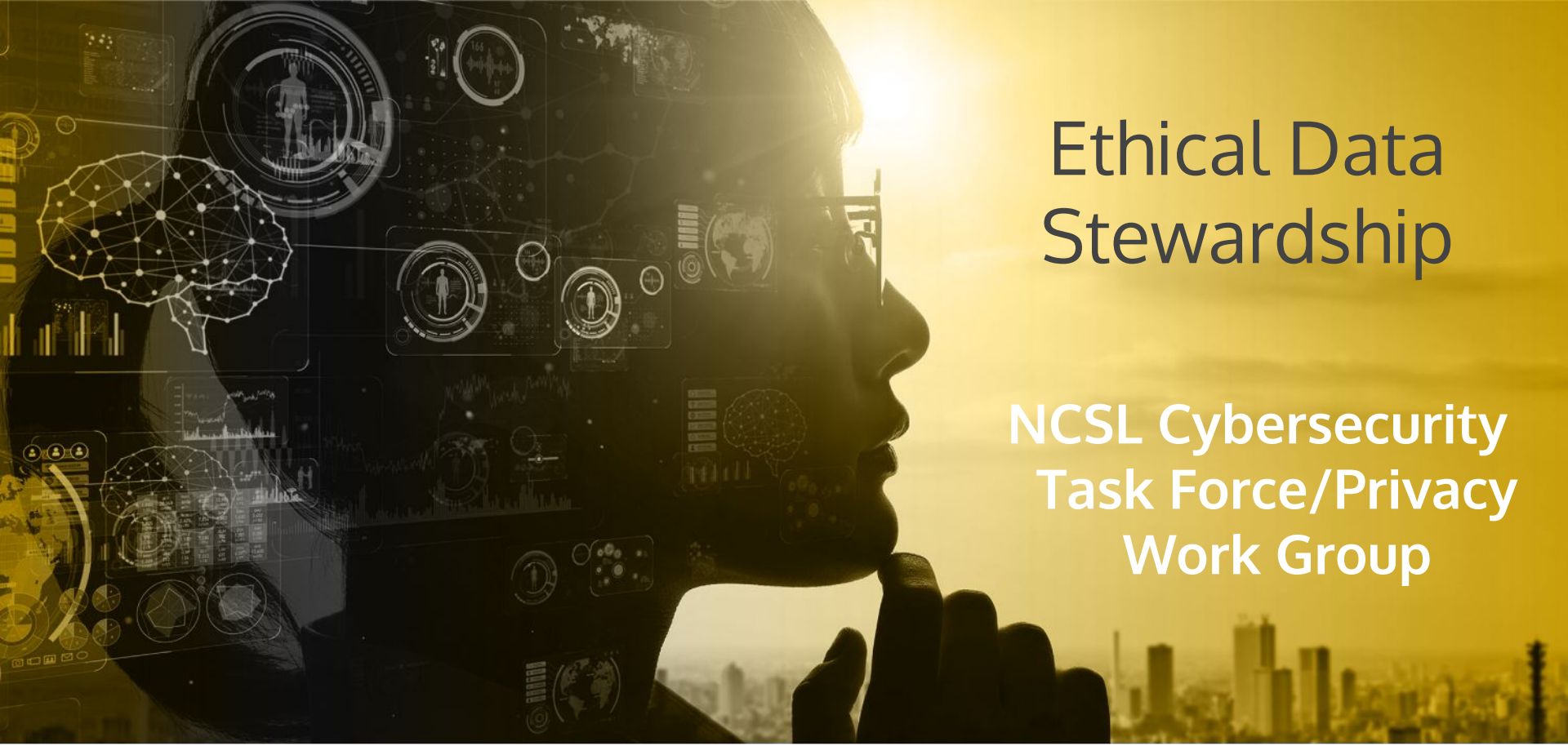
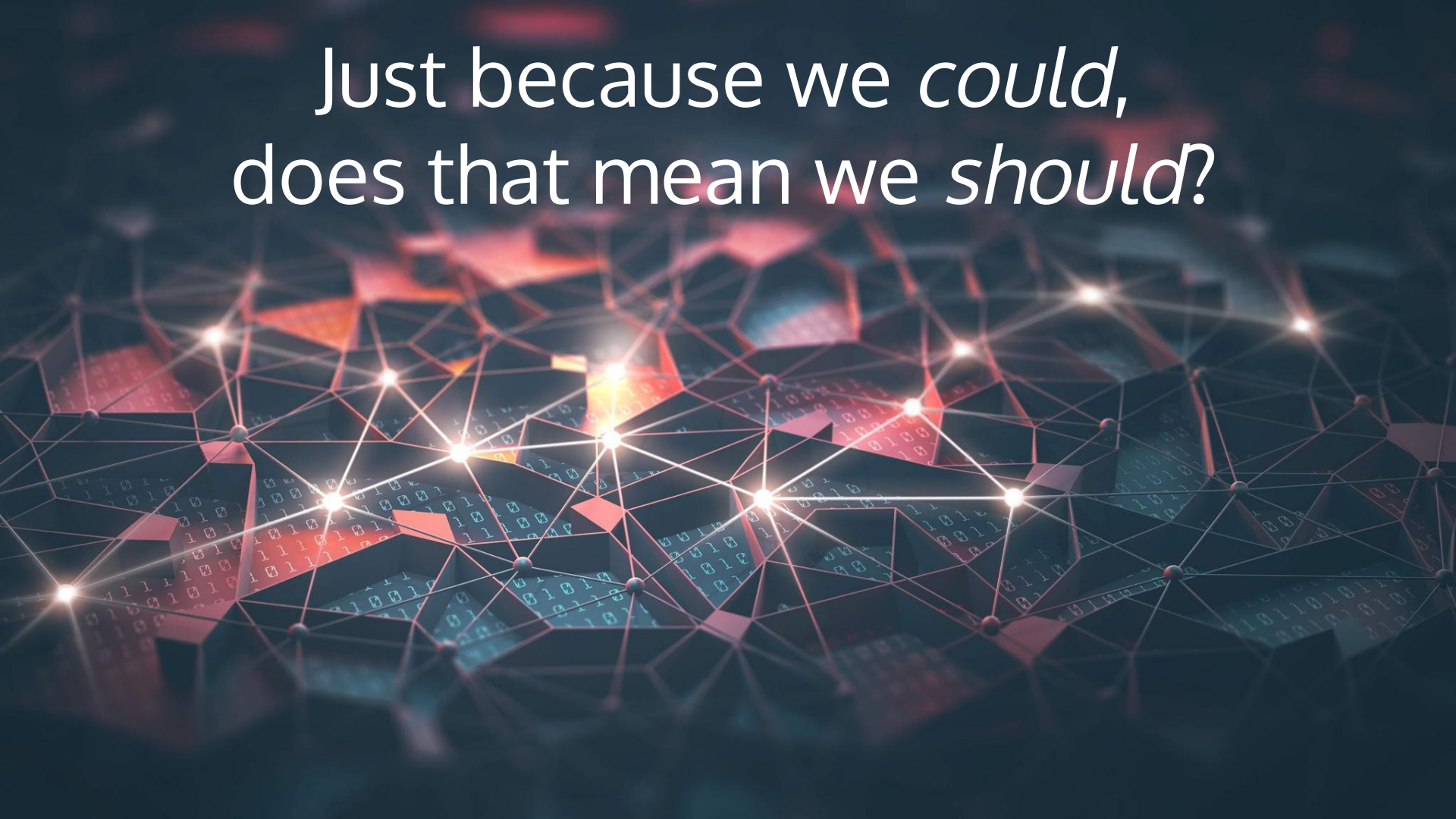


Ethical Data Stewardship

NCSL Cybersecurity Task Force/Privacy Work Group

Michael Best

NCSL
NATIONAL CONFERENCE OF STATE LEGISLATURES


Just because we *could*,
does that mean we *should*?

How an AI grading system ignited a national controversy in the U.K.

AXIOS

AI and ethics: One-third of executives are not aware of potential AI bias

TechRepublic.

REPORT

Credit denial in the age of AI

Aaron Klein · Thursday, April 11, 2019

BROOKINGS

Study: Only 18% of data science students are learning about AI ethics

The neglect of AI ethics extends from universities to industry

TNW

Why IBM Decided to Halt all Facial Recognition Development

IBM disapproves of any technology that could lead to racial profiling.

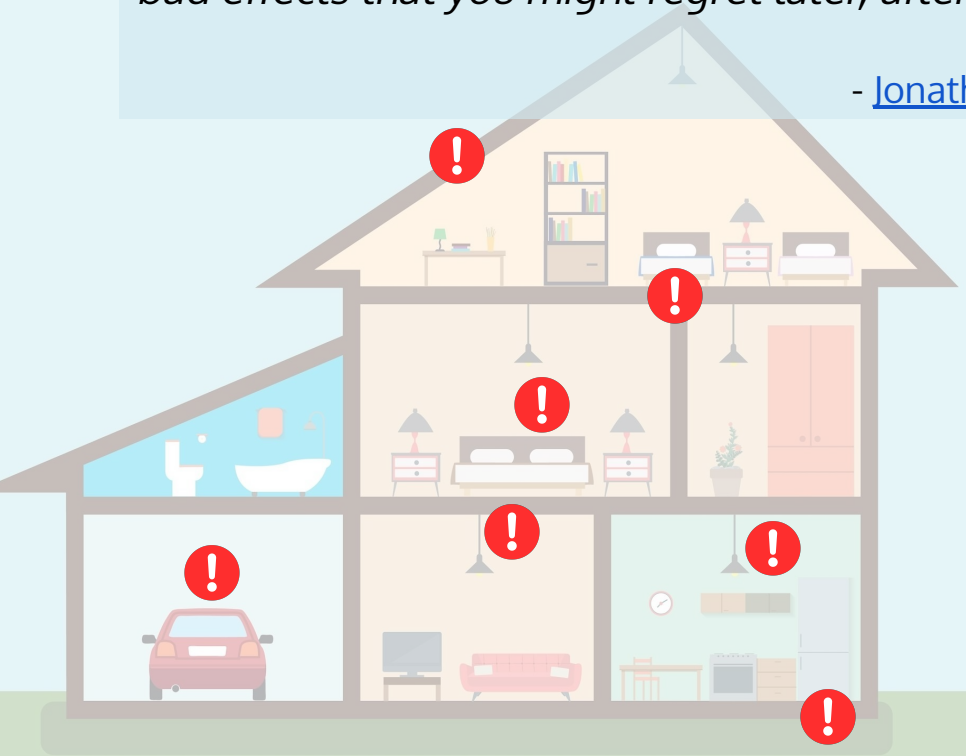
Medium

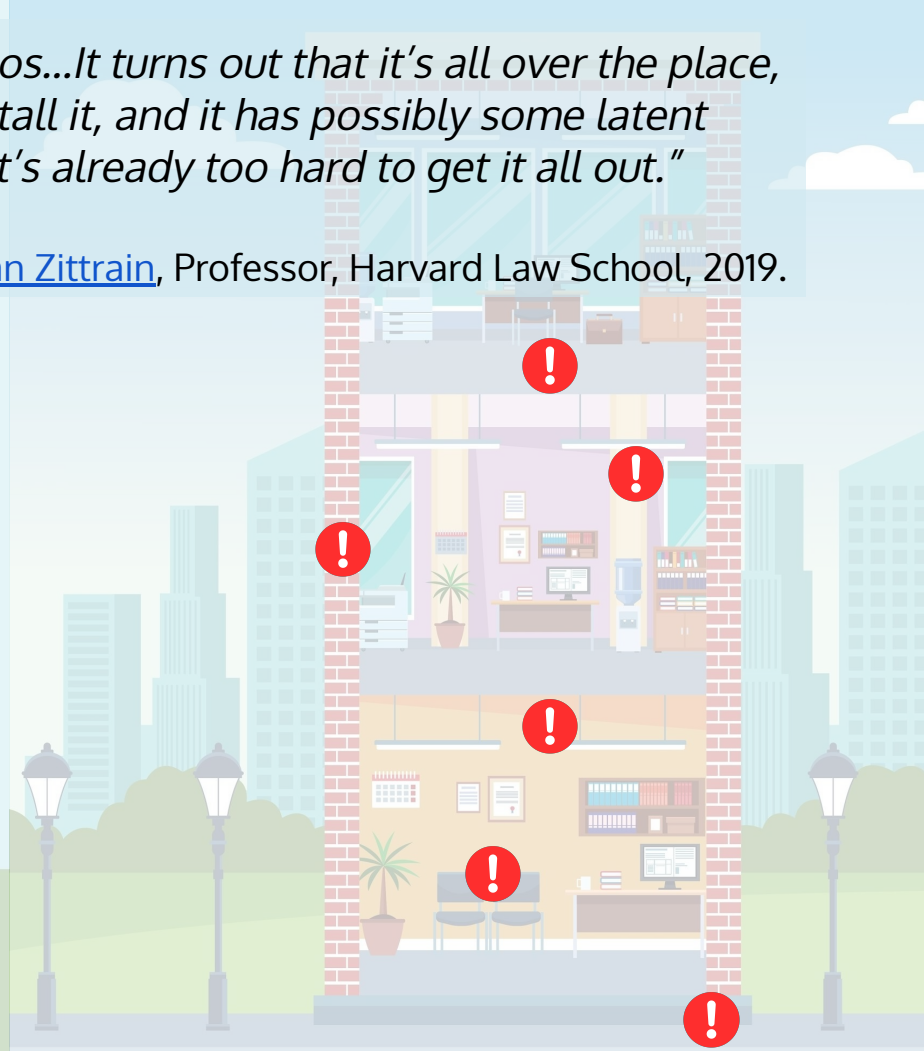
U.S. warns of discrimination in using artificial intelligence to screen job candidates

npr

"I think of machine learning kind of as asbestos...It turns out that it's all over the place, even though at no point did you explicitly install it, and it has possibly some latent bad effects that you might regret later, after it's already too hard to get it all out."

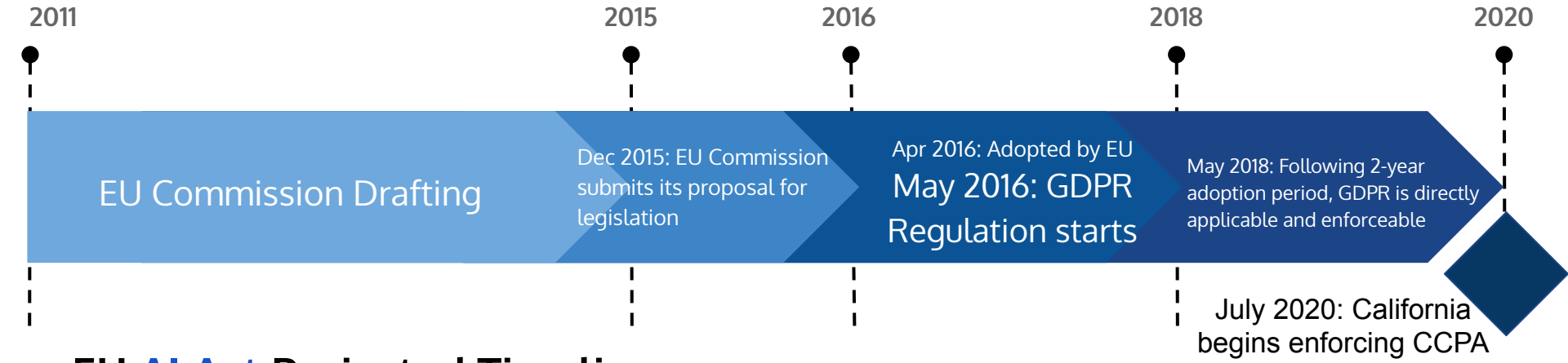
- Jonathan Zittrain, Professor, Harvard Law School, 2019.



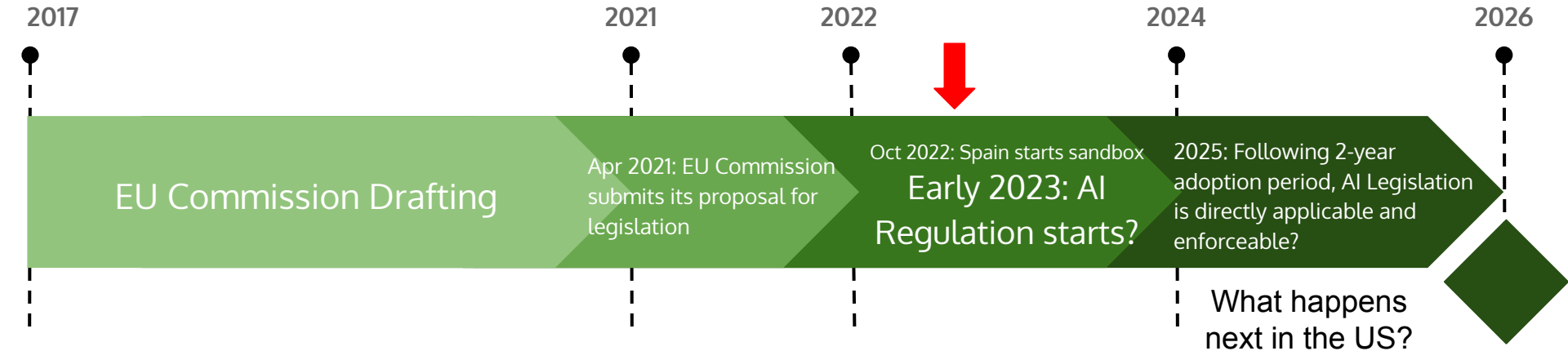



EU Commission's
Concerns for AI
Rogue or Unsafe
Compromising Privacy
Biased & Discriminatory
Damaging to Society
Lacking Accountability
European
Commission

General Data Protection Regulation (GDPR) Actual Timeline



EU AI Act Projected Timeline



High Risk AI Systems (EU Commission)

US Executive Order 13960 (Dec 2020)
9 Principles for designing, developing, acquiring, and using AI

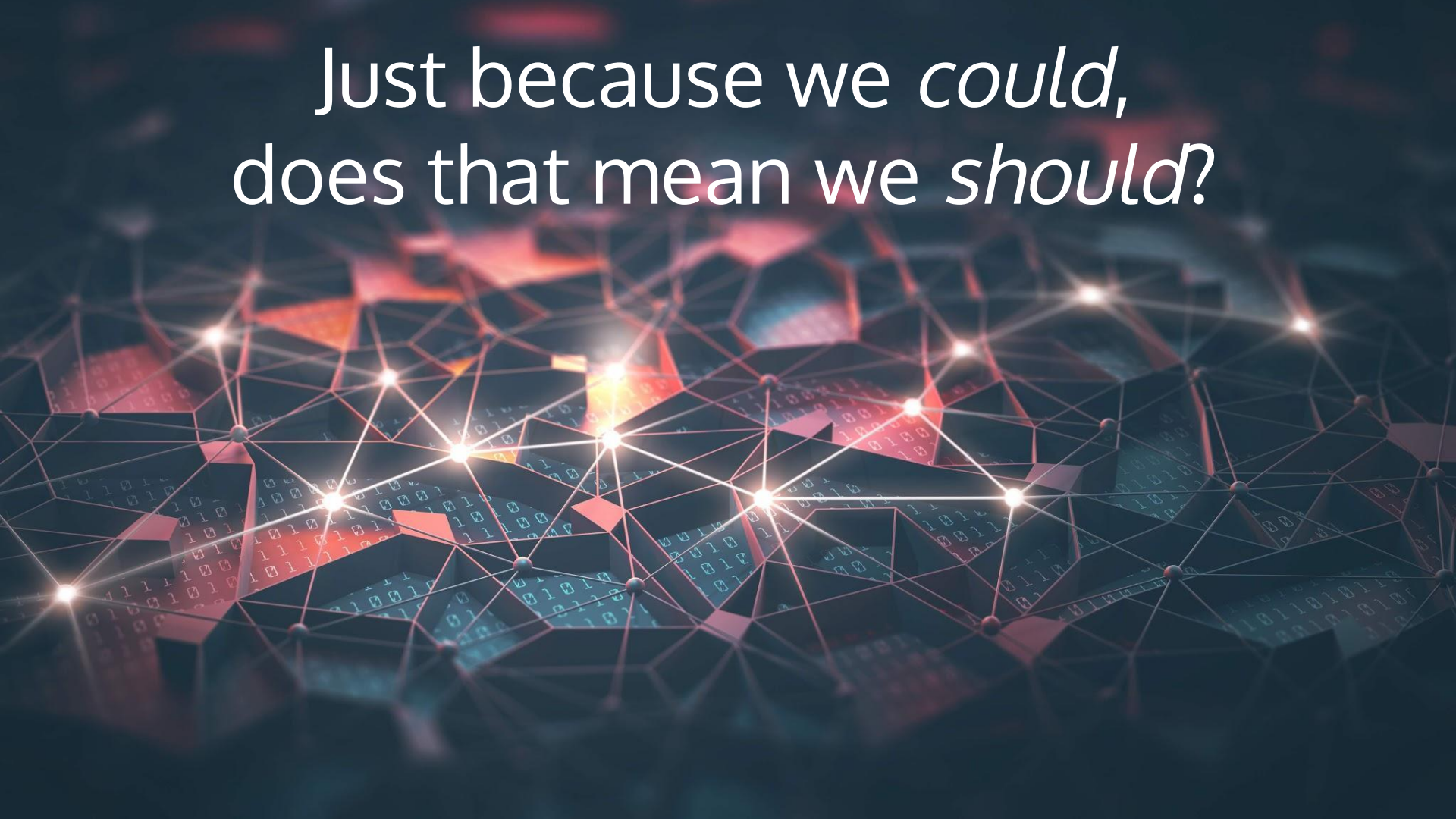
1. Lawful and respectful of our Nation's values
2. Purposeful and performance-driven
3. Accurate, reliable, and effective
4. Safe, secure and resilient
5. Understandable
6. Responsible and traceable
7. Regularly monitored
8. Transparent
9. Accountable

H.R.2231 Algorithmic Accountability Act of 2019

Monitoring Large Accessible Places

Privacy & Security of Personal Information

Human Rights, Unfair Bias, & Discrimination


Just because we *could*,
does that mean we *should*?

Not using people's information illegally is the minimum responsibility for organizations, and it will not be enough to sustain trust, manage risk or be responsive to stakeholder expectations.

- Accenture, 2019

IEEE P7000-2021 Series

Strategic Fit

What: International Standard

Who: Businesses that build or deploy new or revised products/services that include AI/ML.

When: System design, concept, exploration

Read this first.
Educate others.

Ethically Aligned Design (EAD)

Available now.

P7000-2021 Series: Model Process for Addressing Ethical Concerns During System Design

ISO/IEC JTC1/SC42 Artificial Intelligence

Strategic Fit

What: International Standard

Who: Businesses that build or deploy new or revised products/services that include AI/ML.

When: System design, concept, exploration

Published Standards

Standards Under Development

NIST AI Risk Management Framework

Strategic Fit

What: U.S. National Standard

Who:

- Especially businesses with federal contracts.
- Designing, developing, using, deploying, evaluating or governing AI

When: Any stage of the above

Drafting stage; target release in early 2023:

[AI Risk Management Framework]

Others in draft:

[Cybersecurity for IoT]

[Consumer Software Labeling for Cybersecurity]

World Economic Forum

Strategic Fit

What: International Guidance

Who: Organizations and citizens who need broader education and understanding of why ethical AI is important.

When: Anytime. The earlier in product/service lifecycle, the better.

WORLD ECONOMIC FORUM

AI Toolkit for Boards of Directors

AI Procurement in a Box for Government Entities

Ethics by Design

Singapore's Partnership with WEF

Strategic Fit

What: International Guidance; Singapore's Personal Data Protection Commission (PDPC)

Who: Designing, developing, using, deploying, evaluating or governing AI

When: Any stage of the above

WORLD ECONOMIC FORUM

Model Artificial Intelligence Governance Framework

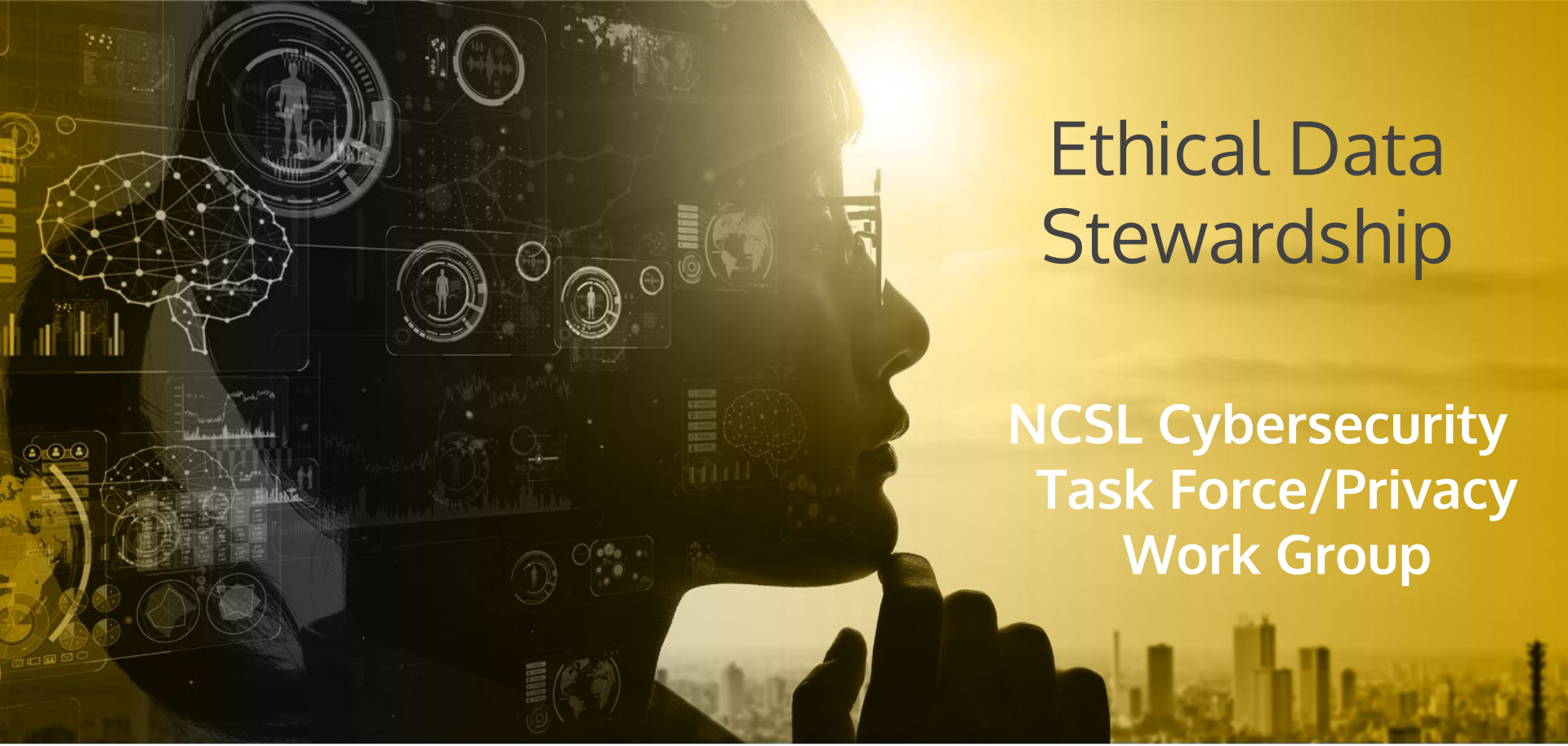
Implementation & Self-Assessment Guide for Organizations

Other Countries

Brazilian Artificial Intelligence Strategy

Pan-Canadian AI Strategy

China, Ethical Norms for the New Generation Artificial Intelligence



Ethical Data Stewardship

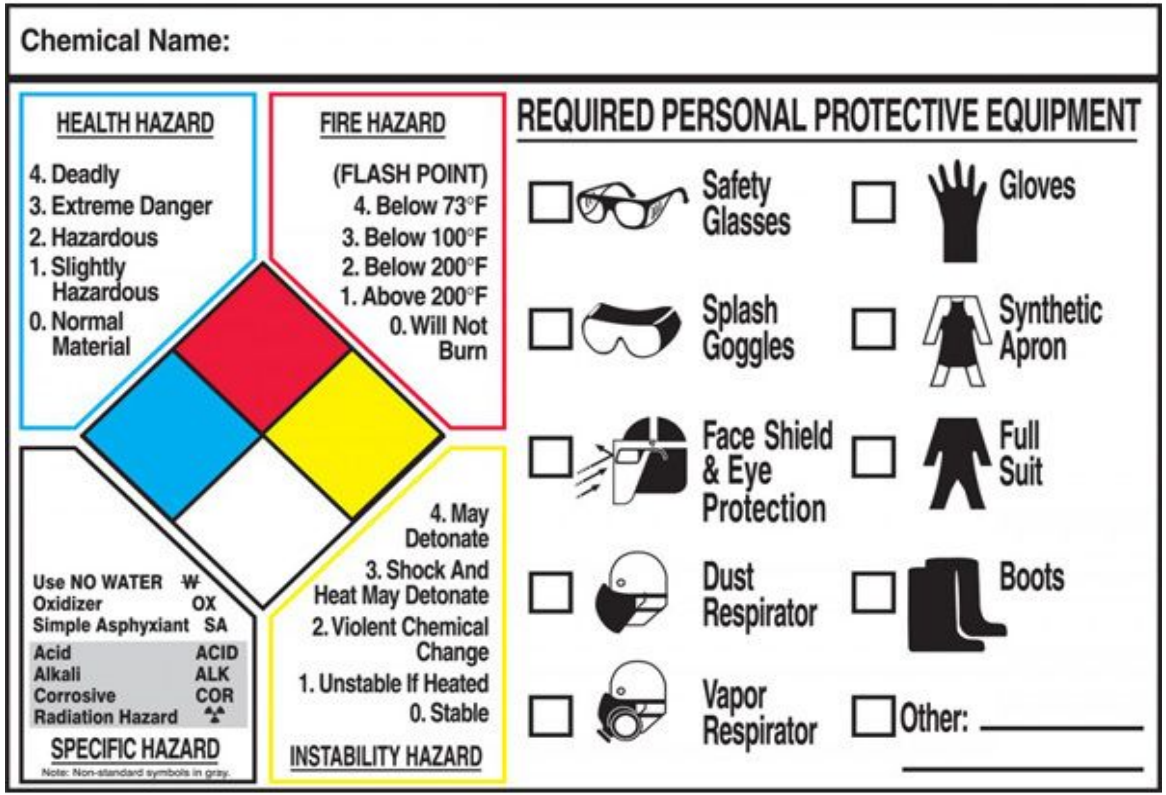
NCSL Cybersecurity Task Force/Privacy Work Group

Michael Best

NCSL
NATIONAL CONFERENCE OF STATE LEGISLATURES

Example: Datasheets for Datasets - Like MSDS for Data!

MATERIAL DATA SAFETY SHEETS



Research Paper: [Datasheets for Datasets]

The Open Ethics Canvas v1.0

Designed For	Designed By	Date	Version

Scope
- What is this product designed for?
- In which context it operates?

Users
- What type of users does this product have? (customers/admins/ etc)
- What are their roles?

Key Stakeholders
- Who are the key stakeholders?
- What influence do they have over the product?
- How do stakeholders interact with each other?
- How is the power distributed?

Values & Interests
- What values do stakeholders/users have?
- Where these values can clash or create tensions?
- What is known at the moment and how assumptions are tested?
- How can you align your technology to the values you want to support/people desire?

Training Data
- How was the training data collected?
- How do you ensure its representativeness?
- Does your training dataset contain personal data?
- Who annotates the data and how quality is controlled?
- What is the data labeling process that you employ?

Personal Data Processing
- Which personal data is collected by the product?
- What is the purpose of collecting personal data?
- How is this data processed? Used? Stored? Deleted?

Explainability
- How is interpretability defined for the system?
- What interpretability methods are used?
- What metrics are used in result interpretation?
- How interpretations of the output are communicated?

Algorithms & Source Code
- Do you use open or proprietary sources
- Who in the team is setting the heuristic influence the output?
- How do you ensure the quality of used t codebases?
- What is your process of making the key choices?

Components & Subprocessi
- Which third parties are engaged by the
- How do you evaluate the potential impa the quality of your product's output?
- How do you check the reliability of your processing contractors?

Human in the Loop (HITL)
- What is the role of a human agent in the validation/verification of the outputs?
- What is the role of a human agent in ref performance?
- What is the decision-making power ass agents responsible for the quality of ou

Impact Assessment
- What potential harms can your product cause? (loss of opportunity, discrimination, economic loss, social stigma, detriment, emotional distress, etc)?
- What are the risks of the product's failure?
- What impact product can cause if deployed at scale?
- How is the product influencing the existing markets?

Regulatory Landscape
- What is the regulatory context in which operates?
- Is the model portable to other market v
- What are the involved regulatory risks?

Changes in Behavior
- Do the automated decisions have significant legal or similar effects on the users/stakeholders?
- How the users may change their behavior after use?
- What are the potentials for power imbalance?

Group Interactions
- What are potential changes in group behavior?
- How is the product addressing group interests?
- What new groups could be born due to the product deployment at scale?

Comments

Open Ethics