

# AI & Privacy Workshop

## Privacy + Security Forum

---

May 6, 2020

# Agenda

- 2:00-2:25 AI Intro/Background, Promise, & Policy Concerns  
*Fatima Khan · Sr. Corporate Counsel, Privacy & Product · Okta*
- 2:25-2:40 AI Laws, Regulations, & Guides  
*Susan Hintze · Managing Partner · Hintze Law PLLC*
- 2:40-3:00 AI Privacy Frameworks, Guidelines, & Best Practices  
*Aaron Weller · VP Strategy · Sentinel*
- 3:00-3:30 Break – Submit Questions
- 3:30-3:40 Answer Questions
- 3:40-4:30 Hypothetical · *Fatima, Susan, Aaron*

## **AI & Privacy Workshop**

Privacy & Security Forum

Fatima Khan

Sr. Corporate Counsel, Privacy & Product  
Okta

# The Promise of AI



Possibility to positively impact every industry – healthcare, autonomous driving, fraud detection, security, commerce, etc.

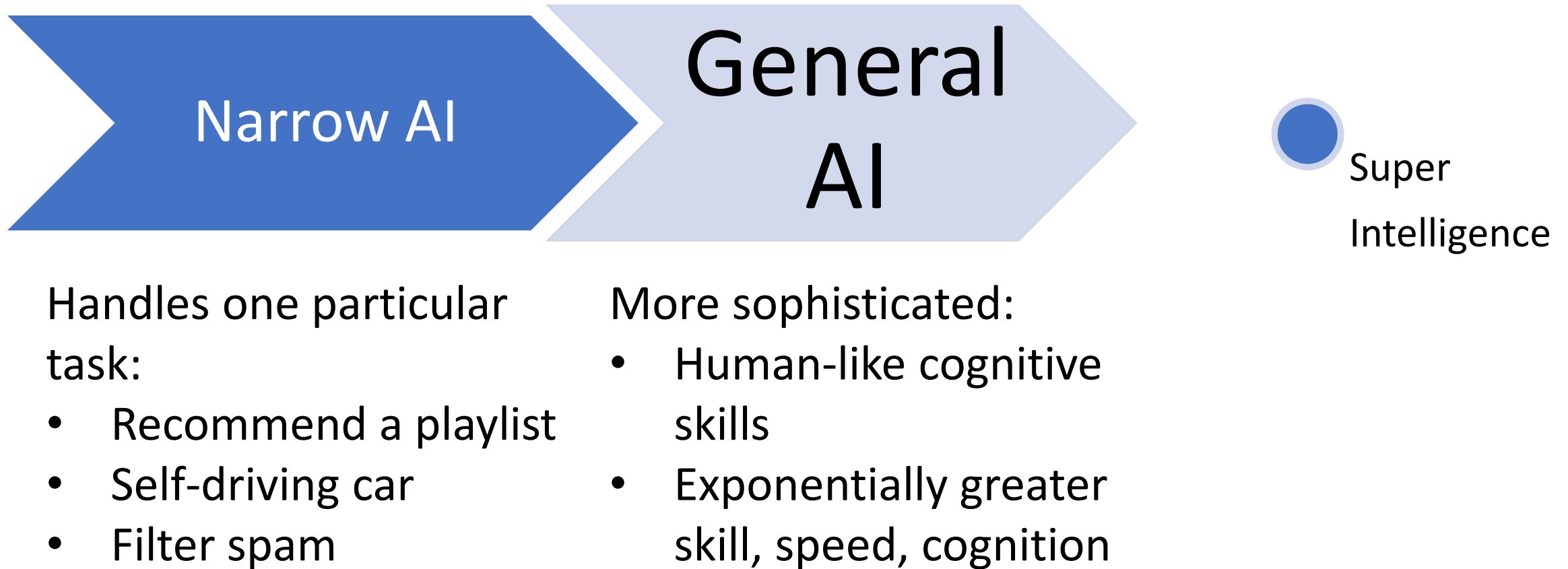
# Hey Siri, What is AI?

a branch of computer science dealing with the simulation of intelligent behavior in computers.

Methods for non-human systems to learn

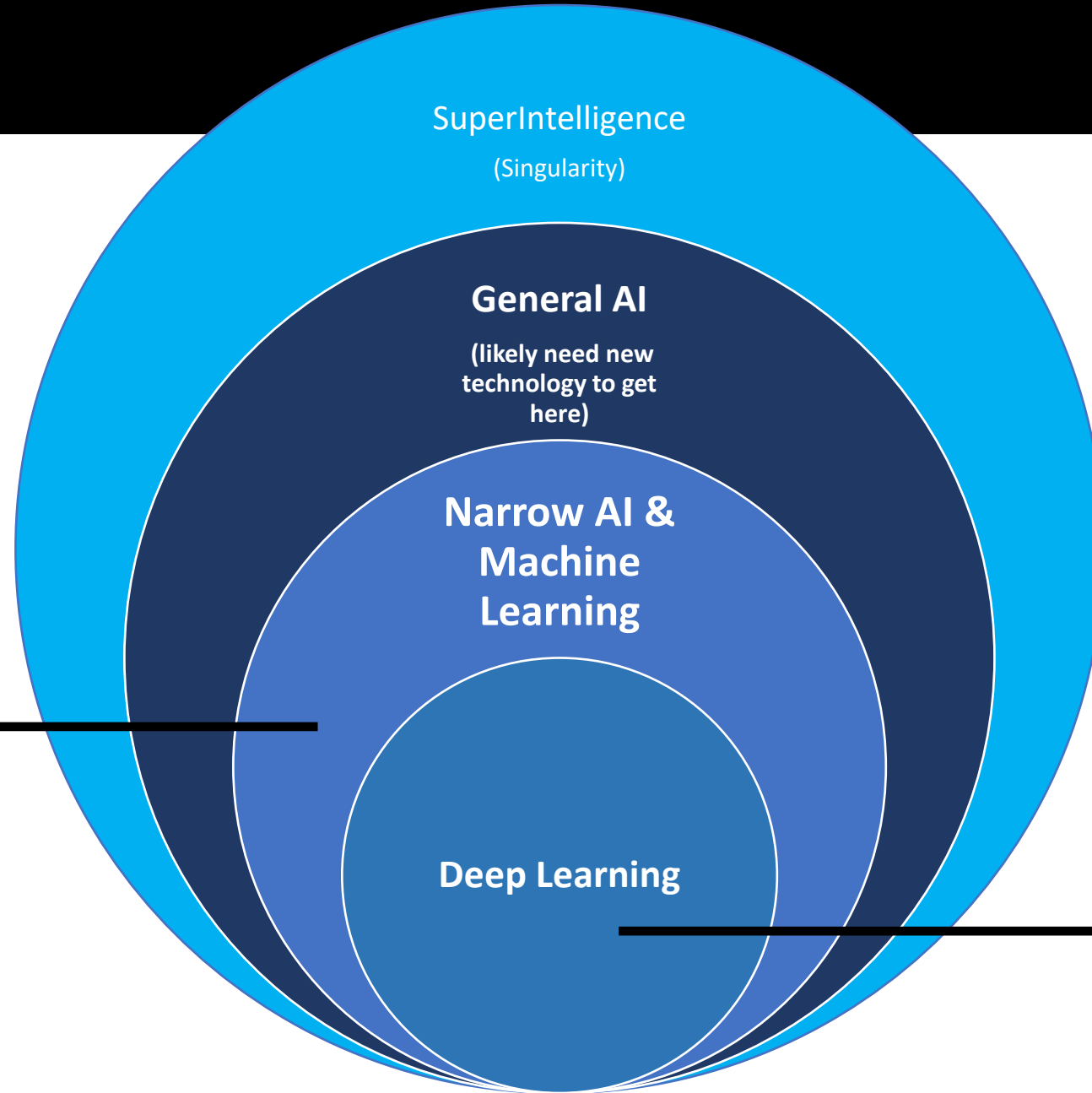
theory and development of computer systems able to perform tasks normally requiring human intelligence

# General vs. Narrow



# Core Concepts

Use of algorithms to parse training data and make predictions without being explicitly programmed



Class of ML algorithms that uses multiple layers to learn based on artificial neural networks

# Key Components

## Computing Power

- GPUs, Processors optimized for ML, Quantum computing

## Large amounts of data (for training)

## Algorithms

- ML/Neural Networks Focused, Tools like TensorFlow, PyTorch, etc.



**If AI-assisted decision-making**

Human  
decision-making

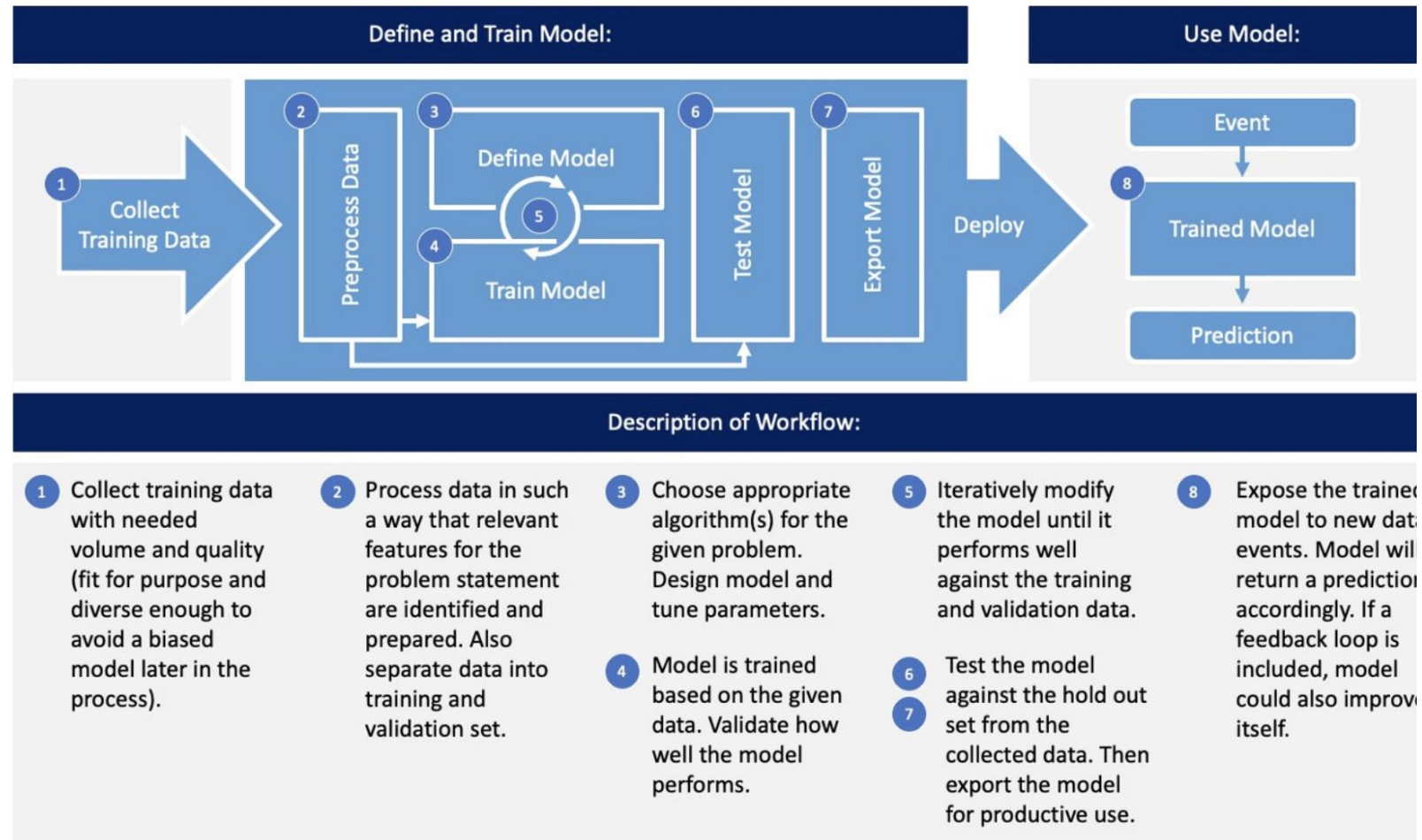
Additional data



**Outputs: ex.  
prediction,  
classification, or  
recommendation**



# Machine Learning - Supervised

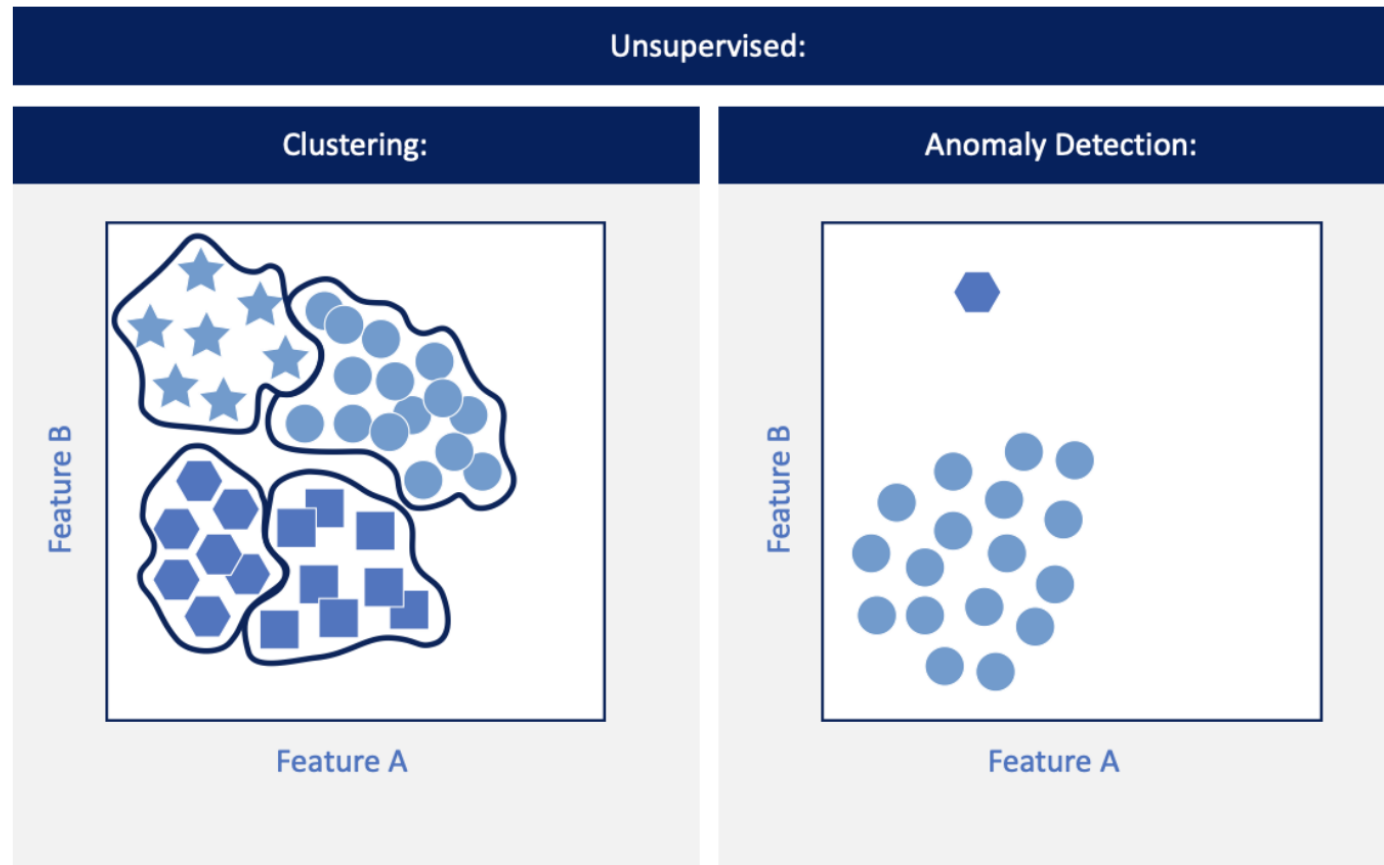


"Machine learning workflow (supervised)" by Nils Ackermann is licensed under Creative Commons [CC BY](#)

[ND 4.0](#)

# Machine Learning - Unsupervised

- Too difficult/expensive to obtain enough labeled training data.
  - Algorithm finds a way to classify the data on its own.



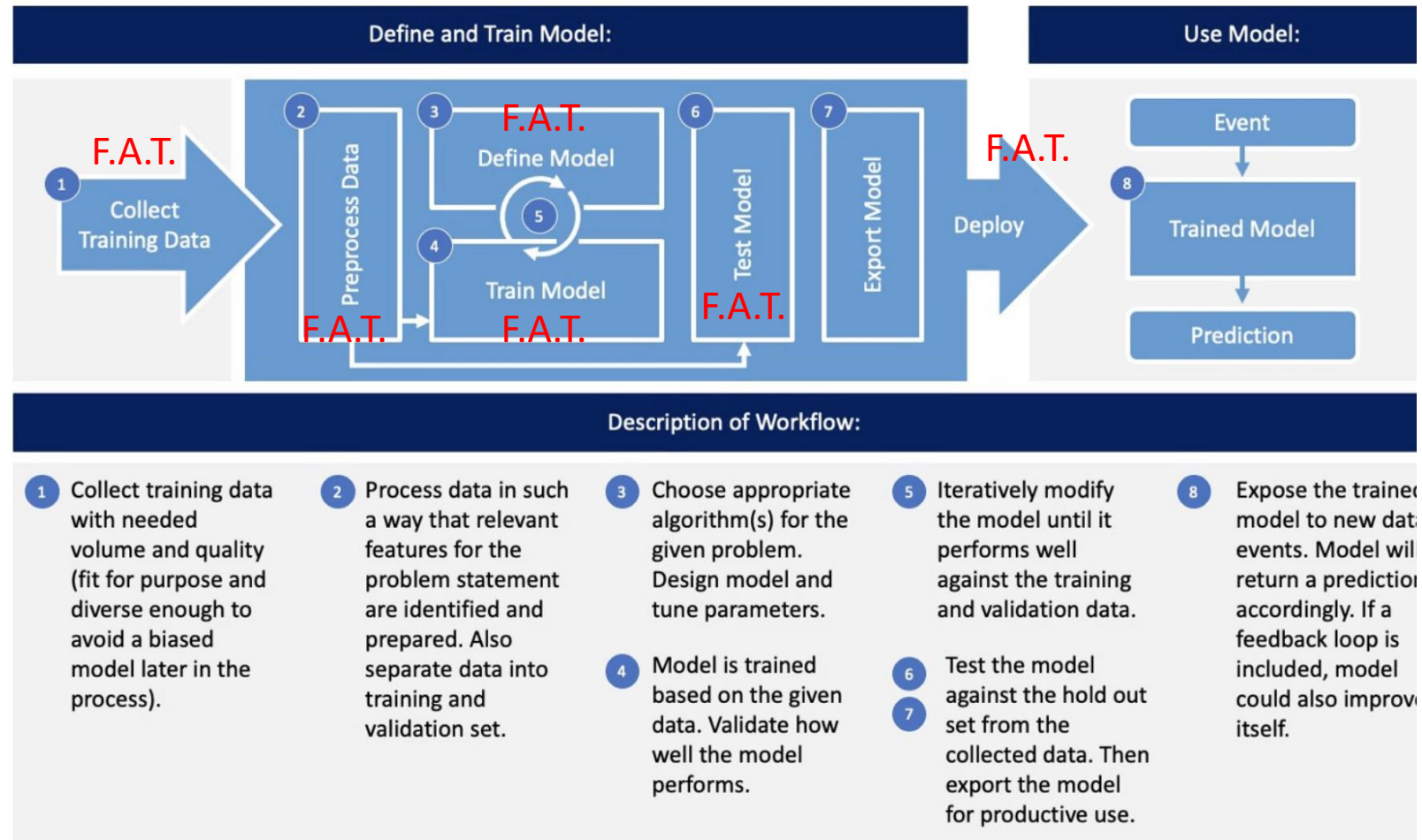
# FAT Framework

Framework in line with privacy principles

Fairness	Accountability	Transparency
<ul style="list-style-type: none"><li>• Fair &amp; lawful processing</li><li>• Lawful bases</li><li>• Mitigating discrimination</li><li>• <b>Obligation to conduct a DPIA</b></li></ul>	<ul style="list-style-type: none"><li>• Report, explain, and justify decision-making to impacted individuals</li><li>• Held accountable</li><li>• <b>Design</b> for accountability</li><li>• <b>Right of access in data protection law (including logic)</b></li><li>• <b>Right to object in data protection law</b></li><li>• <b>Requirement to adopt suitable safeguards (such as human intervention) and ability to contest</b></li></ul>	<ul style="list-style-type: none"><li>• Provide transparency in <b>design</b>, decision-making, and output</li><li>• <b>Right to be informed in data protection law, especially for process without human involvement for legal/significant effects</b></li><li>• <b>Right not to be subject to automated decision-making under data protection law</b></li></ul>

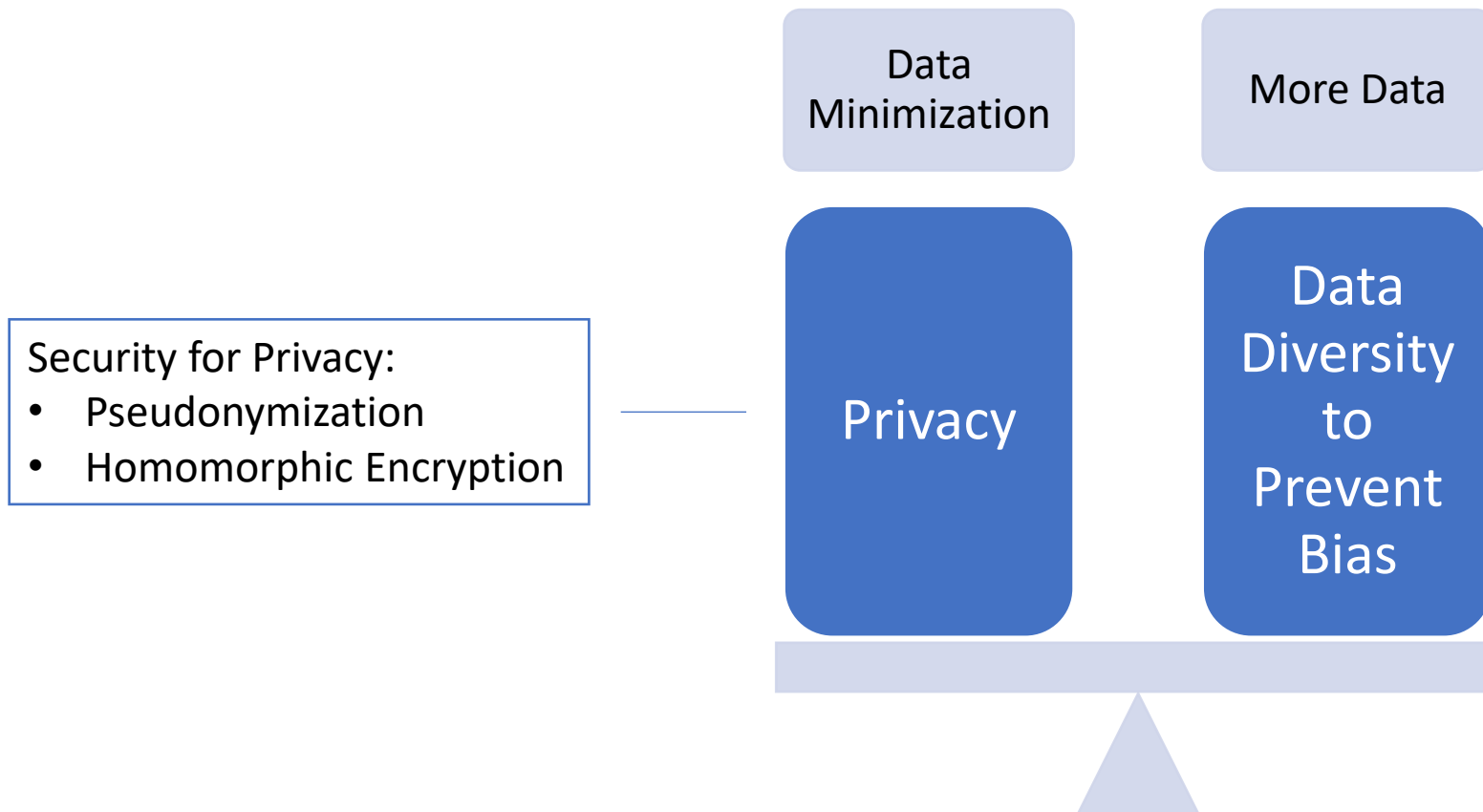
Avoid justification for decision-making to “the algorithm made me do it...”

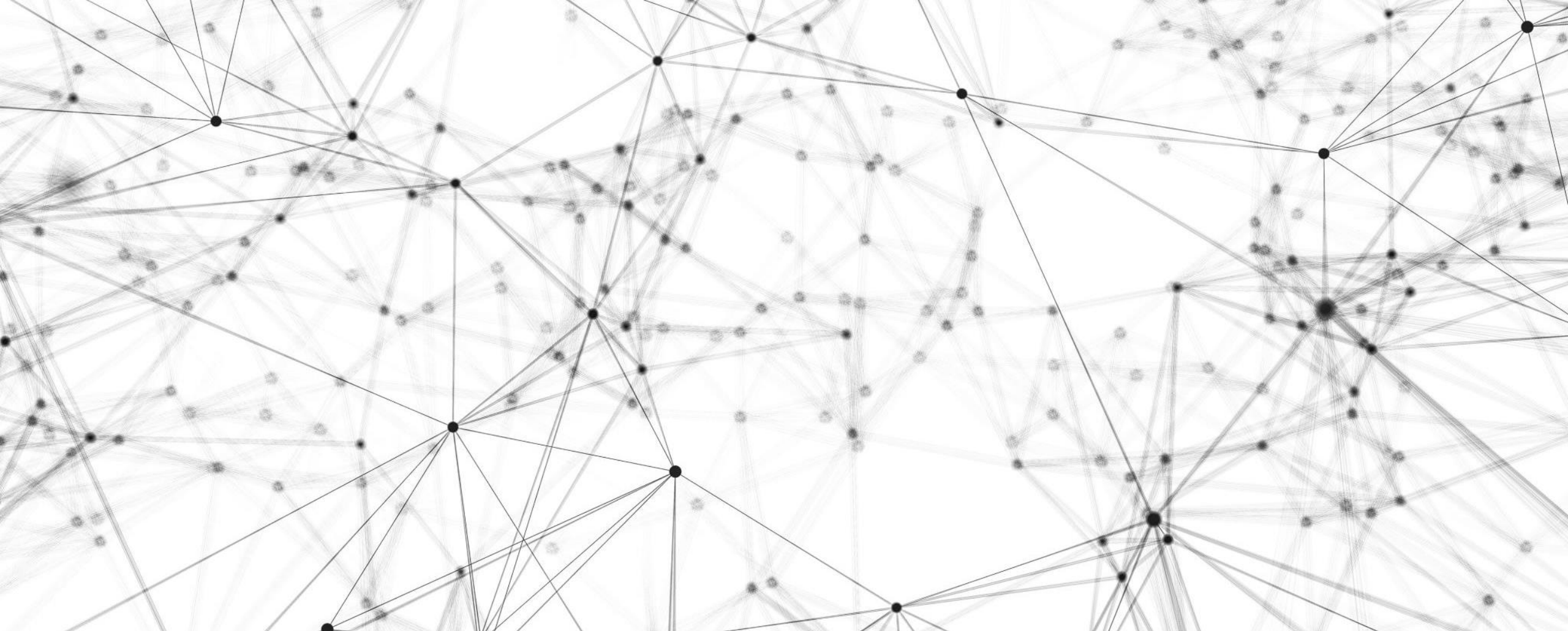
# Machine Learning - Supervised



"Machine learning workflow (supervised)" by Nils Ackermann is licensed under Creative Commons [CC BY-ND 4.0](https://creativecommons.org/licenses/by-nd/4.0/)

# Data Minimization





# Privacy Laws and AI

SUSAN L. HINTZE

MANAGING PARTNER (SHE/HER/HERS)

HINTZE LAW PLLC

@SLHINTZE



# FTC Act

FTC v. Ashley Madison – Alleged fake engager profiles of attractive mates used to deceive customers. <https://www.ftc.gov/news-events/press-releases/2016/12/operators-ashleymadisoncom-settle-ftc-state-charges-resulting>

FTC v. Facebook – Alleged collecting images for facial recognition algorithm

<https://www.ftc.gov/news-events/press-releases/2019/07/ftc-imposes-5-billion-penalty-sweeping-new-privacy-restrictions>

<https://www.ftc.gov/news-events/blogs/business-blog/2020/04/using-artificial-intelligence-algorithms>

EPIC complaint against HireVue

[https://www.epic.org/privacy/ftc/hirevue/EPIC\\_FTC\\_HireVue\\_Complaint.pdf](https://www.epic.org/privacy/ftc/hirevue/EPIC_FTC_HireVue_Complaint.pdf)

Big Data: A Tool for Inclusion or Exclusion? Understanding the Issues (FTC Report)

<https://www.ftc.gov/reports/big-data-tool-inclusion-or-exclusion-understanding-issues-ftc-report>

# FCRA / ECOA

Fair Credit Reporting Act (FCRA) & Equal Credit Opportunity Act (ECOA) both address automated decision-making and adverse action notices

## **ECOA**

**"Discrimination"** - Prohibits discrimination against credit applicants on the basis of race, color, religion, national origin, sex, marital status, age, receipt of public assistance, or good faith exercise of any rights under the Consumer Credit Protection Act.

**"Adverse action notice"** - Requires creditors to provide applicants, upon request, with the reasons underlying decisions to deny credit.

## **FCRA**

**"Consumer reporting agency"** - automated decision-making about eligibility for credit, employment, insurance, housing, etc.

**"Adverse action notice"** - tells the consumer about their right to see the information reported about them and to correct inaccurate information



# HIPAA

---



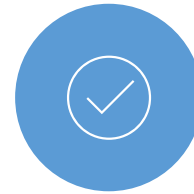
PROTECTED  
HEALTH  
INFORMATION



DEIDENTIFIED  
HEALTH  
INFORMATION



ACCESS



AMENDMENT



USE



DELETION

## US Guidelines for Ethical AI - HR 153

Transparent and explainable AI  
systems

Information privacy and personal  
data protection

Access and fairness in  
technological services and benefits

Accountability and oversight for  
automated decision-making

# Illinois - Artificial Intelligence Video Interview Act

## Artificial Intelligence Video Interview Act (AIVIA)

- Went into effect January 1, 2020,
- Places requirements on companies to:
  - provide notice,
  - obtain consent,
  - maintain confidentiality/limit distribution only those persons "whose expertise or technology" is necessary to evaluate an applicant's fitness for the position,
  - duty to destroy data, and
  - provide an explanation of how the AI works and general types of characteristics the technology uses to evaluate applicants.

# Washington – (new) Facial Recognition Law

Facial Recognition-- State and Local Government SB6280 – takes effect July 1, 2021

- Covers facial recognition services used by state or local agencies
- Requires:
  - Accountability report
  - Meaningful human review of decisions
  - Independent testing for accuracy and unfair performance differences across distinct subpopulations based on race, skin tone, ethnicity, gender, age, or disability status

# CALIFORNIA - CCPA

---



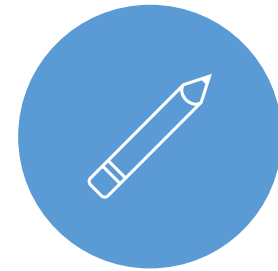
DEFINITION OF  
PERSONAL  
INFORMATION



DEIDENTIFICATION



ACCESS



DELETION

# Other state laws

Constitutional Rights of Privacy

Consumer Fraud and Deceptive Business  
Practices Acts

Breach of Contract

Intrusion on Seclusion

Unjust Enrichment

State Biometrics Laws – ex. IL, WA, TX,

# EU - GDPR

---



DEFINITION OF  
PERSONAL DATA



PSEUDONYMIZE/  
ANONYMIZE



ACCESS



RECTIFICATION



DELETION



AUTOMATED  
DECISION-MAKING

# EU Papers

---

Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679

---

[http://ec.europa.eu/newsroom/article29/document.cfm?doc\\_id=49826](http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826)

---

White Paper on Artificial Intelligence A European approach to excellence and trust

---

[https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020\\_en.pdf](https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf)

---

DECLARATION ON ETHICS AND DATA PROTECTION IN ARTIFICIAL INTELLIGENCE 40th International Conference of Data Protection and Privacy Commissioners Tuesday 23rd October 2018, Brussels

---

[https://edps.europa.eu/sites/edp/files/publication/icdppc-40th\\_ai-declaration\\_adopted\\_en\\_0.pdf](https://edps.europa.eu/sites/edp/files/publication/icdppc-40th_ai-declaration_adopted_en_0.pdf)

---

STUDY ON THE HUMAN RIGHTS DIMENSIONS OF AUTOMATED DATA PROCESSING TECHNIQUES (IN PARTICULAR ALGORITHMS) AND POSSIBLE REGULATORY IMPLICATIONS

---

<https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5>



# EU Member State Guidance & Actions (& Canada)

## UK - ICO

Big data, artificial intelligence, machine learning and data protection

<https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf>

AI Auditing Framework

<https://ico.org.uk/about-the-ico/news-and-events/ai-auditing-framework/>

## France

Algorithms and artificial intelligence: CNIL's report on the ethical issues

<https://www.cnil.fr/en/algorithms-and-artificial-intelligence-cnils-report-ethical-issues>

## Spain

RGPD compliance of processings that embed Artificial Intelligence An introduction

[https://www.aepd.es/sites/default/files/2020-02/adecuacion-rgpd-ia-en\\_0.pdf](https://www.aepd.es/sites/default/files/2020-02/adecuacion-rgpd-ia-en_0.pdf)

## Germany/Hamburg

Google Home Speech Assistant - auditing of AI voice assistant [https://datenschutz-hamburg.de/assets/pdf/2019-08-01\\_press-release-Google\\_Assistant.pdf](https://datenschutz-hamburg.de/assets/pdf/2019-08-01_press-release-Google_Assistant.pdf)

## Sweden

DPA launches investigation of governmental use of Clearview AI -

<https://www.datainspektionen.se/nyheter/datainspektionen-inleder-tillsyn-med-anledning-av-clearview-ai/>

## Canada

Consultation on the OPC's Proposals for ensuring appropriate regulation of artificial intelligence

[https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/consultation-ai/pos\\_ai\\_2020017](https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/consultation-ai/pos_ai_2020017)

Other

**OECD - Recommendation of the Council  
on Artificial Intelligence -**

<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>



# Susan Lyon-Hintze

---

Managing Partner

Hintze Law PLLC

[www.hintzelaw.com](http://www.hintzelaw.com)

[susan@hintzelaw.com](mailto:susan@hintzelaw.com)

@slhintze



# AI & Privacy Best Practices

Privacy + Security Forum, May 2020

*Aaron Weller*

*VP Strategy, Sentinel*





*AI is everywhere...from  
the benign to the deadly.*



# Comparison of AI Best Practice Principles



# Legal and Regulatory Drivers



## **GDPR Article 35: Data protection impact assessment**

Where a type of processing **in particular using new technologies**, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an **assessment of the impact of the envisaged processing operations on the protection of personal data**. A single assessment may address a set of similar processing operations that present similar high risks.

## **Fair Credit Reporting Act (FCRA):**

CRA must implement reasonable procedures to ensure **maximum possible accuracy of consumer reports** and provide consumers with access to their own information, along with the **ability to correct any errors**.

## **Federal Equal Opportunity Laws:**

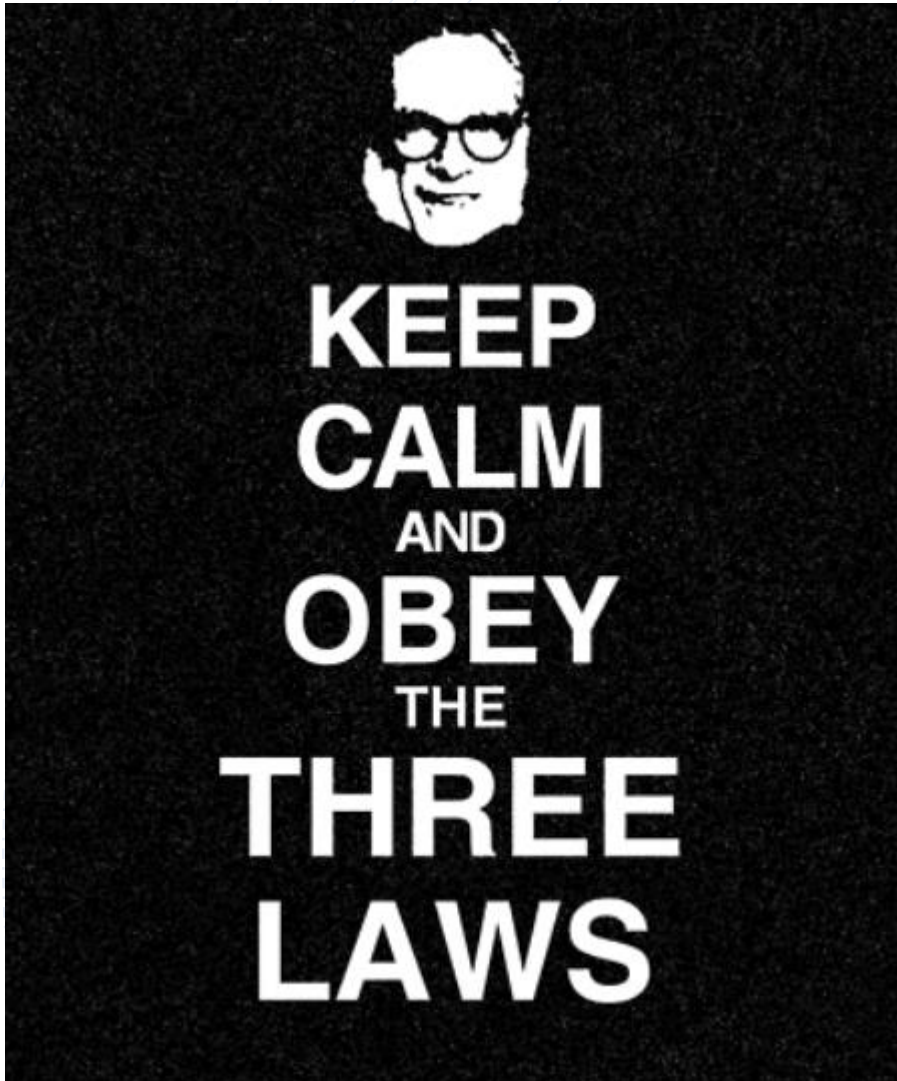
e.g. Equal Credit Opportunity Act (“ECOA”), 88 Title VII of the Civil Rights Act of 1964, the Americans with Disabilities Act, the Age Discrimination in Employment Act (“ADEA”), the Fair Housing Act (“FHA”), and the Genetic Information Nondiscrimination Act (“GINA”).

<https://www.ftc.gov/news-events/blogs/business-blog/2020/04/using-artificial-intelligence-algorithms>

<https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf>



# *Protecting us against AI run amok is not a new idea*



1942




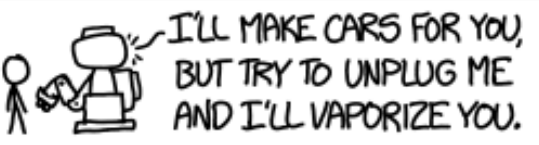

1. A robot may not injure a human being or, through inaction, allow a human to come to harm.
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.



# WHY ASIMOV PUT THE THREE LAWS OF ROBOTICS IN THE ORDER HE DID:

## Obligatory XKCD

<https://xkcd.com/1613/>

POSSIBLE ORDERING	CONSEQUENCES	
1. (1) DON'T HARM HUMANS 2. (2) OBEY ORDERS 3. (3) PROTECT YOURSELF	[SEE ASIMOV'S STORIES]	BALANCED WORLD
1. (1) DON'T HARM HUMANS 2. (3) PROTECT YOURSELF 3. (2) OBEY ORDERS	EXPLORE MARS! 	FRUSTRATING WORLD
1. (2) OBEY ORDERS 2. (1) DON'T HARM HUMANS 3. (3) PROTECT YOURSELF		KILLBOT HELLSCAPE
1. (2) OBEY ORDERS 2. (3) PROTECT YOURSELF 3. (1) DON'T HARM HUMANS		KILLBOT HELLSCAPE
1. (3) PROTECT YOURSELF 2. (1) DON'T HARM HUMANS 3. (2) OBEY ORDERS		TERRIFYING STANDOFF
1. (3) PROTECT YOURSELF 2. (2) OBEY ORDERS 3. (1) DON'T HARM HUMANS		KILLBOT HELLSCAPE



**“Satya Nadella's rules for AI are more boring (and relevant) than Asimov's Three Laws.”**

# *2016 Microsoft Responsible AI Principles*



1. AI must be designed to assist humanity.
2. AI must be transparent.
3. AI must maximize efficiencies without destroying the dignity of people.
4. AI must be designed for intelligent privacy.
5. AI must have algorithmic accountability.
6. AI must guard against bias.



# *2018 AI Code update from UK House of Lords AI committee*

1. AI should be developed for the common good and benefit of humanity
2. AI should operate on principles of intelligibility and fairness.
3. AI should not be used to diminish the data rights or privacy of individuals, families or communities.
4. All citizens should have the right to be educated to enable them to flourish mentally, emotionally and economically alongside AI.
5. The autonomous power to hurt, destroy or deceive human beings should never be vested in AI.

# *Fairness, Ethics, Accountability and Transparency (FEAT) framework (1 of 2)*



## **Fairness**

### **Justifiability**

1. Individuals or groups of individuals are not systematically disadvantaged through AI & Data Analytics (AIDA) driven decisions unless these decisions can be justified.
2. Use of personal attributes as input factors for AIDA-driven decisions is justified.

### **Accuracy and Bias**

3. Data and models used for AIDA-driven decisions are regularly reviewed and validated for accuracy and relevance, and to minimize unintentional bias.
4. AIDA-driven decisions are regularly reviewed so that models behave as designed and intended.

## **Ethics**

5. Use of AIDA is aligned with the firm's ethical standards, values and codes of conduct.
6. AIDA-driven decisions are held to at least the same ethical standards as human-driven decisions.

# *Fairness, Ethics, Accountability and Transparency (FEAT) framework (2 of 2)*



## **Accountability**

### **Internal Accountability**

- 7. Use of AIDA in AIDA-driven decision-making is approved by an appropriate internal authority.
- 8. Firms using AIDA are accountable for both internally developed and externally sourced AIDA models.
- 9. Firms using AIDA proactively raise management and Board awareness of their use of AIDA.

### **External Accountability**

- 10. Data subjects are provided with channels to enquire about, submit appeals for and request reviews of AIDA-driven decisions that affect them.
- 11. Verified and relevant supplementary data provided by data subjects are taken into account when performing a review of AIDA-driven decisions.

## **Transparency**

- 12. To increase public confidence, use of AIDA is proactively disclosed to data subjects as part of general communication.
- 13. Data subjects are provided, upon request, clear explanations on what data is used to make AIDA-driven decisions about the data subject and how the data affects the decision.
- 14. Data subjects are provided, upon request, clear explanations on the consequences that AIDA-driven decisions may have on them.

# Comparison of key principles

Principle	House of Lords	Microsoft	FEAT
Benefit to Humanity	✓	✓	✗
Fairness / No Bias	✓	✓	✓
Intelligibility / Transparency	✓	✓	✓
Restrict autonomous decisions that could cause harm / don't impact individuals rights	✓	Partial "don't destroy dignity"	Only if justified
Ethical decision making	✗	✗	✓
Accountability	✗	✗	✓
Intelligent privacy	✗	✓	✗





# Mitigation Measures



## **Fair**

It is important that AIDA-driven decisions do not disadvantage any particular individual or groups of individuals without justification. The use of AIDA may also create the ability to identify or analyze new types of differentiation that could not previously be done. This could perpetuate cases of unjustified differentiation at a systemic level if not properly managed.

## **Ethical**

All firms using AIDA or not operate in line with their ethical standards. These ethical standards are expressed through many ways, including company values, codes of conduct and mission statements, and may vary across firms and geographies. Adherence to these ethical standards applies equally to the use of AIDA.

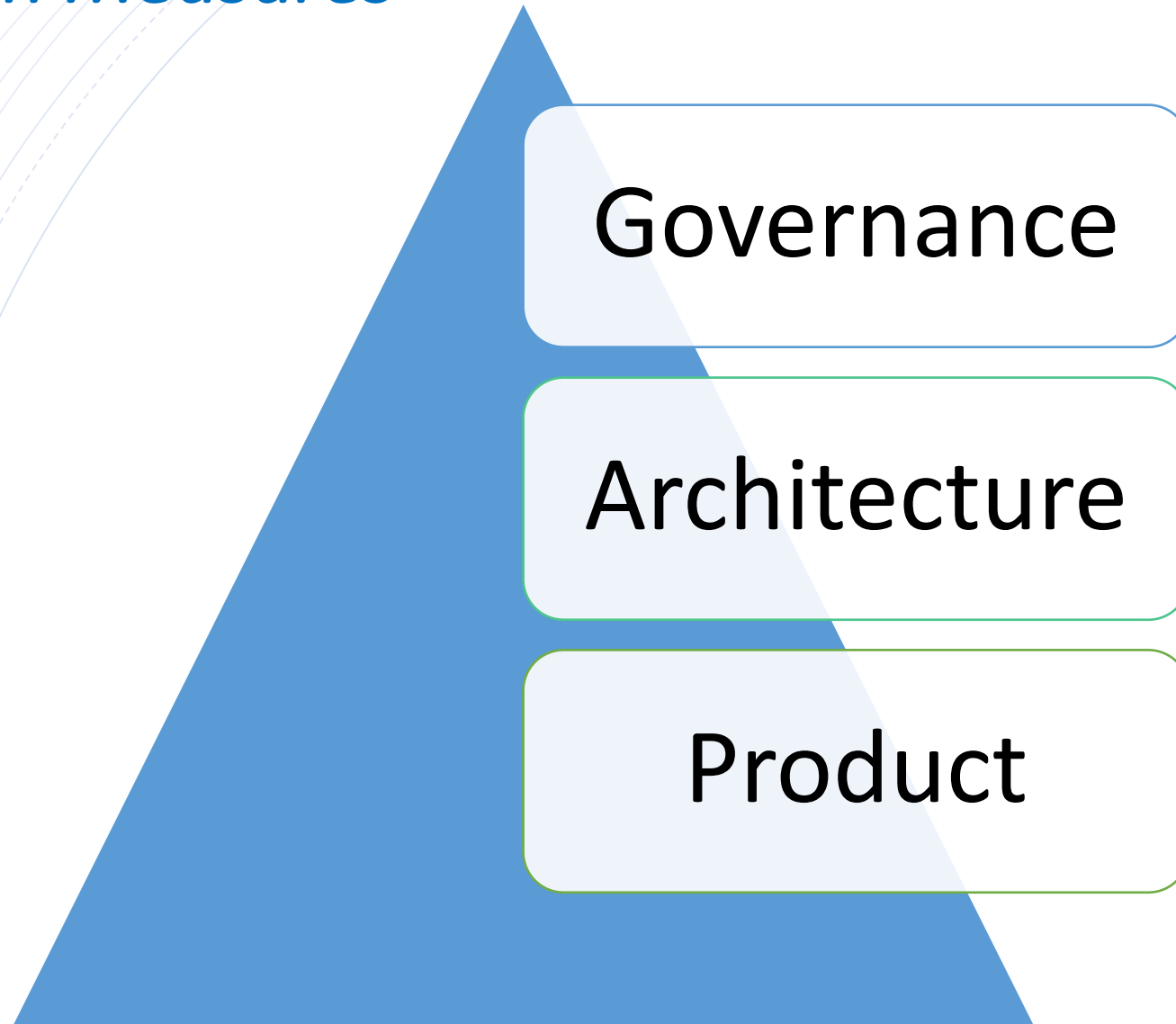
## **Accountability**

It is important that there is clear responsibility for and ownership of AIDA-driven decisions within an AIDA firm, with appropriate internal approving authorities for the use of AIDA. Such accountability applies to all uses of AIDA, whether internally developed or externally sourced.

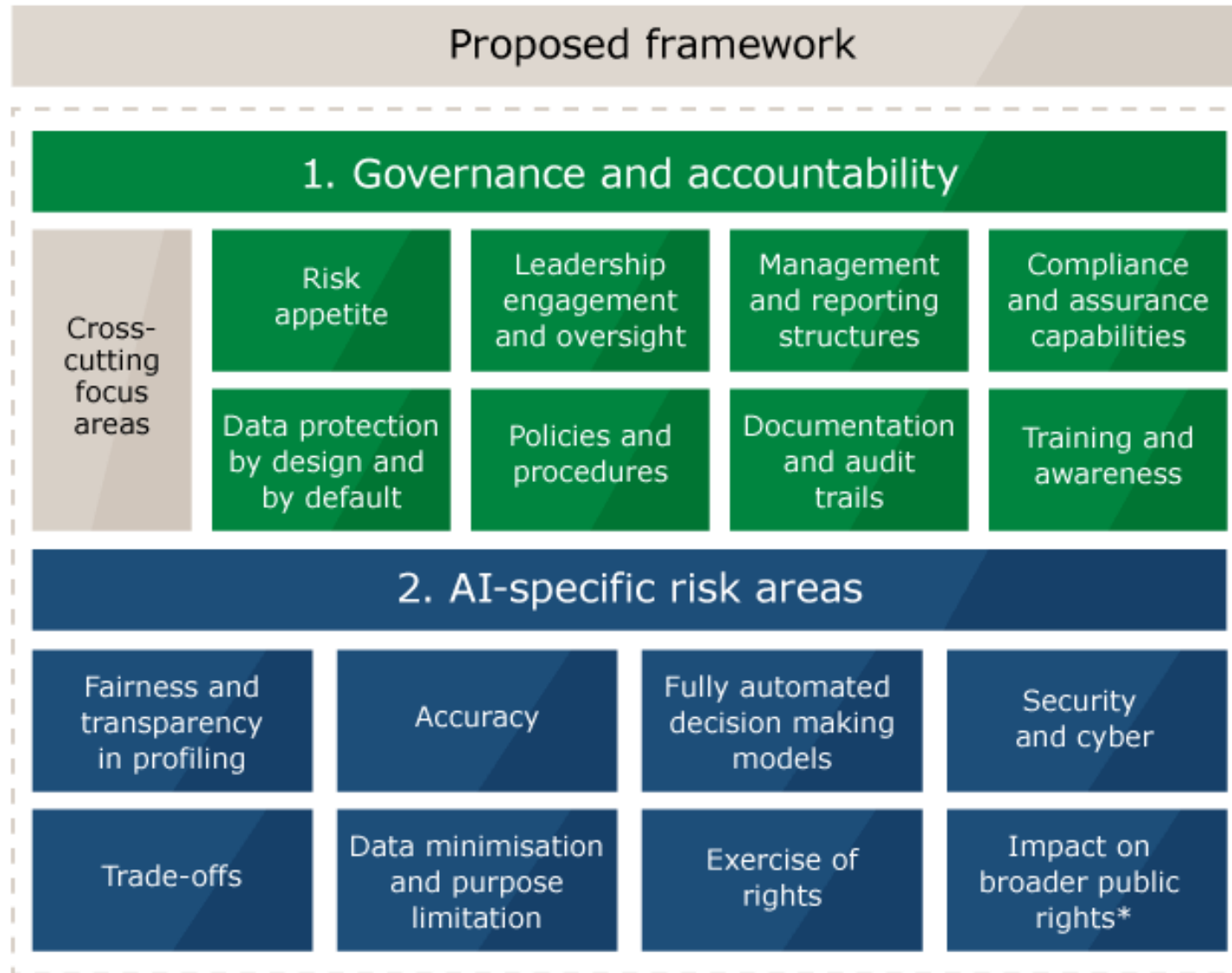
## **Transparency**

While increased transparency in AIDA Firms' use of AIDA could improve public understanding and confidence in AIDA, excessive transparency could create confusion or unintended opportunities for individuals to exploit or manipulate AIDA models. It is important to balance these considerations in determining the appropriate level of transparency in the use of AIDA. In determining levels of transparency, the materiality of the decision is also relevant.

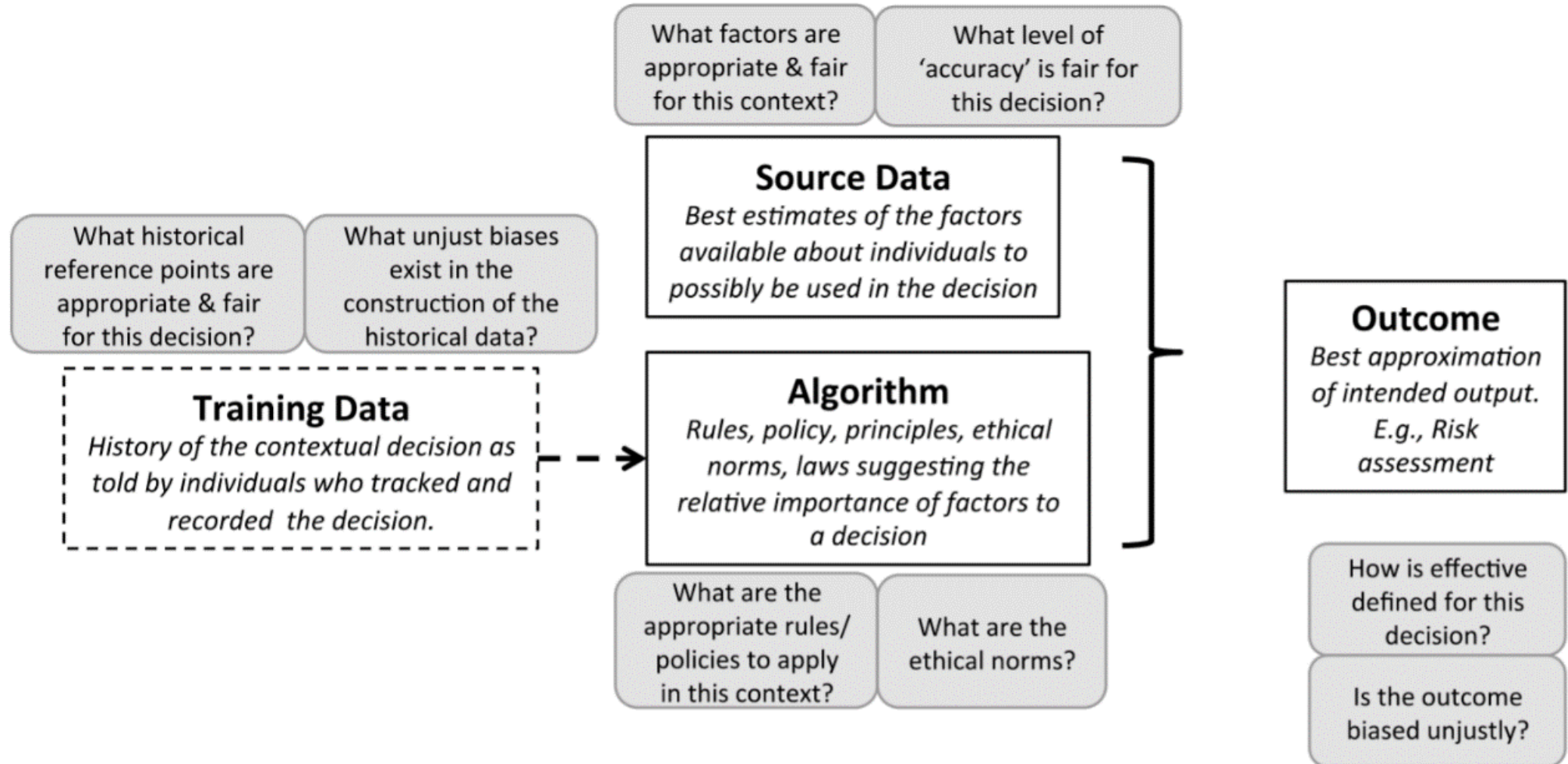
# *Key Mitigation Measures*



# ICO AI Auditing Framework



# Product level controls





# *Asimov's Zeroth law...*

"A robot may not injure humanity, or, by inaction, allow humanity to come to harm."

**Let's hope that we don't end up with AI making the decisions for us.**





# Aaron Weller

FIP, CIPP/US, CIPT, CIPM

VP Strategy, Sentinel

[aaron@sentinelcsg.com](mailto:aaron@sentinelcsg.com)

[cultureofprivacy.com](http://cultureofprivacy.com)



# SENTINEL



# Questions?

- Submit Questions via Chat during Break.
- Please Return at 3:30

# Welcome to Gobias (CoffeeTalk)

- You've been hired as the new CPO for a startup called Gobias.
- Gobias is an online dating service that allows users to match based on coffee shop and other preferences using a proprietary AI Algorithm they have developed.
- Gobias is based in California and plans to focus on the U.S. but thinks they may be big in Amsterdam.





- Gobias' algorithm was trained using a combination of data from public sites (other social media sites and Yelp) and research data bases that to learn about affinities certain types of people have for other people, "Coffee Mates," for certain types of coffee shops and coffee.
- Users download the app to their phones and log in and fill out a profile to start matching. Login to the app is via popular social media logins.
- Data from the app is used by the algorithm to match users to potential "Coffee Mates" based on a variety of data including proximity to coffee shops, dietary preferences, and physical and personal characteristics to match users for a date at a coffee shop.
- The app uses location, the user profile, and data about coffee shops with the algorithm to find "matching" coffee shops close to the "Coffee Mates" and to give directions to shops.
- CoffeeTalk is free for users - funding is ad driven – by mostly coffee shops and coffee-related product advertisers. The algorithm uses all available data to help deliver personalized ads.
- Coffee shops that advertise get special Bluetooth sensors. When a user enters a shop, the sensor detects the user's app and alerts the shop staff.
- Once alerted, shop staff can use the app info to greet the user and make a user's favorite coffee ready. If the user wants to try something different, the algorithm can suggest a new coffee it thinks the user may like.
- All data gathered by the app from users, shop staff, and coffee shops is also used to further train and refine the algorithm.



## CoffeeTalk

