WHERE FROM HERE?

a conversation series about the great acceleration of 2020 and what it means for the future

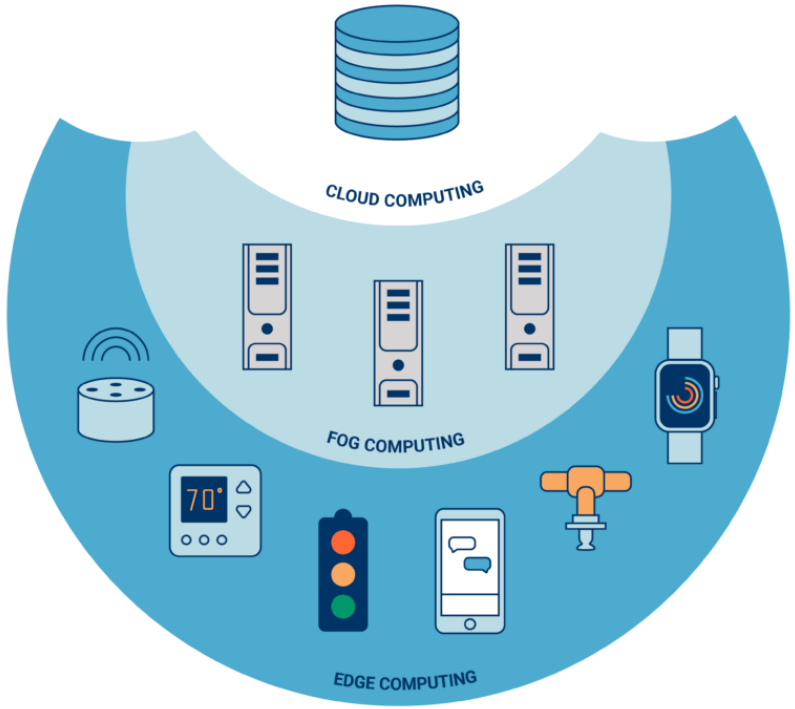
orrick

DATA AS AN ASSET – HOW TO ACQUIRE IT AND USE IT SAFELY

orrick

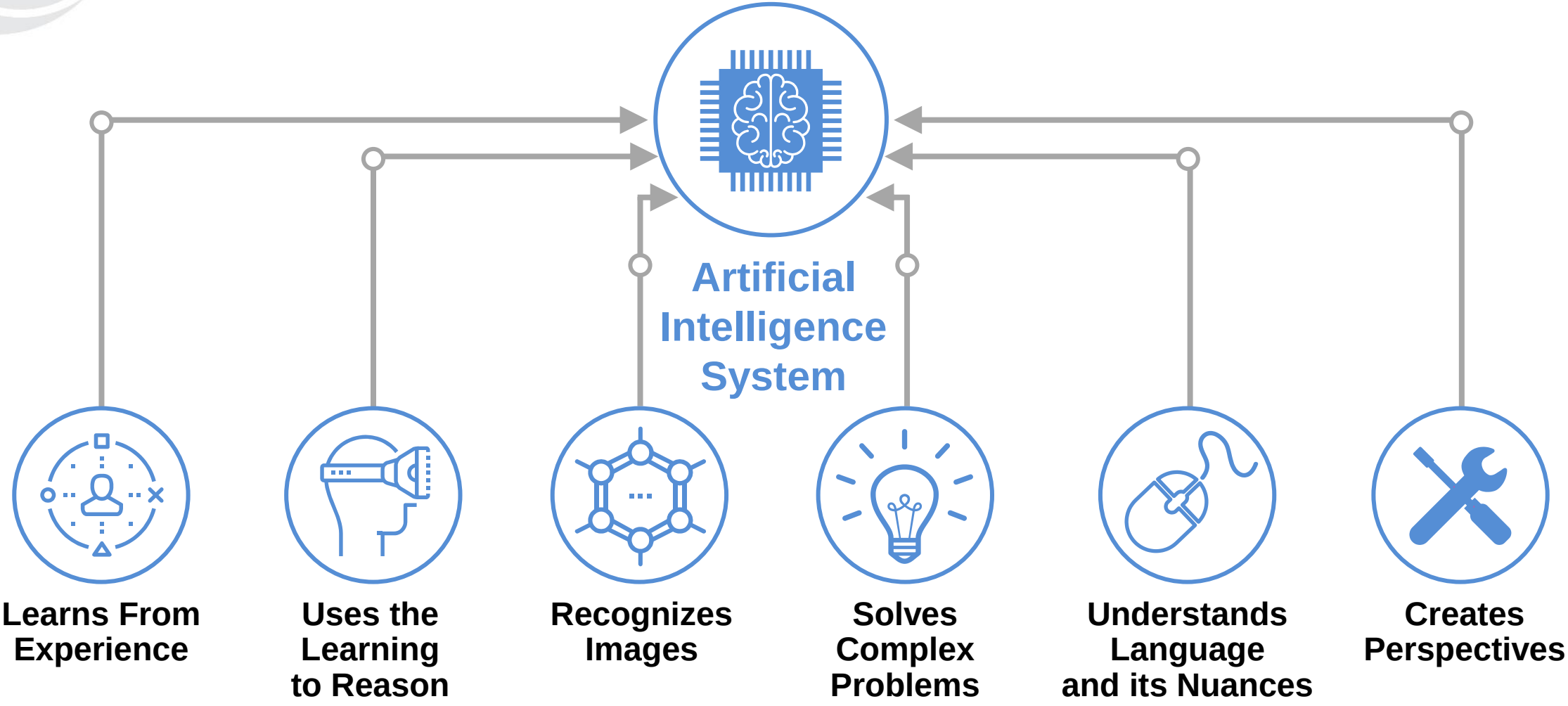
AI: Why Is It Real Now?

Mass Distribution of Computing Capability



Rapidly Declining Cost of Storage

AI: Capabilities

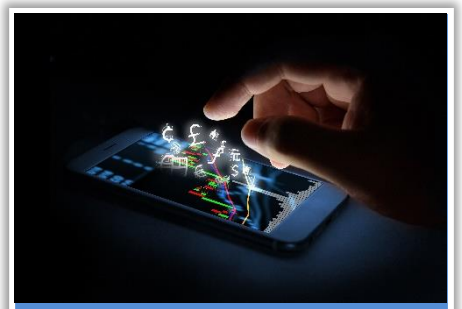


AI: Impact

AI is already being used for many services and products

HEALTH CARE:
Diagnostic Platform

CYBERSECURITY:
Anomaly Detection



FINTECH:
Robo-Investment Advisors

INTELLIGENT CARS

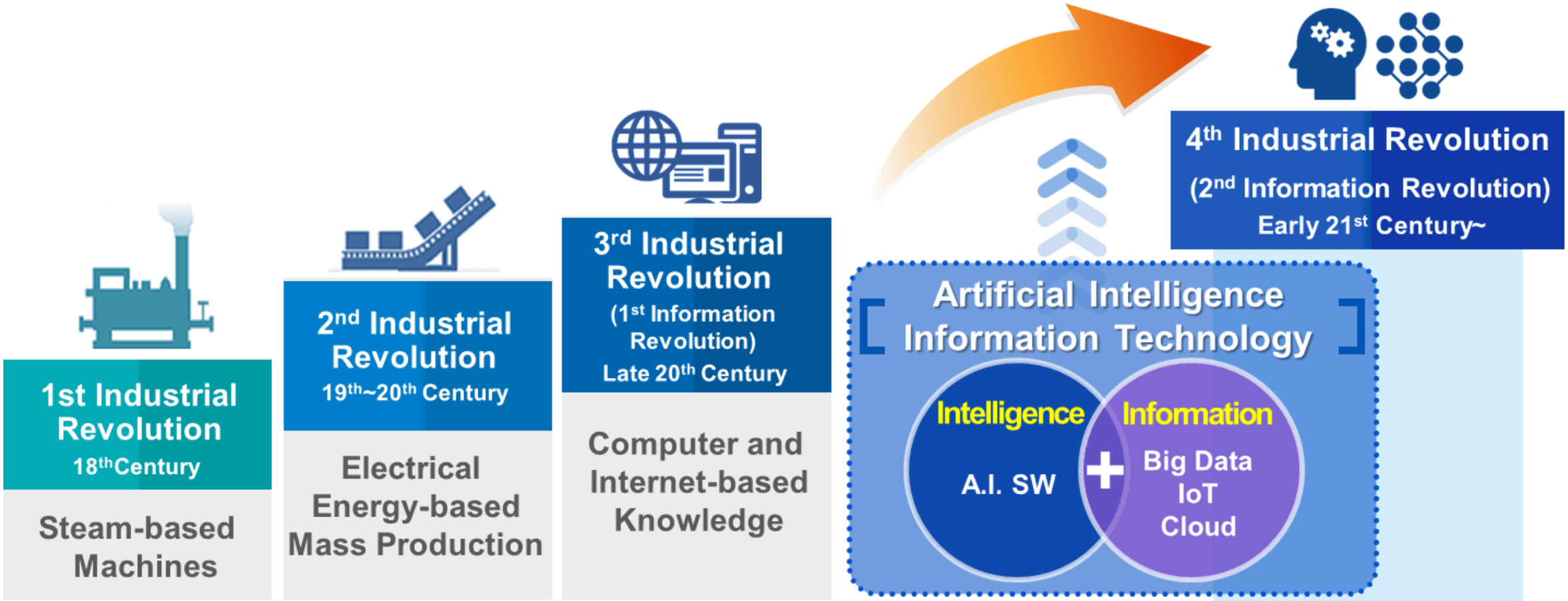
AGRI-FOOD:
Smart Farming

PERSONALIZED MARKETING:
Chatbots

VIDEO GAMES:
Customized Player Interaction

VIRTUAL ASSISTANT

AI: Impact


1st Industrial Revolution
18th Century
Steam-based Machines
2nd Industrial Revolution
19th~20th Century
Electrical Energy-based Mass Production
3rd Industrial Revolution
(1st Information Revolution)
Late 20th Century
Computer and Internet-based Knowledge
4th Industrial Revolution
(2nd Information Revolution)
Early 21st Century~
Artificial Intelligence
Information Technology
Intelligence
A.I. SW
Information
Big Data
IoT
Cloud

Agenda

- Privacy and Security Issues
- Intellectual Property & Other Problems in Data Acquisition
- AI Ethical Issues

PRIVACY & SECURITY

Why is Data Privacy and Security Relevant?

DATA PROTECTION PRINCIPLES	AI CHARACTERISTICS
Transparency and notice *Disclosing the collection, use, sharing, and storage practices for personal information*	Development and implementation can involve complex processing activities making disclosure difficult
Choice *Specifying the choices people have about how their personal information is used*	Complex systems make informed choice out of reach for some individuals, and implementation of choice may be difficult for some systems
Minimization *Limiting the personal information collected to the minimum necessary*	AI systems often rely on large data sets to learn, and developers may not know which data is relevant
Use and retention *Use is limited to the reasons for collection and retained for no longer than is necessary*	AI systems often rely on third-party data to learn, and it can be difficult to know how long data will be relevant in connection with AI
Access and accuracy *Providing users with the ability to access and correct their personal information*	Verifying accurate data outputs may be difficult in a “Blackbox” model, and access to underlying data may be complicated
Security *Taking reasonable steps to protect personal information*	Physical and logical separation may be difficult where AI systems rely on large data sets to train and verify models

Comprehensive Data Protection Laws

Comprehensive data protection legislation has been gaining momentum in recent years, with federal and state governments passing laws across the globe.

General Data Protection Regulation

The European Union's comprehensive privacy regulation enforceable as law in all EU Member States as of May 2018

California Consumer Privacy Act

The state of California's comprehensive privacy law that took effect January 2020 and is set for significant amendment in January 2023 due to CPRA

Lei Geral de Proteção de Dados Pessaoais

Brazil's comprehensive privacy law inspired by the EU's GDPR, which took effect in September 2020

The U.S. Congress continues to consider comprehensive federal data protection legislation. In addition, many states considered CCPA-like bills in 2020 and may propose similar bills in 2021.

Subject Matter Specific Laws

An AI system may be governed by subject matter specific laws when its input data or the purposes for which it is processing personal information trigger a particularly sensitive subtype of personal information:

Biometric Information:

- "Biometric information" means a retina or iris scan, fingerprint, voiceprint, or scan of hand or face geometry which is used to identify an individual.
- General state laws in IL, TX & WA.
- IL's Biometric Information Privacy Act is the most comprehensive requiring:
  - Disclosure of biometric information being collected, its specific purpose and the length of term it will be stored (no greater than 3 years)
  - Obtain a written release executed by the data subject before collection
- BIPA class actions are popular due to private right of actions with "liquidated damages."

Health Information:

- "Personal health information" means individually identifiable health information that:
  - A covered entity creates/receives
  - Is maintained or transmitted in any form; and
  - Is related to an individual's past, present or future health.
- HIPAA Privacy Rule
  - Notify individuals about privacy rights and how PHI can be used;
  - Implement privacy procedures;
  - Train employees on procedures;
  - Designate a privacy lead.
- HIPAA Security Rule
  - Requires appropriate administrative, physical and technical safeguards.

Financial Information:

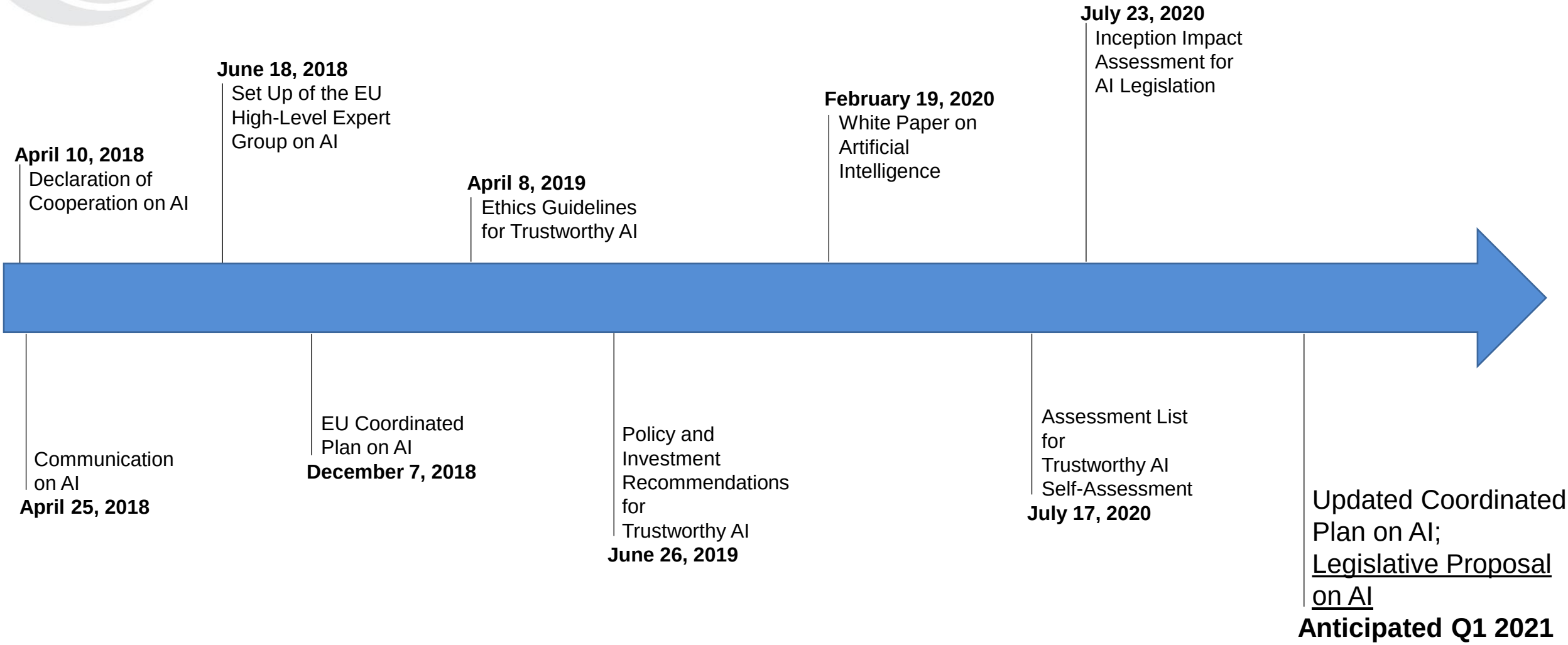
- "Nonpublic personal information" is information provided by a consumer, derived from a transaction with a consumer or otherwise obtained in connection with providing a financial product or service to that consumer.
- GLBA Privacy Rule
  - Notify individuals about privacy rights and how NPI can be used;
  - Limit and allow consumers to opt-out of certain NPI sharing.
- GLBA Safeguards Rule
  - Develop a written information security program implementing technical, administrative and physical safeguards.

Children's Information:

- Children's "Personal Information" is individually identifiable information collected online about an individual under the age of 13.
- Must:
  - Provide online notice of data practices;
  - Provide parents direct notice;
  - Permit parents to review data collected and opt-out;
  - Do not condition participation on unnecessary data collection;
  - Maintain reasonable procedures to protect data.

But What's Next . . .

European Union is Leading the Way



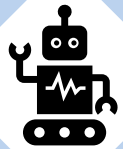
Predicting the Focus of EU AI Regulation

Civil and Tort Liability

- Not necessary to give AI systems their own legal personality.
- Harm caused by AI systems can and should be attributed to existing legal persons.
- Strict liability may be an option in non-private environments that have the capability of typically causing significant harm.
- Alternatively, or in combination, burden shifting regimes that are more favorable to plaintiffs may be appropriate for all AI.

Consumer and Data Protection Laws

- Apply existing consumer and data protection laws to AI more concretely:
  - GDPR rights in relation to automated-decision;
  - Scrutinize legal bases for processing personal data;
  - Reassess which protective measures are appropriate.
- Consider new laws addressing sufficient market access to data, partial systems automation and processing of non-personal data.



Non-Discrimination Laws

- Accidental indirect discrimination is more likely to occur than intentional discrimination.
- Existing laws create rebuttable open-ended standards.
- AI systems can make it difficult to identify biased outcomes.
- EU law may be updated to impose a greater obligation to verify the absence of unjust bias and improve applicability of non-discrimination enforcement mechanisms to AI.

Competition Rules

- EU views access to data as a key ingredient for a fair and competitive market, particularly for data's value for AI.
- Volume of data and its value likely to be taken into consideration when applying rules on anti-competitive behavior, abuse of dominance or collusion, and when evaluating mergers.
- Competition remedies are likely to be designed to provide the market better access to even proprietary data.

Adapting AI Systems to Address Anticipated Regulation

Artificial intelligence will create **new risks** for companies implementing the advanced technology, but companies taking **proactive steps** to adapt AI systems to **properly address anticipated regulation** may find themselves with a **competitive advantage**.

Data and AI Governance

Develop robust data and AI governance structures that oversee the proper use of AI systems and the data used in their training and implementation.

Diverse Team of AI Developers and Operators

Engage diverse and inclusive AI teams capable of considering a diversity of opinions in AI development, identify potential biases in AI implementation, and address potential disparate impacts in AI outcomes.

Data Protection by Design & by Default

Embed appropriate measures designed to implement data protection principles through the lifecycle of an AI system and to ensure that, by default, only personal data which are necessary for use in the AI system are processed.

Secure AI Systems and Databases

Adopt custom-fit security measures based on the level and type of risks that arise from the specific AI processing activities, considering not only potential loss of data but also manipulation of AI outcomes.

Promote Transparency

Provide clear notice of the data being collected, how that data is intended to be used, and when the individual is or will be interacting with an AI system, taking into consideration protection of IP and competitive procedures.

Human Intervention

Identify when and how to implement human intervention or oversight into AI systems to maximize the benefits of AI operations while addressing concern for unchecked, automated systems.

IP & OTHER DATA ACQUISITION ISSUES

Data Acquisition & Use

- Data is the fuel that drives ML.
  - Ensuring proper access and use rights to data is often an overlooked concern
  - Data can be acquired from a wide variety of sources
    - Published datasets
    - Data from customers/corporate partners
    - Data scraped from public sources
    - Synthetic data
- Each source has its own mix of issues

IP: AI DEVELOPMENT AND DEPLOYMENT

Protect/Avoid Protection

- Datasets
- Data structures
- Synthetic Data
- Improved algorithms
- Outputs

Potentially Relevant Forms of IP

Dataset Cases

- *hiQ v. LinkedIn* – Startup sues platform alleging right to scrape public data for use in algorithm-based service.
- *CCC v. Tractable* – Large insurance service company sues startup alleging improper acquisition of estimate data for use in insurance ML.
- *Compulife Software Inc. v. Newman* – Automated acquisition of entire database one record at a time may violate trade secret laws.
- *U.S. v. Van Buren* – Police Officer's use of license/registration databases for non-work purpose may violate CFAA.

Data Acquisition & Use

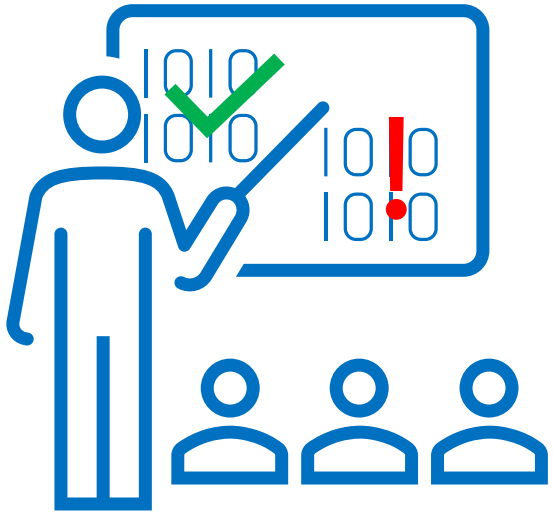
- Potential impediments
  - Terms of Use
  - Computer Fraud & Abuse Act (CFAA) & state law equivalents
  - Online trespass
  - DMCA anti-circumvention prohibitions

- Recommendation: implement a data counseling practice and a data inventory approach

DATA ACQUISITION & USE

Policy Development

- A data use and tracking policy allows an organization to:
  - Manage data use risk
  - Comply with data usage obligations
  - Prepare for any needed diligence and data privacy compliance
- A basic policy is implemented through:
  - Intake forms, decision matrix, & tracking database
  - Options for management
    - Developer self management, managed by legal, hybrid



AI ETHICAL ISSUES

Seven Requirements for Trustworthy AI

1. Human Agency and Oversight

*AI systems should **empower human beings** to hold some autonomy in relation to decisions made by AI models. In addition, the AI system should **involve human oversight**, which can be achieved by having humans involved in determining when to use the system, monitoring the system and/or having input in every cycle of the system.*

2. Technical Robustness and Safety

*AI systems should be **secure**, have a **fallback plan in case of error or attack**, and be **safe, accurate, reliable and reproducible**.*

3. Privacy & Data Governance

Adequate governance should include respect for privacy, quality and integrity of data, and ensure restricted and legitimized access to data

4. Transparency

AI data, systems and models should be transparent, traceable, and capable of explanation catered to the stakeholder concerned, with proper communication regarding the use of AI models and their capabilities and limitations.

5. Diversity, Non-Discrimination and Fairness

AI systems should be accessible to all regardless of disability, avoid marginalizing groups or incorporating unfair biases, and involve diverse stakeholders.

6. Societal and Environmental Well-Being

AI systems should benefit all human beings, they should be environmentally friendly, and societal impact should be carefully understood

7. Accountability

AI systems should be accompanied by mechanisms designed to make algorithms auditable (with known data and design processes), to document and minimize potential negative impacts to human well-being and dignity, and to provide individuals with adequate redress when unjust adverse impacts occur.

Ethical AI Guiding Principles

AI systems must be able to facilitate full and effective human self-determination over an individual's person and his or her rights, and should follow human-centric design principles leaving meaningful opportunity for human choice and oversight.

AI systems should neither cause nor exacerbate harm or otherwise adversely affect human beings, taking due consideration of vulnerable populations, asymmetrical power positions, the natural environment and all living beings.

The development, deployment and use of AI systems must be fair both substantively (e.g., ensuring equal and just distribution of benefits/costs without discrimination) and procedurally (e.g., decisions made by AI systems should be contestable and redressable).

The processes used by AI systems must be transparent, their capabilities and purposes openly communicated, and their decisions—to the extent possible—explainable to those directly and indirectly affected.

Principles for the Stewardship of AI Applications

1. Public Trust In AI

Aimed at ensuring that regulatory and non-regulatory approaches to AI promote reliable, robust, and trustworthy AI applications that will contribute to public trust in AI. The appropriate approaches should balance the nature of the potential risk against the appropriate mitigation.

2. Public Participation

Agencies are advised to create opportunities for the public to participate in the rulemaking process, to the extent feasible and consistent with legal requirements.

3. Scientific Integrity and Information Quality

Regulatory and non-regulatory use of AI should "leverage scientific and technical information and processes." Such information should be held to "a high standard of quality, transparency, and compliance."

4. Risk Assessment and Management

Under this principle, a risk-based approach to using AI should be applied consistently across all agencies and across all applicable technological platforms instead of "hazard-based" and precautionary approaches to AI that "could unjustifiably inhibit innovation."

5. Benefits and Costs

Agencies are directed to "carefully consider the full societal costs, benefits, and distributional effects before considering regulations" and weigh whether implementation will alter, mitigate, or aggravate errors in existing operations.

6. Flexibility

Agencies developing regulatory and non-regulatory approaches are advised to pursue AI use and application that is both evidence- and performance-based, as well as sufficiently flexible to adapt to developments in technology.

7. Fairness and Non-Discrimination

Agencies weighing regulations or non-regulatory uses and application of AI should scrutinize whether such use and application will reduce or increase bias and discrimination and impact "public trust and confidence in AI."

8. Disclosure and Transparency

In addition to complying with existing laws and regulations, agencies should strive to maximize disclosure and transparency on a case-specific basis in order to increase public trust and confidence in the use and application of AI.

9. Safety and Security

This principle directs agencies to promote the development of AI uses and applications "that are safe and secure and operate as intended and should encourage the consideration of safety and security issues throughout the AI design, development, deployment, and operation process," particularly as regards" confidentiality, integrity, and availability of the information processed, stored, and transmitted by AI systems."

10. Interagency Coordination

Agencies are advised to coordinate and share information about the success of using AI in order "advance American innovation and growth in AI, while appropriately protecting privacy, civil liberties, and American values and allowing for sector- and application-specific approaches when appropriate."

OECD Principles on Artificial Intelligence

- **1** – AI should benefit people and the planet by driving inclusive growth, sustainable development and well-being;
- **2** – AI systems should be designed in a way that respects the rule of law, human rights, democratic values and diversity, and they should include appropriate safeguards – for example, enabling human intervention where necessary – to ensure a fair and just society;
- **3** – There should be transparency and responsible disclosure around AI systems to ensure that people understand when they are engaging with them and can challenge outcomes;
- **4** – AI systems must function in a robust, secure and safe way throughout their lifetimes, and potential risks should be continually assessed and managed; and
- **5** –Organizations and individuals developing, deploying or operating AI systems should be held accountable for their proper functioning in line with the above principles.

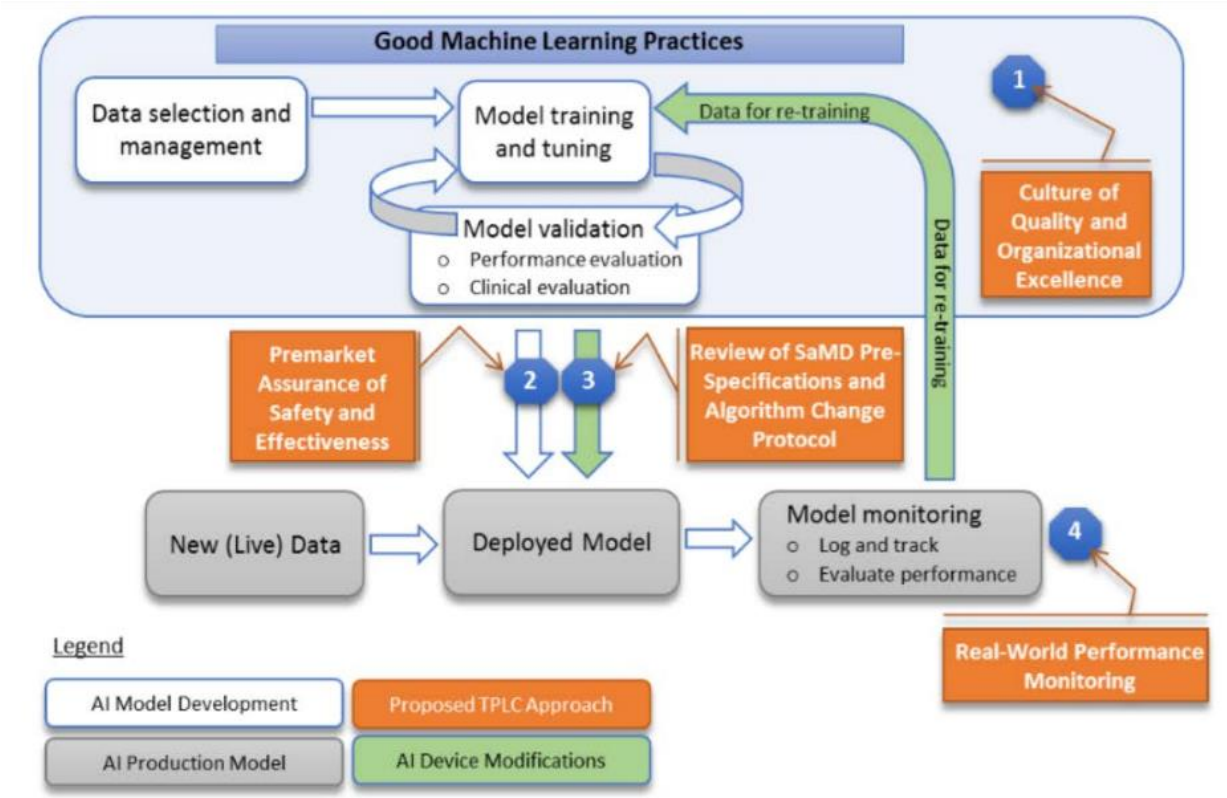
AI Ethics/Policy Issues

- **Real World Examples** – The FDA's "Action Plan" for good machine learning practices (GMLP) and the Federal Reserve's focus on bias.
  - **Explainability** – Increased calls for explainability of AI systems to understand how systems make their predictions, including how and why they might fail. (Significant products liability concerns).
  - **Bias** – AI models reflect the data used to train them, if data reflects real-world bias then models are also likely to reflect such bias. Models should be both *nondiscriminatory* and *effective* for all.

Explainability

- **Explainability –**
  - Validation
    - Can you trust the AI's results
  - Regulatory approval
    - Needed to determine efficacy and failure modes
  - Products liability
    - More information provided to downstream vendors makes the AI system a tool they can make reasonable decisions about.

FDA GMLP

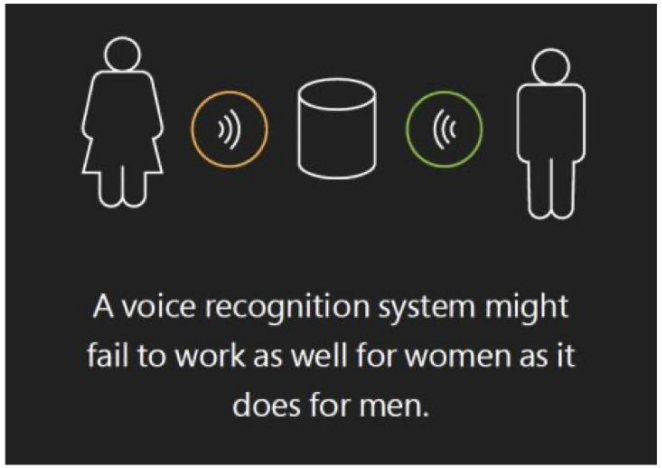


AI ETHICS/POLICY ISSUES

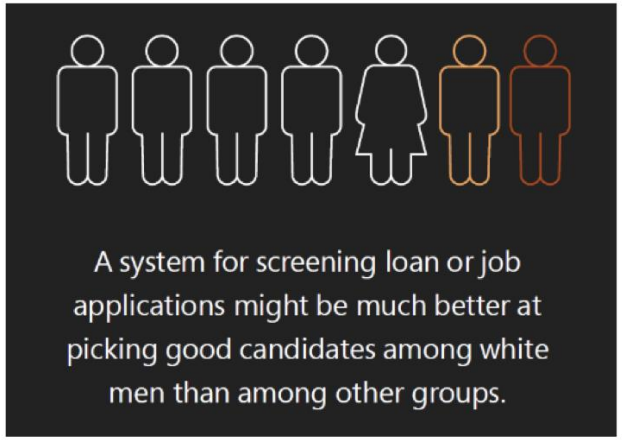
Bias

- Bias types
  - QoS harms
    - AI less effective for certain groups.
  - Allocation harm
    - Grants resources disproportionately
- Bias may not be solvable
  - Focus is often on mitigation of disproportionate results
    - Use synthetic data to even out dataset representation
    - Refactor models to provide a fairer outcome

Example of a quality-of-service harm



Example of an allocation harm



* https://www.microsoft.com/en-us/research/uploads/prod/2020/05/Fairlearn_WhitePaper-2020-09-22.pdf

Key Takeaways

- Artificial intelligence has been and will continue to be a hot button topic in the U.S. and around the world.
- While AI comes in many forms, there is a clear desire to regulate situations where systems and models take the place of human operators in making decisions that impact individuals.
- The intrinsic value data provides to AI models suggests data privacy & security will play a key role in future regulation of AI, particularly where data protection principles and AI characteristics collide.
- Existing data privacy and security laws already govern the use of AI models to some extent; especially comprehensive privacy laws governing all processing of personal data (like GDPR and CCPA).
- However, the European Union is leading the way in developing AI-specific regulation, which is likely to focus on ethical and robust AI designed to respect the fundamental rights of individuals (e.g., Privacy).
- The United States is also developing its own AI-specific regulation strategy but places a larger emphasis on promoting innovation and growth of artificial intelligence technologies.
- **Proactive steps to adapt AI systems to properly address anticipated regulation assists to mitigate risk but also presents an opportunity for competitive advantage.**