HAS IMPACT ON PEOPLE'S LIVES (1)

A 2016 ProPublica study found that a risk assessment tool used in criminal sentencing is more likely to misclassify black defendants as high risk.

- Black defendants were more likely to be incorrectly flagged as higher risk of committing future crimes compared to white defendants
- Only 20% of the time did the algorithm correctly predict violent recidivism
- Defendants have no way to verify or challenge the risk scores assigned to them
- Judges may place undue weight on the scores due to cognitive biases like automation bias.



These charts show that scores for white defendants were skewed toward lower-risk categories. Scores for black defendants were not. (Source: ProPublica analysis of data from Broward County, Fla.)

Source: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

OUR INDUSTRY HAS IMPACT ON PEOPLE'S LIVES (2)

RICE UNIVERSITY'S
BAKER INSTITUTE
FOR PUBLIC POLICY

Working Paper

Automation Does Not Kill Jobs; It Increases Inequality

Dagobert L. Brito, Ph.D.
Baker Institute Rice Faculty Scholar
Peterkin Professor of Political Economy Emeritus, Rice University

and

Robert F. Curl, Ph.D.
Baker Institute Rice Faculty Scholar
Pitzer–Schlumberger Professor of Natural Sciences Emeritus,
University Professor Emeritus and Professor of Chemistry Emeritus, Rice University

A 2020 Baker Institute study focused on the effects of automation on jobs.

It discusses two main effects:
- increased inequality and
- economic growth.

Automation lowers the cost of labor for some tasks, which increases the profit share of capital owners – widening the gap between the rich and the poor.

At the same time, automation spurs economic growth.

However, to benefit from this growth, workers need skills that cannot be automated.

Source: https://www.bakerinstitute.org/research/report-automation-does-not-kill-jobs-it-increases-inequality

OUR INDUSTRY HAS IMPACT ON PEOPLE'S LIVES (3)

Eli Pariser's TED Talk "Beware online 'filter bubbles'" explores the concerning phenomenon of personalized filters on the internet creating "filter bubbles" that can isolate people and limit their exposure to diverse perspectives.

Potential Societal Harms:
- Filter bubbles can isolate people and prevent exposure to different viewpoints that challenge or broaden their worldviews.
- This limits the internet's potential to connect diverse people and ideas.
- It can be detrimental for individuals and democracy by amplifying societal divisions and polarization.

Source: https://www.youtube.com/watch?v=B8ofWFx525s

YOUR IMPACT IN 3 KEY AREAS

Foundational AI Ethics

Focus:
Broad principles guiding ethical AI development.

Actionable Steps:
Learning frameworks, identifying ethical concerns.

Outcome:
Building ethical considerations into the AI development process.

Bias in AI

Focus:
Identifying and mitigating biases within AI systems.

Actionable Steps:
Detecting biases, applying debiasing techniques.

Outcome:
Developing fairer and more trustworthy AI systems.

Explainable AI (XAI)

Focus:
Making AI models more interpretable.

Actionable Steps:
Choosing and applying XAI techniques for specific models.

Outcome:
Improving transparency and understanding of AI decisions.

Learn More: 3 Standard Deviations

AGENDA

Introductions	Why?	What?	How?	Q & A
• Cecilia Dones, 3 Standard Deviations, Founder	Our Impact • Industry • Firms Your Impact	Current Regulations • GDPR, CCPA • EU AI Act Philosophy	Frameworks • IEEE • NIST • ODI • IAPP Thought Leaders	• Contact Me

A REGULATORY PERSPECTIVE: CIVIL LAW VS. COMMON LAW

EU **US**

	Civil law	Common law
The basic law-making fact	Codification of regulations	Decisions of judges
The basic source of law	Legislative enactments	Legal precedents and legislation
The entity that creates the law	Legislative body	Legislative body and judges
Division of law	Private and public law	Common law and equity
Origin of law	Roman law	Ius commune

Source: Pelc, Justyna & Torchała, Piotr & Suścicka, Beata & Łabowska, Magdalena & Popiel, Małgorzata. (2020). MARS SOCIETY - HOW SHOULD POLICY AND ORGANIZATION IN MARS COLONY LOOK LIKE?.

CONSUMER EXPECTATIONS: PRIVACY VS. CONFIDENTIALITY

People

It is about the individual's rights regarding their information.

Data

It is about the individual's data and how it is protected from others.

DATA PROTECTION LAW ESSENTIALS

	GDPR	CCPA	CPRA	VCDPA	CPA
Q1. What are the basics?					
Name	General Data Protection Regulation	California Consumer Privacy Act	California Privacy Rights Act	Consumer Data Protection Act	Colorado Privacy Act
Citation	EU/2016/679	Cal. Civ. Code § 1798.100 et seq.	Cal. Civ. Code § 1798.100 et seq.	Va. Code § 59.1-571 et seq.	Colo. Rev. Stat. § 6-1-1301 et seq.
Jurisdiction	European Union	California	California	Virginia	Colorado
Model	Opt-in	Opt-out	Opt-out	Opt-out	Opt-out
Sector	Non-sectoral	Non-sectoral	Non-sectoral	Non-sectoral	Non-sectoral
Effective date(s)	May 25, 2018	Jan. l, 2020	Dec. 16, 2020; Jan. l, 2023	Jan. l, 2023	Jul. l, 2023
Q2. Whose data is protected?					
Statutory term	Data subject	Consumer	Consumer	Consumer	Consumer
Defined as	Natural person in the EU	Natural person who is a CA resident	Natural person who is a CA resident	Natural person who is a VA resident	Individual who is a CO resident

LEARN MORE: Bloomberg Law; https://assets.bbhub.io/bna/sites/7/2021/12/Comparison-Table-Privacy-Law-FAQs.pdf

LAW ESSENTIALS: AI DEVELOPER

	GDPR	CCPA	CPRA	VCDPA	CPA
Q3. What types of data are protected?					
Statutory term	Personal data	Personal information	Personal information	Personal data	Personal data
Defined as	Any information relating to an identified or identifiable natural person	Information that identifies, relates to, describes, is reasonably capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular consumer or household	Information that identifies, relates to, describes, is reasonably capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular consumer or household	Any information that is linked or reasonably linkable to an identified or identifiable natural person	Information that is linked or reasonably linkable to an identified or identifiable individual
Definition excludes de-identified data	GDPR uses the term "pseudonymized," rather than "de-identified." According to Recital 26, personal data that has undergone pseudonymization-which could be attributed to a natural person by the use of additional information-should be considered personal data	Yes, but see provisions regarding reidentification of deidentified information. Cal. Civ. Code §1798.148	Yes, but see provisions regarding reidentification of deidentified information. Cal. Civ. Code §1798.148. Moreover, the CPRA authorizes the attorney general to update the definition of "deidentifed." Cal. Civ. Code §l798.l85(a)	Yes, but special requirements apply to de-identified data. See Va. Code§ 59.1-581	Yes, but special requirements apply to de-identified data. See Colo. Rev. Stat.§ 6-1-1307
Definition excludes publicly available info	No	Yes	Yes	Yes	Yes
Definition excludes aggregate info	Not specified, but Recital 162 indicates that the GDPR applies to the processing of personal data for statistical purposes	Yes	Yes	Not specified	Not specified

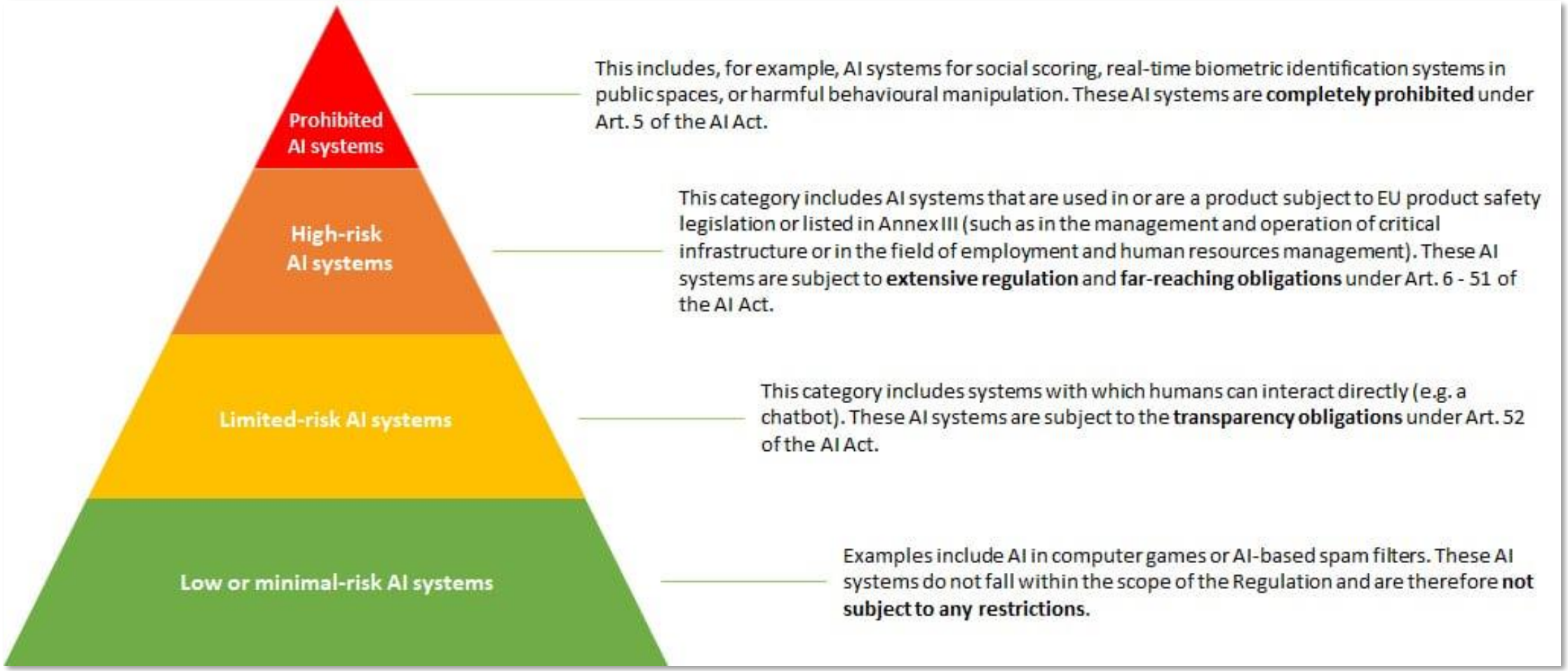
LEARN MORE: Bloomberg Law; https://assets.bbhub.io/bna/sites/7/2021/12/Comparison-Table-Privacy-Law-FAQs.pdf

EU AI ACT OVERVIEW

- “Safeguards agreed on general purpose artificial intelligence
- Limitation for the of use biometric identification systems by law enforcement
- **Bans on social scoring and AI used to manipulate or exploit user vulnerabilities**
- Right of consumers to launch complaints and receive meaningful explanations
- Fines ranging from 35 million euro or 7% of global turnover to 7.5 million or 1.5% of turnover”

Source: https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai;

EU AI ACT RISK LEVELS



Source: https://www2.deloitte.com/dl/en/pages/legal/articles/ki-verordnung-eu.html

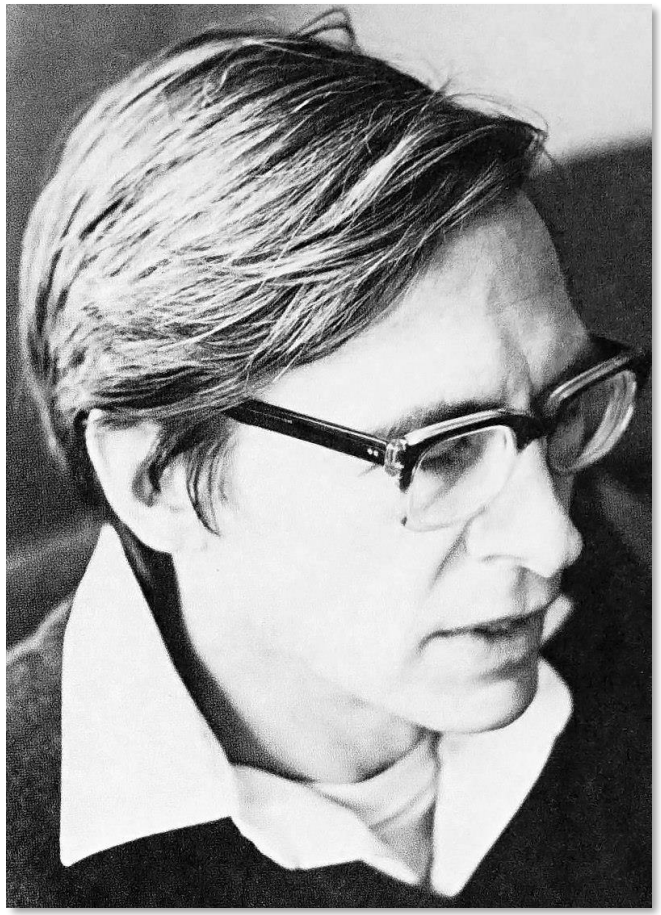
A PHILOSOPHICAL PERSPECTIVE: GUIDING PRINCIPLES

John Rawls' Justice: Just society = fair social contract designed in ignorance of your own advantages/disadvantages (veil of ignorance).

Original Position: Imagine designing society from scratch, blind to who you'll be (race, wealth, etc.).

Veil of Ignorance: Ensures fairness by removing self-interest based on unknown position.

Two Principles of Justice:
•Everyone gets basic liberties (as long as they don't limit others').
•Social/economic advantages should help the worst off the most.



Source: https://plato.stanford.edu/entries/rawls/

AGENDA

Introductions

- Cecilia Dones,
  3 Standard Deviations,
  Founder

Why?

Our Impact
- Industry
- Firms

Your Impact

What?

Current Regulations
- GDPR, CCPA
- EU AI Act

Philosophy

How?

Frameworks
- IEEE
- NIST
- ODI
- IAPP

Thought Leaders

Q & A

- Contact Me

DATA SCIENTISTS ETHICAL TOOLKIT: IEEE

Excerpt From Executive Summary The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems

Eudaimonia, as elucidated by Aristotle, is a practice that defines human well-being as the highest virtue for a society. Translated roughly as "flourishing," the benefits of eudaimonia begin by conscious contemplation, where ethical considerations help us define how we wish to live.

Whether our ethical practices are Western (Aristotelian, Kantian), Eastern (Shinto, Confucian), African (Ubuntu), or from a different tradition, by creating autonomous and intelligent systems that explicitly honor inalienable human rights and the beneficial values of their users, we can prioritize the increase of human well-being as our metric for progress in the algorithmic age. Measuring and honoring the potential of holistic economic prosperity should become more important than pursuing one-dimensional goals like productivity increase or GDP growth.

Source: https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_executive_summary_v2.pdf

DATA SCIENTISTS ETHICAL TOOLKIT: NIST

Excerpt From Artificial Intelligence Risk Management Framework (AI RMF 1.0)

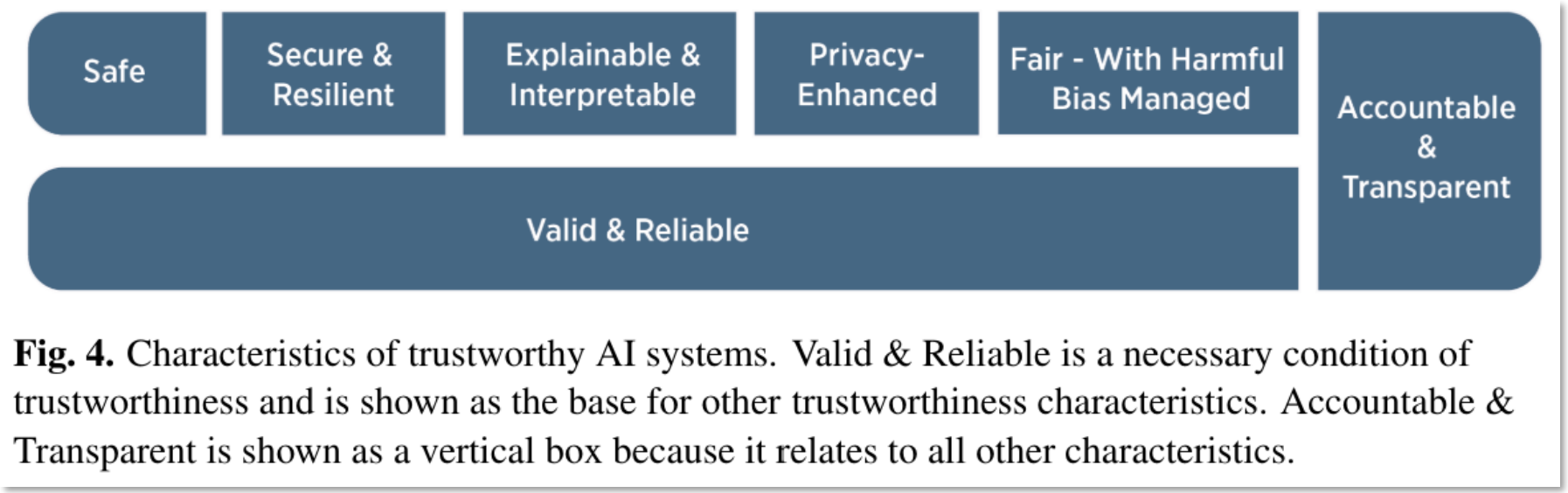


Fig. 4. Characteristics of trustworthy AI systems. Valid & Reliable is a necessary condition of trustworthiness and is shown as the base for other trustworthiness characteristics. Accountable & Transparent is shown as a vertical box because it relates to all other characteristics.

Source: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf

DATA SCIENTISTS ETHICAL TOOLKIT: ODI

Excerpt From Ethical Data Science Maturity Model Workbook

Levels of maturity

The model builds up across five levels of maturity:

1. **Initial –** the desirable processes are non-existent or ad hoc, with no organisational oversight.
2. **Repeatable –** processes are becoming refined and repeatable, but only within the scope of individual teams or projects. There are no organisational standards.
3. **Defined –** processes are standardised within the organisation based on best practices identified internally or from external sources. Knowledge and best practices start to be shared internally. However the processes may still not be widely adopted.
4. **Managed –** the organisation has widely adopted the standard processes and begins to monitor them using defined metrics.
5. **Optimising –** the organisation is attempting to optimise and refine its process to increase efficiency within the organisation and, more widely, within its business sector.

Source: https://theodi.org/insights/tools/

DATA SCIENTISTS ETHICAL TOOLKIT: IAPP

Excerpt From The EU AI Act: Where to start

The EU AI Act countdown has begun. What should you do now?

1. **Consider the big picture.** Join the "EU AI Act: A Major Moment in the Digital World" LinkedIn Live and check out the "Global AI Law and Policy Tracker."
2. **Learn what happens next and outline compliance deadlines.** Check out the "Key Dates for EU AI Act Implementation" infographic.
3. **Review key components and brief your executives.** Review the EU AI Act cheat sheet and the EU AI Act 101 infographic.
4. **Understand how others are preparing.** Join the "Preparing to Implement the EU AI Act" LinkedIn Live.
5. **Gain knowledge to build a compliance plan.** Attend the EU AI Act Comprehensive at the IAPP Global Privacy Summit 2024.
6. **Upskill your team.** Register for AI Governance Professional training.
7. **Validate your readiness, knowledge and skill.** Get AIGP certified.
8. **Hear from enforcers and gauge your risk profile.** Join the "The EU AI Act: A View From the Lawmaker on Next Steps" LinkedIn Live.
9. **Benchmark your AI governance plans against peers:** Attend AI Governance Global 2024 and check out the "IAPP-EY Professionalizing Organizational AI Governance Report."
10. **Stay up to date.** Sign up for the AI Governance Dashboard for the latest on the EU AI Act and more.

Source: https://iapp.org/about/ai-governance-center/

LET'S REVISIT YOUR IMPACT IN 3 KEY AREAS

Foundational AI Ethics

Focus:
Broad principles guiding ethical AI development.

IEEE (Institute of Electrical and Electronics Engineers)

Additional Resources:
- EU Ethics Guidelines for Trustworthy AI
- Montréal Declaration for a Responsible Development of Artificial Intelligence

Bias in AI

Focus:
Identifying and mitigating biases within AI systems.

NIST (National Institute of Standards and Technology)
IAPP (International Association of Privacy Professionals)

Additional Resources:
- AI NOW Institute
- Data & Society Research Institute

Explainable AI (XAI)

Focus:
Making AI models more interpretable.

ODI (Open Data Institute)

Additional Resources:
- Local Interpretable Model-agnostic Explanations (LIME)
- **SH**apley **A**dditive ex**P**lanations

HOW YOU CAN STAY UP TO DATE

Institutes To Follow

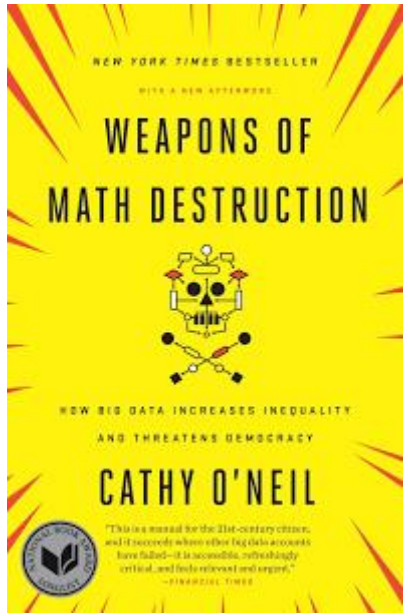
The Alan Turing Institute

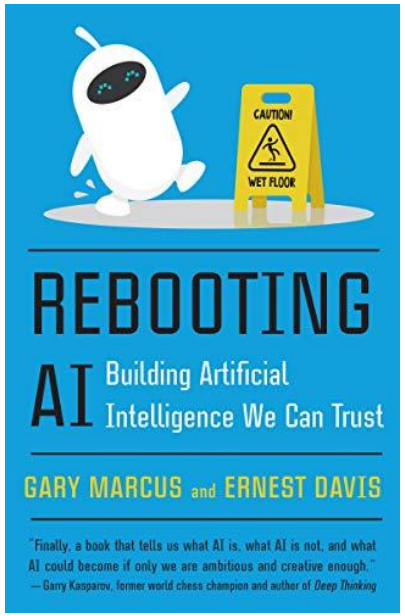
HAI
Stanford University
Human-Centered
Artificial Intelligence

MAIEI

all tech is human

People To Know





Learn More: 3 Standard Deviations

LEARN MORE: RESOURCES

Algorithmic Accountability: A Primer
https://datasociety.net/wp-content/uploads/2018/04/Data_Society_Algorithmic_Accountability_Primer_FINAL-4.pdf

THE FUTURE OF EMPLOYMENT: HOW SUSCEPTIBLE ARE JOBS TO COMPUTERISATION?
https://oms-www.files.svdcdn.com/production/downloads/academic/The_Future_of_Employment.pdf

Data Science Association's Code of Conduct
https://datascienceassn.org/code-of-conduct.html

Aligning AI Systems to Ideas of Justice
https://pubmed.ncbi.nlm.nih.gov/37094137/

Participating in the IEEE
https://standards.ieee.org/participate/

NIST Crosswalks
https://airc.nist.gov/AI_RMF_Knowledge_Base/Crosswalks

REFLECTION

1

How confident are you in navigating the AI Ethics space?

(1 – Not Confident 5 – Confident)

2

Who should be responsible for AI Ethics?

(1 – Government, 2 – Firm, 3 – Self)

3

What are you taking away from this talk?

AGENDA

Introductions

- Cecilia Dones,
  3 Standard Deviations,
  Founder

Why?

Our Impact
- Industry
- Firms

Your Impact

What?

Current Regulations
- GDPR, CCPA
- EU AI Act

Philosophy

How?

Frameworks
- IEEE
- NIST
- ODI
- IAPP

Thought Leaders

Q & A

- Contact Me

THANK YOU

IF YOU WOULD LIKE TO COLLABORATE WITH ME, PLEASE FEEL FREE TO FIND ME ON LINKEDIN OR EMAIL ME:

CECILIA.DONES@3STANDARDDEVIATIONS.COM