IEEE SA
STANDARDS ASSOCIATION

KNOW THYSELF AND CHOOSE YOUR AI PARTNERS

SEPTEMBER 2022

Dr Clara Neppel
Senior Director European Operations

IEEE

IEEE SA
STANDARDS ASSOCIATION

ADVANCING TECHNOLOGY FOR HUMANITY

ABOUT IEEE

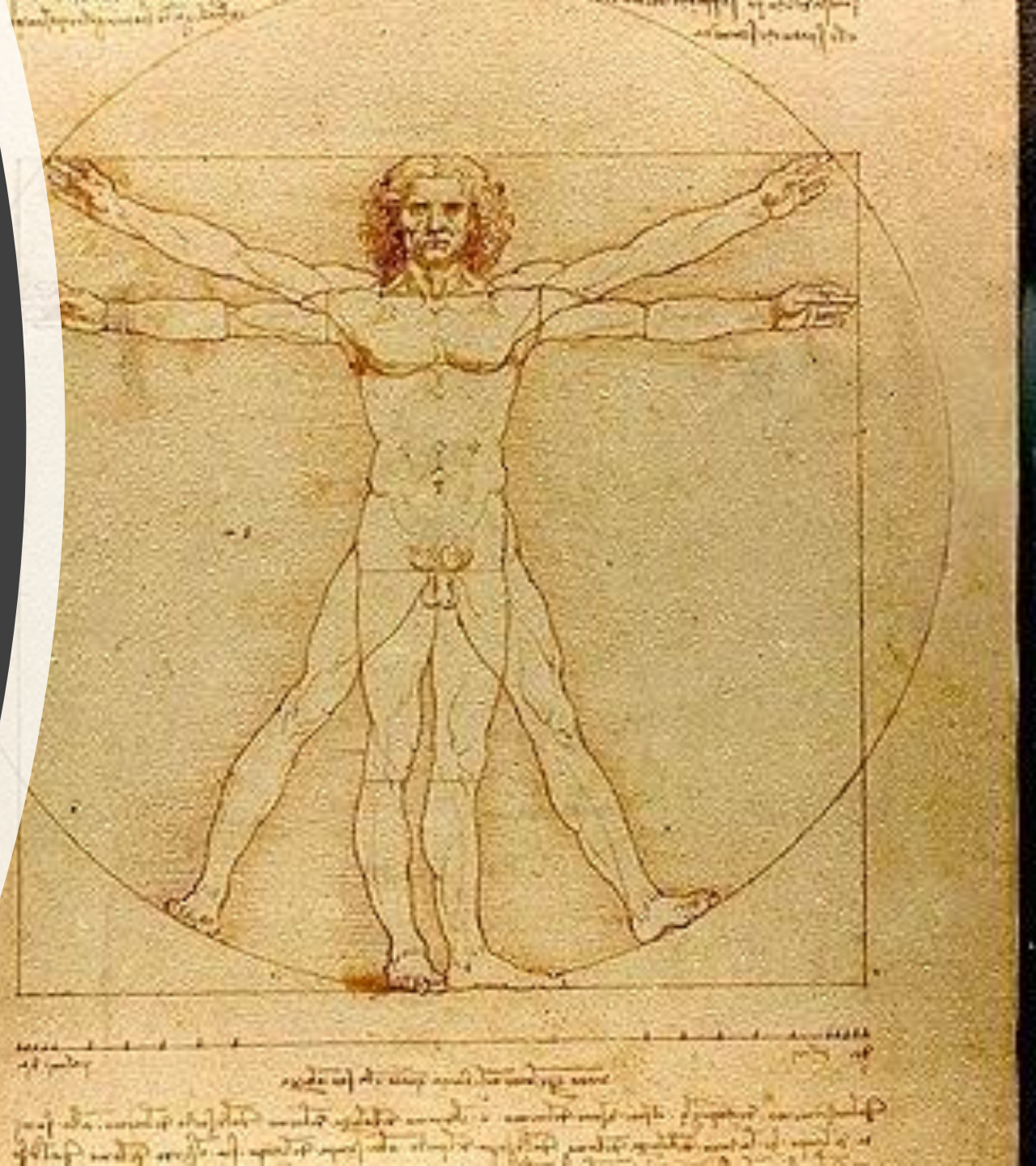
- 400,000 Members
- 160 Countries
- 46 Technical Societies and Councils
- 1900+ Annual Conferences
- Global Humanitarian Efforts
- Developing market relevant open standards and solutions

IEEE

Digital Humanism

“Humanism is a philosophical stance that emphasizes the potential and agency of human beings, individually and socially”

(Wikipedia)



Breakthroughs

Ethical challenges

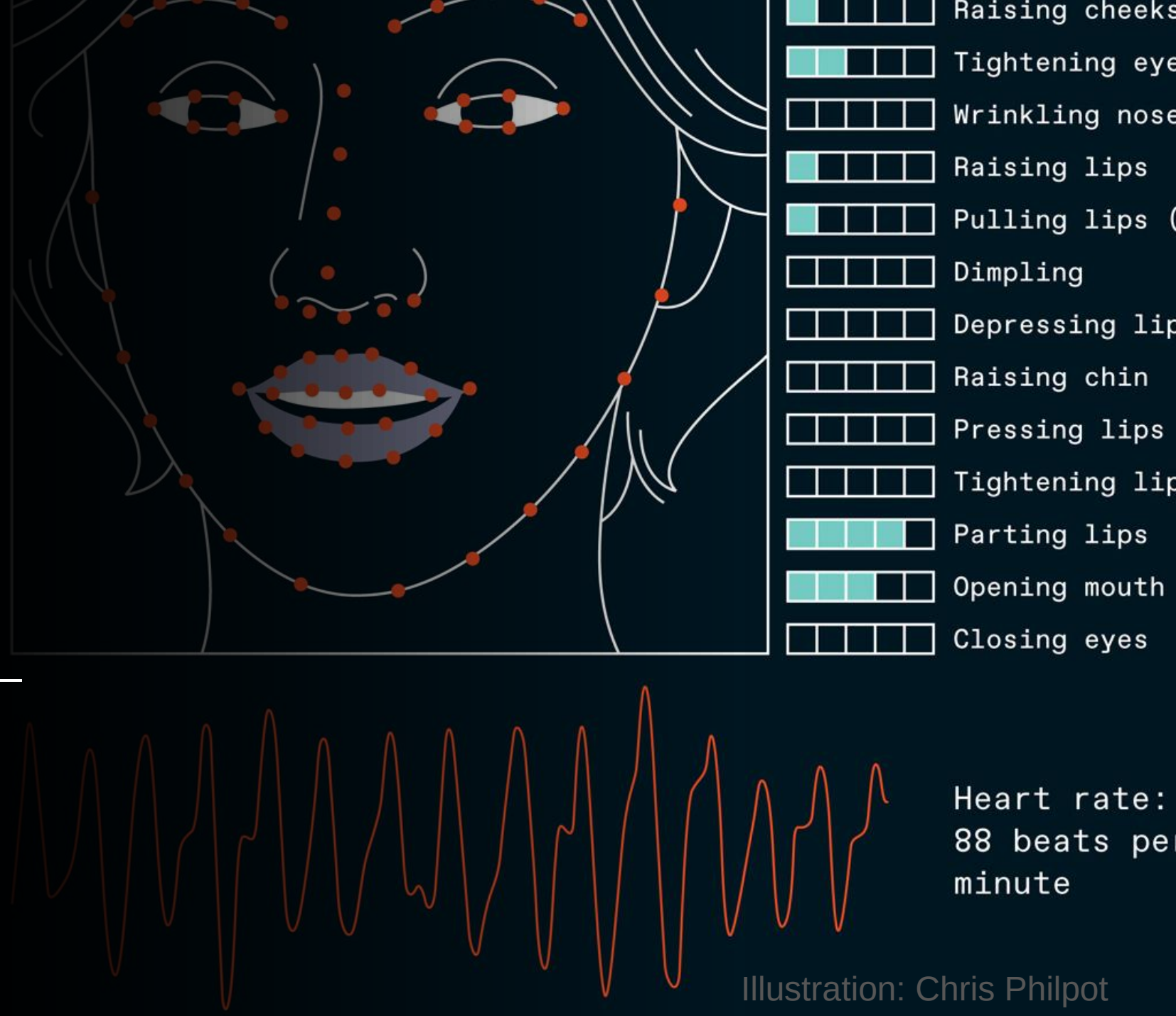


Illustration: Chris Philpot

IEEE Spectrum queried DALL-E 2 for an image of "a technology journalist writing an article about a new AI system that can create remarkable and strange images."

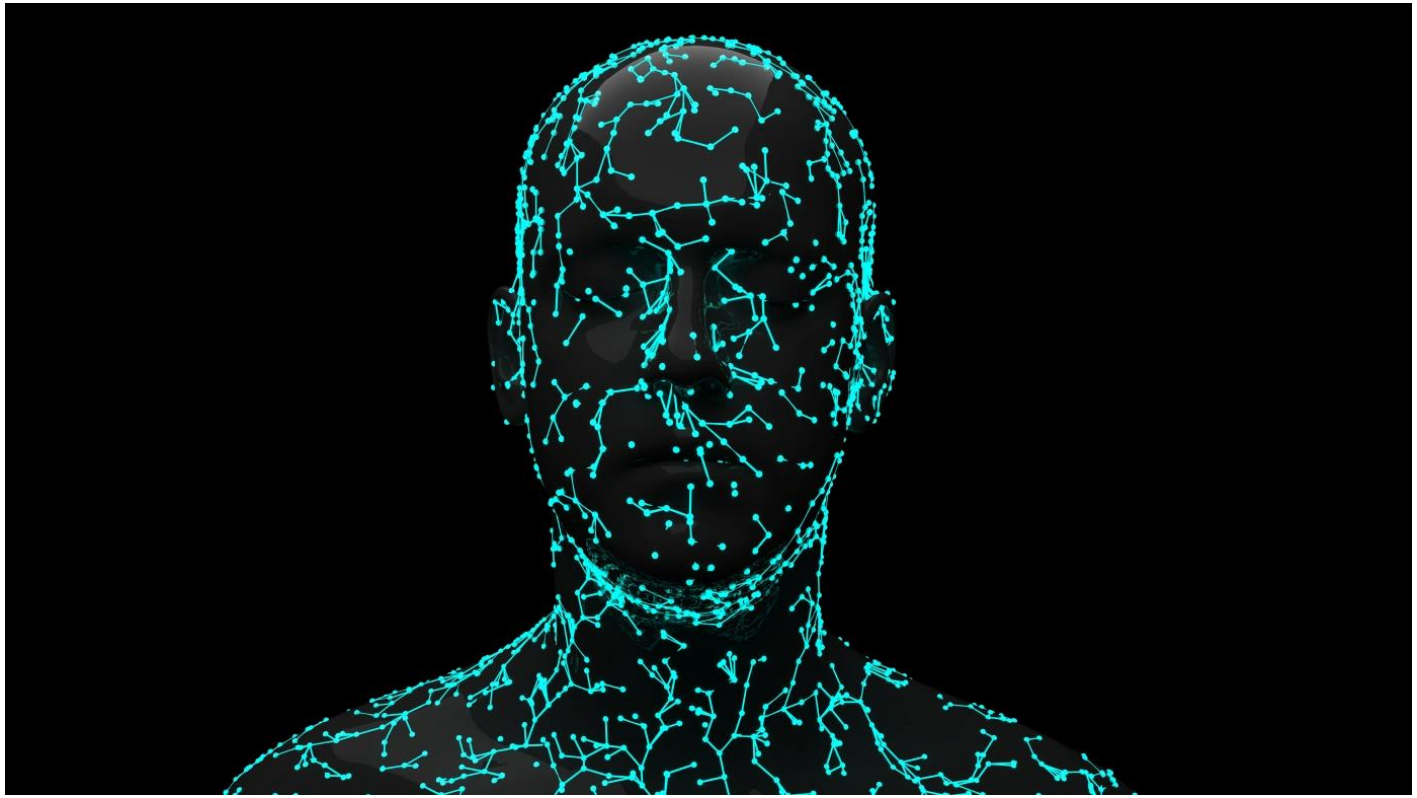
“Metaverse”

Growing Number of Extended Identities and Realities

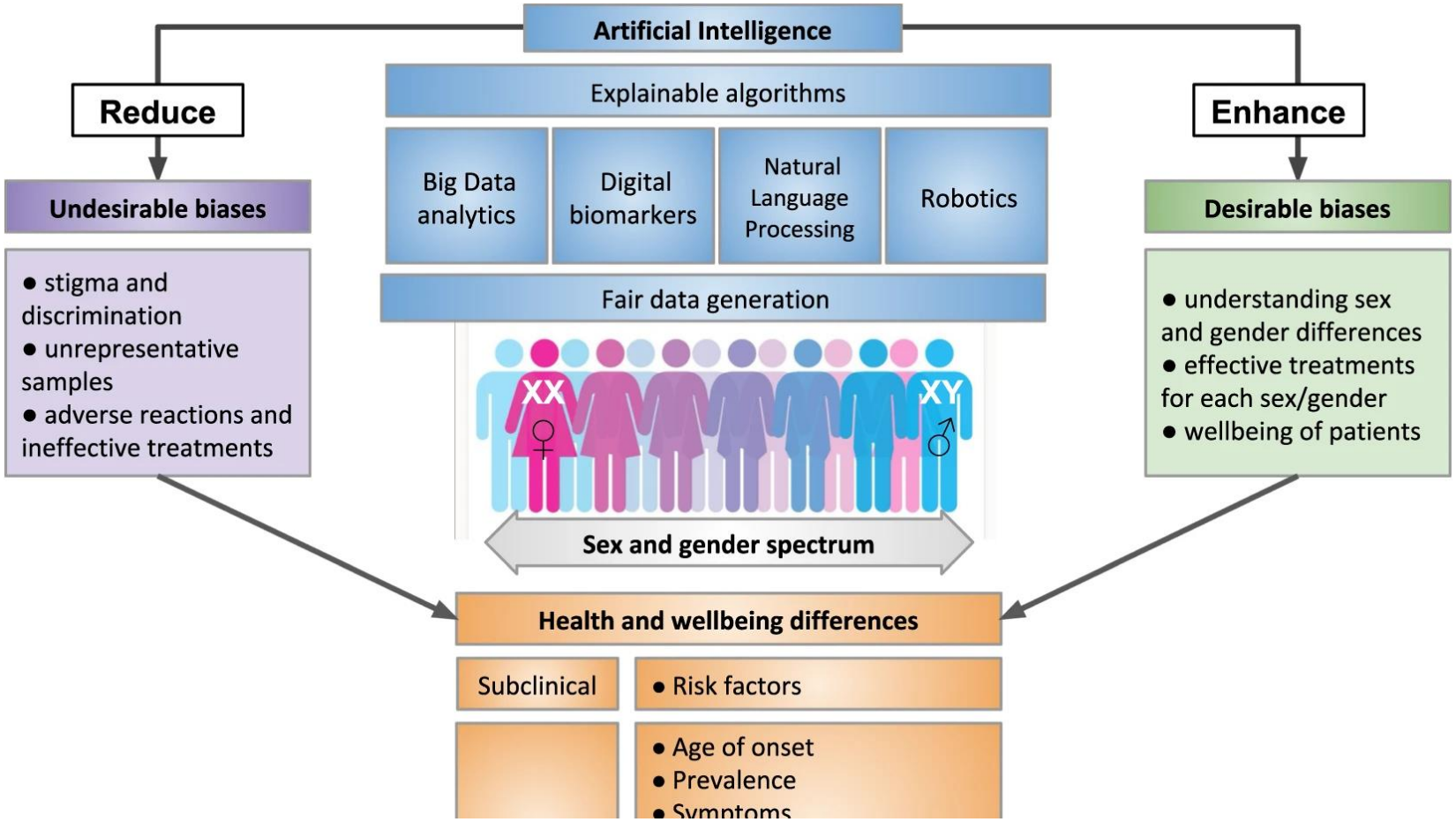
”agency of human beings, individually and socially” ?

Clearview AI banned from selling its facial recognition software to most US companies

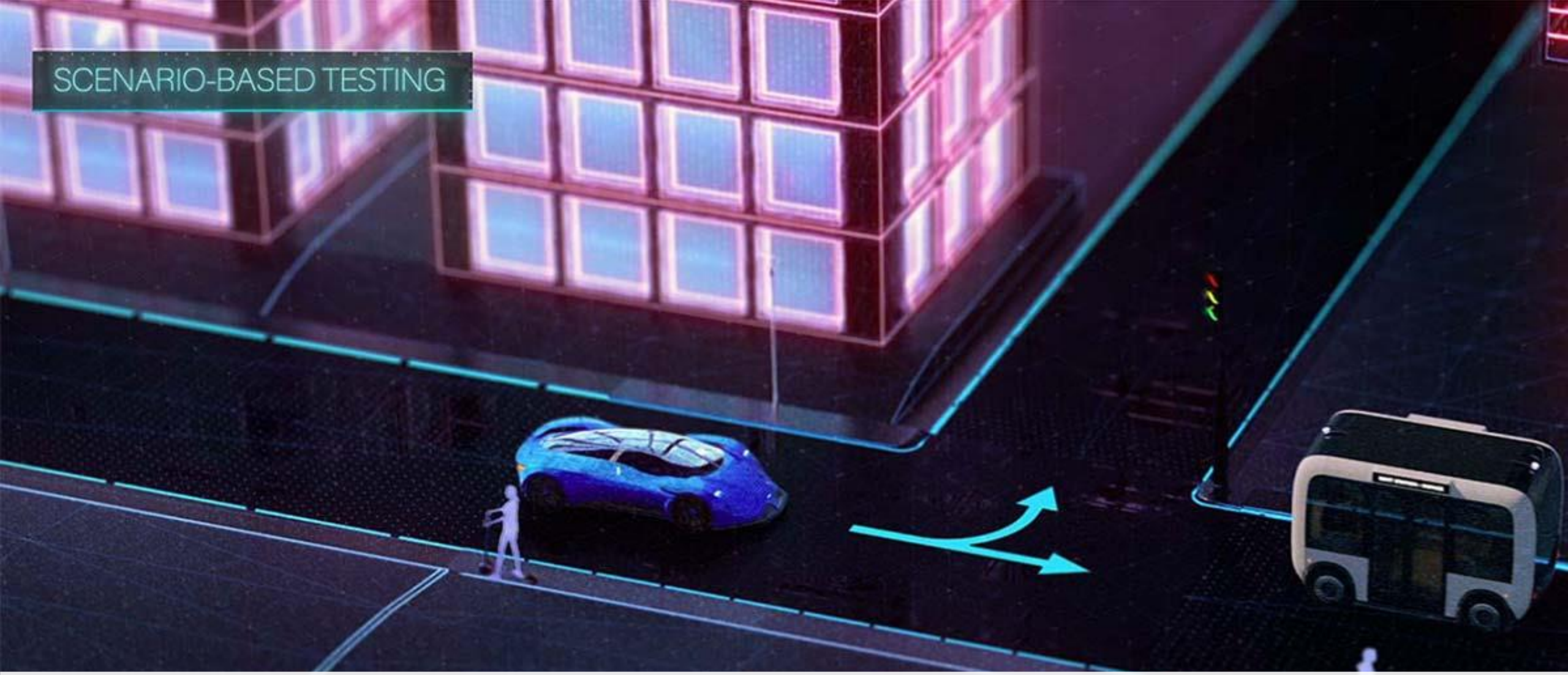
Taylor Hatmaker@tayhatmaker / 9:05 PM GMT+2•May 9, 2022



Healthcare



https://www.nature.com/articles/s41746-020-0288-5



Mobility

Fair Synthetic Data ?

@NVIDIA StyleGAN

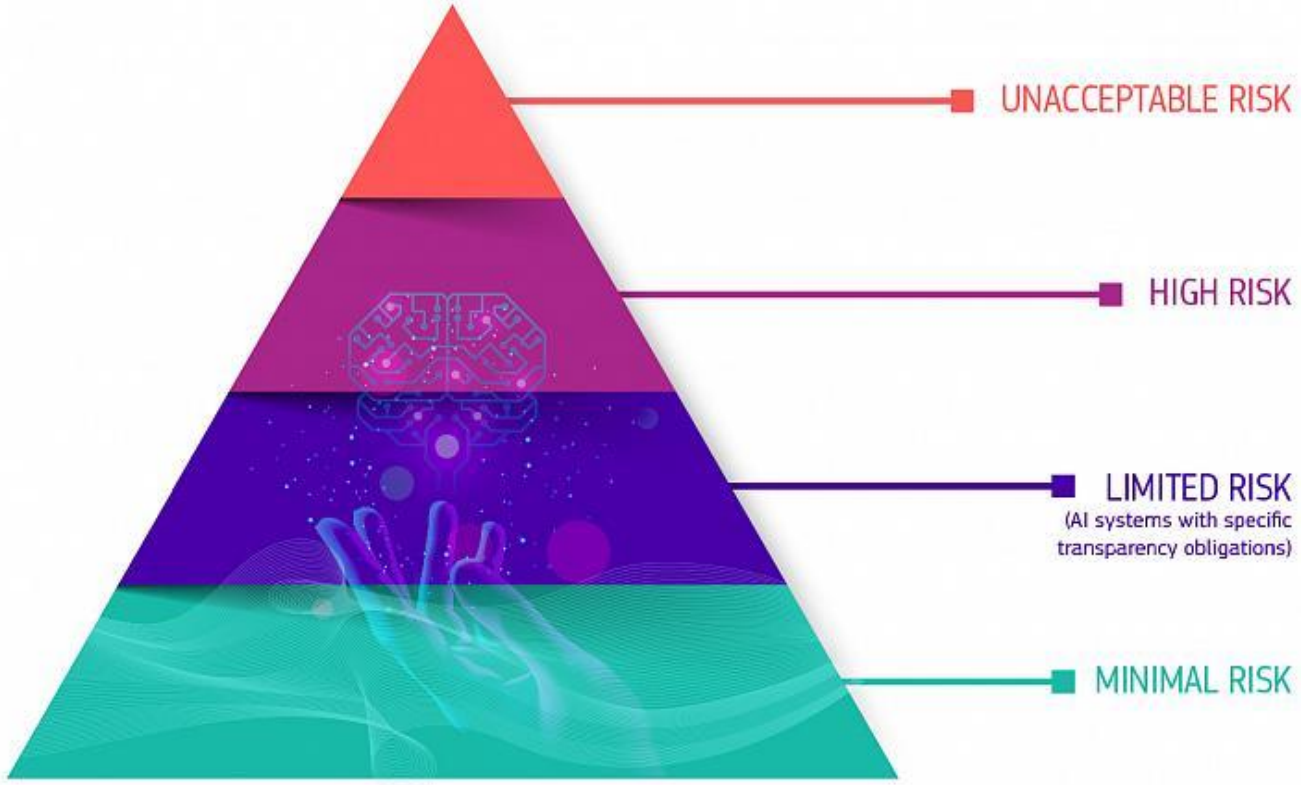
Safety/Security or Ethical Risks

Clear distinction possible?

TRUST

Redefining Risk

Proposal for EU AI Act



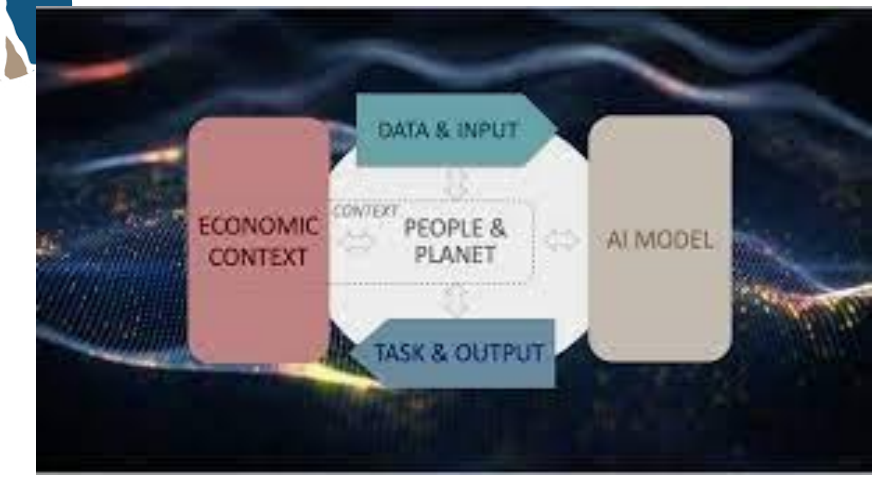
New Metrics of Success

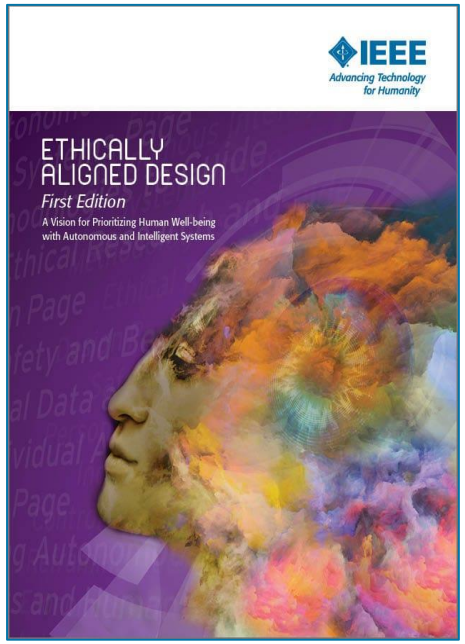
For socio-technical systems

IEEE SA
STANDARDS ASSOCIATION

From Principles To Practice

OECD

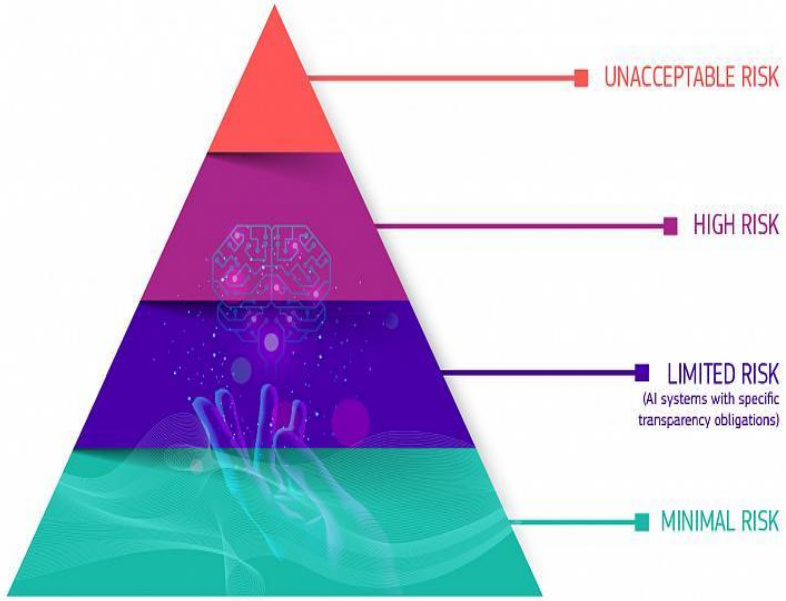




GPAI / THE GLOBAL PARTNERSHIP ON ARTIFICIAL INTELLIGENCE

IEEE

Proposed EU AI ACT



"Mandatory requirements applicable to the design and development of certain AI systems before they are placed on the market that will be further operationalised through harmonised technical standards"

High Risk - horizontal mandatory requirements and conformity assessment

Limited Risk - Transparency obligations

Minimal Risk - Voluntary codes of conduct

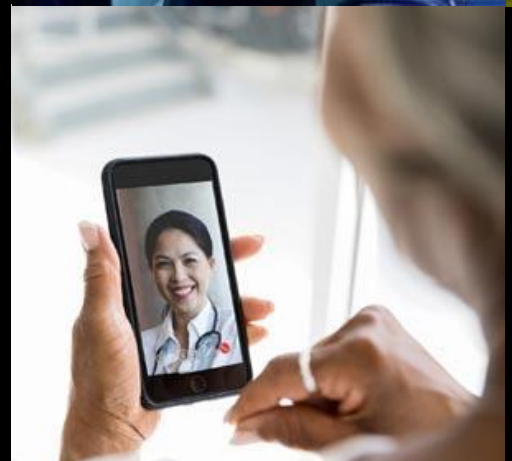
Harmonized standards to bring the necessary level of detail

- Data and Data Governance
- Risk Management
- Quality Management Systems
- Accuracy, Robustness & Cybersecurity
- Human Oversight
- Record Keeping
- Transparency & Provision of Information to Users

IEEE SA
STANDARDS ASSOCIATION

A STANDARD IS . . .

- An agreed way of doing something, such as making a product, managing a process, delivering a service or supplying materials
- Usually, a formal document that established uniform engineering or technical criteria, methods, process and practices
- Result from the collective work by experts in a field and provide a consensus at the time when the standards were developed
- Have a prominent role in our daily lives—they are the unforeseen force behind innovation



IEEE

Transparency and Interoperability

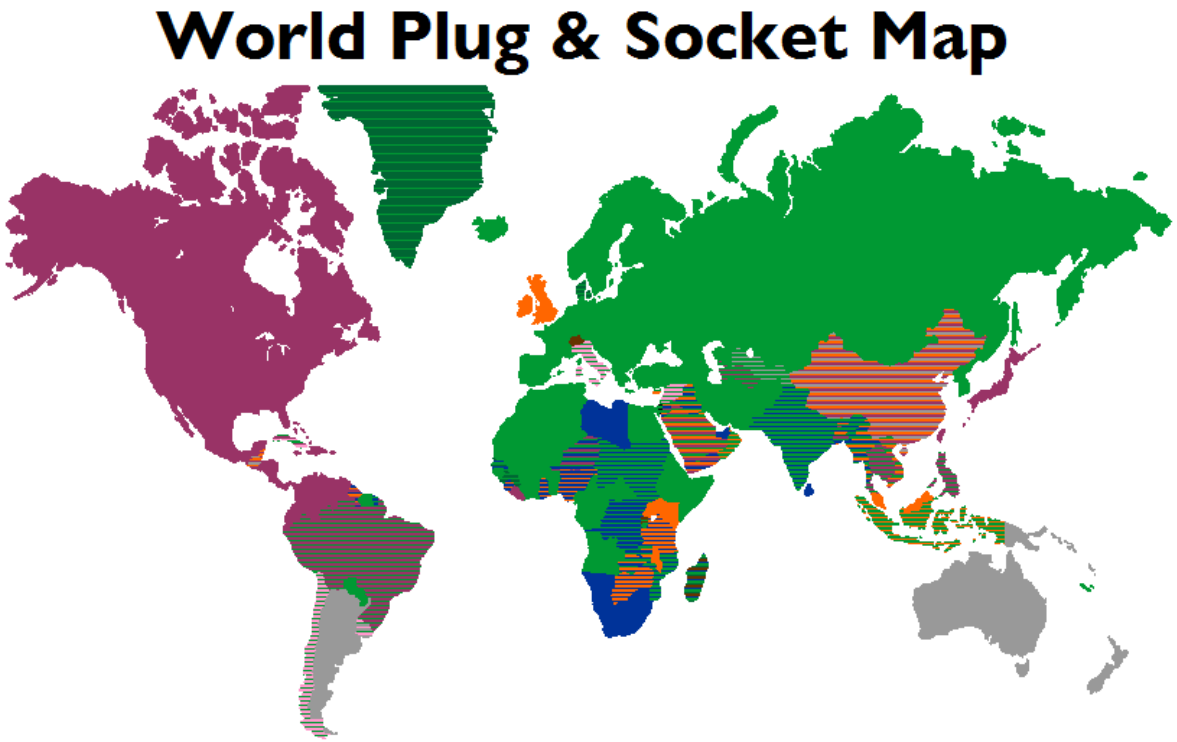
EU member states agree specs for coronavirus app interoperability

Single Standard vs Many Standards

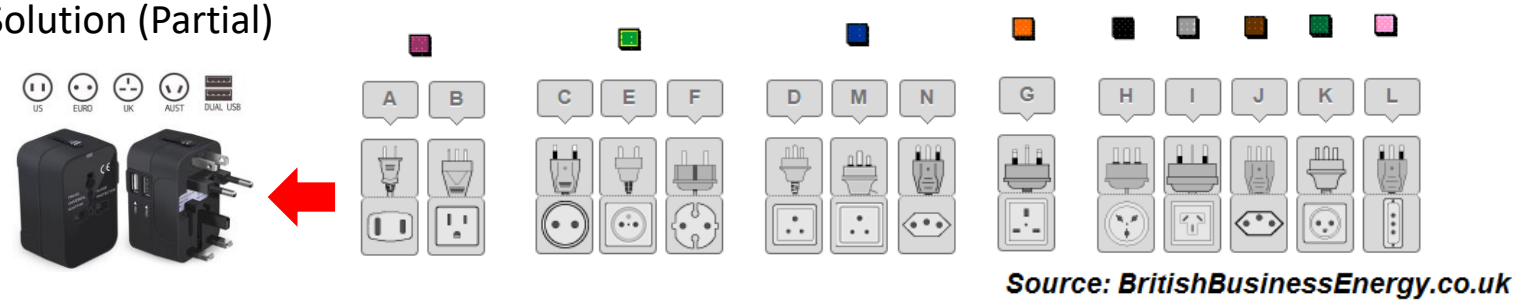
Single Universal Standard

IEEE
802.11

Many Standards



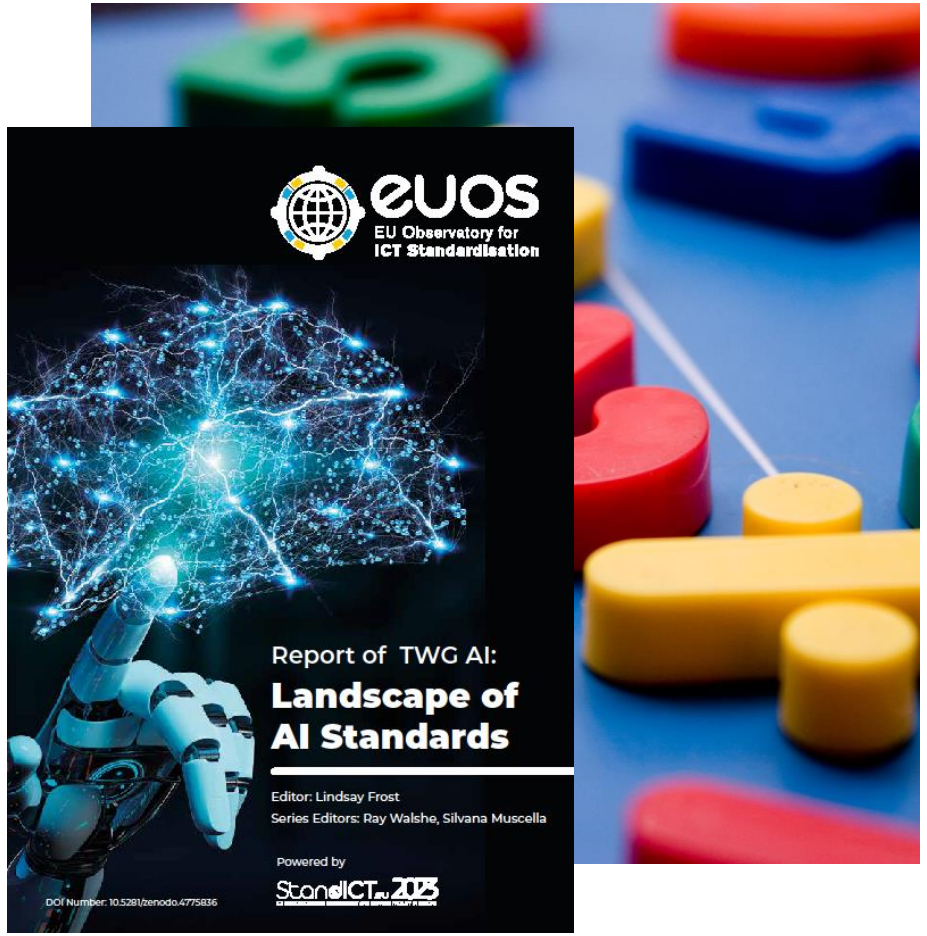
Solution (Partial)



Source: BritishBusinessEnergy.co.uk

State of play in the EU AI regulatory process

- Negotiation
- EC prepares standardization mandate to the ESOs (CEN, CENELEC, ETSI)
  - Uncontroversial elements first
  - Update to reflect outcome of negotiations
- Preliminary roadmap definition activities by ESOs
- Able to leverage existing standards and technical specifications from international SDOs
  - Avoid duplication of work
  - Speed
  - Global reach
- Technical analysis of existing initiatives, such as
  - ISO/IEC JTC1 SC42
  - IEEE



IEEE SA
STANDARDS ASSOCIATION

PROGRAMS

Make the implicit explicit and verifiable for the entire AI ecosystem

STANDARDS

Developing a growing series of standards that promote innovation, foster interoperability and honor human values.

- **Technical**
- **Socio-technical**

CERTIFICATION

Developing metrics and processes towards the implementation of a certification methodology.

- **Transparency**
- **Accountability**
- **Algorithmic bias**
- **Respect of Privacy**
- **Responsible Governance**

GOVERNANCE

Support responsible Artificial Intelligence Systems innovation through proper governance mechanisms for:

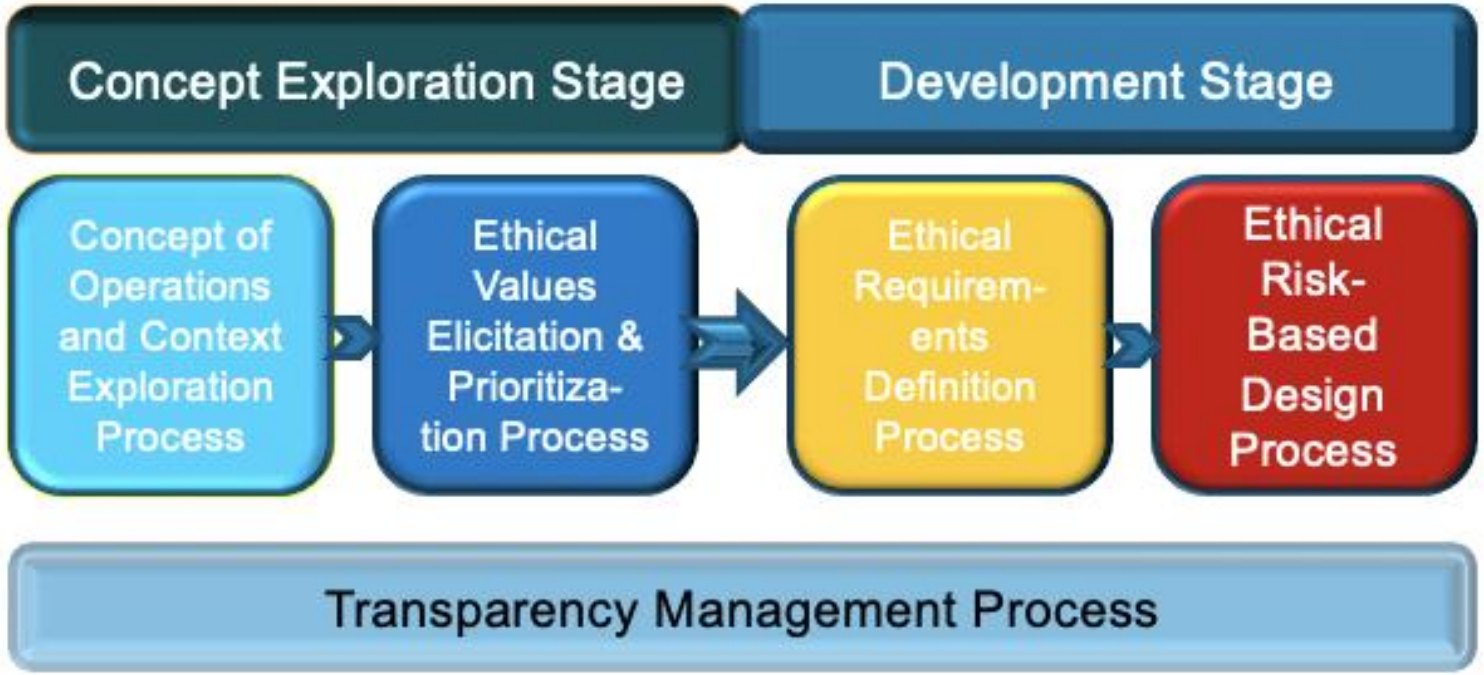
- **Business**
- **Municipalities/Cities**
- **Public Sector/ Governments**

https://standards.ieee.org/initiatives/artificial-intelligence-systems/

IEEE

Risk Assessment and Mitigation starts already at the System Design Phase

IEEE 7000



The protection of Union values and fundamental rights in the AI Act implies the scrutiny of a vast scope of potential context-specific value breaches.

EU Charter of Fundamental Rights

- Human Dignity (Article 1)
- Respect for private life (Articles 7)
- Protection of personal data (Article 8)
- Non-discrimination (Article 21)
- Equality between women and men (Article 23)
- Freedom of expression (Article 11)
- Freedom of assembly (Article 12)
- Presumption of innocence (Articles 47 and 48),
- Rights to fair and just working conditions (Article 31),
- A high level of consumer protection (Article 28),
- The rights of the child (Article 24)
- …

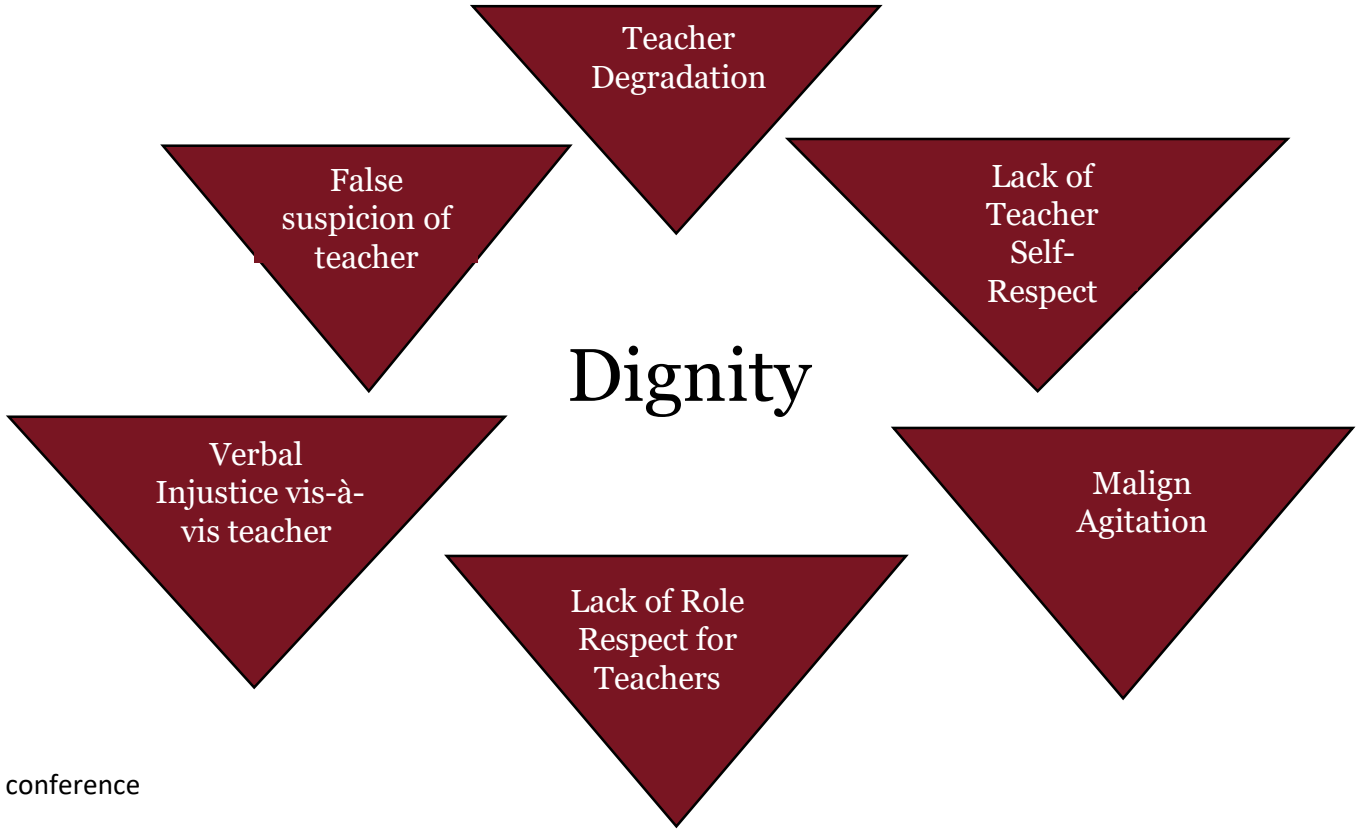
Union Values and HLEG values

- Human Agency & Oversight/Control
- Technical Robustness & Safety
- Privacy & Security
- Diversity, Fairness & Nondiscrimination
- Societal & Environmental Wellbeing
- Openness & Transparency
- Health
- Trust
- Dignity
- Accessibility
- Accountability/Responsibility

One IEEE 7000™ case study showed that a system like a simple teacher-rating app has six different ways in which it can breach teacher dignity.

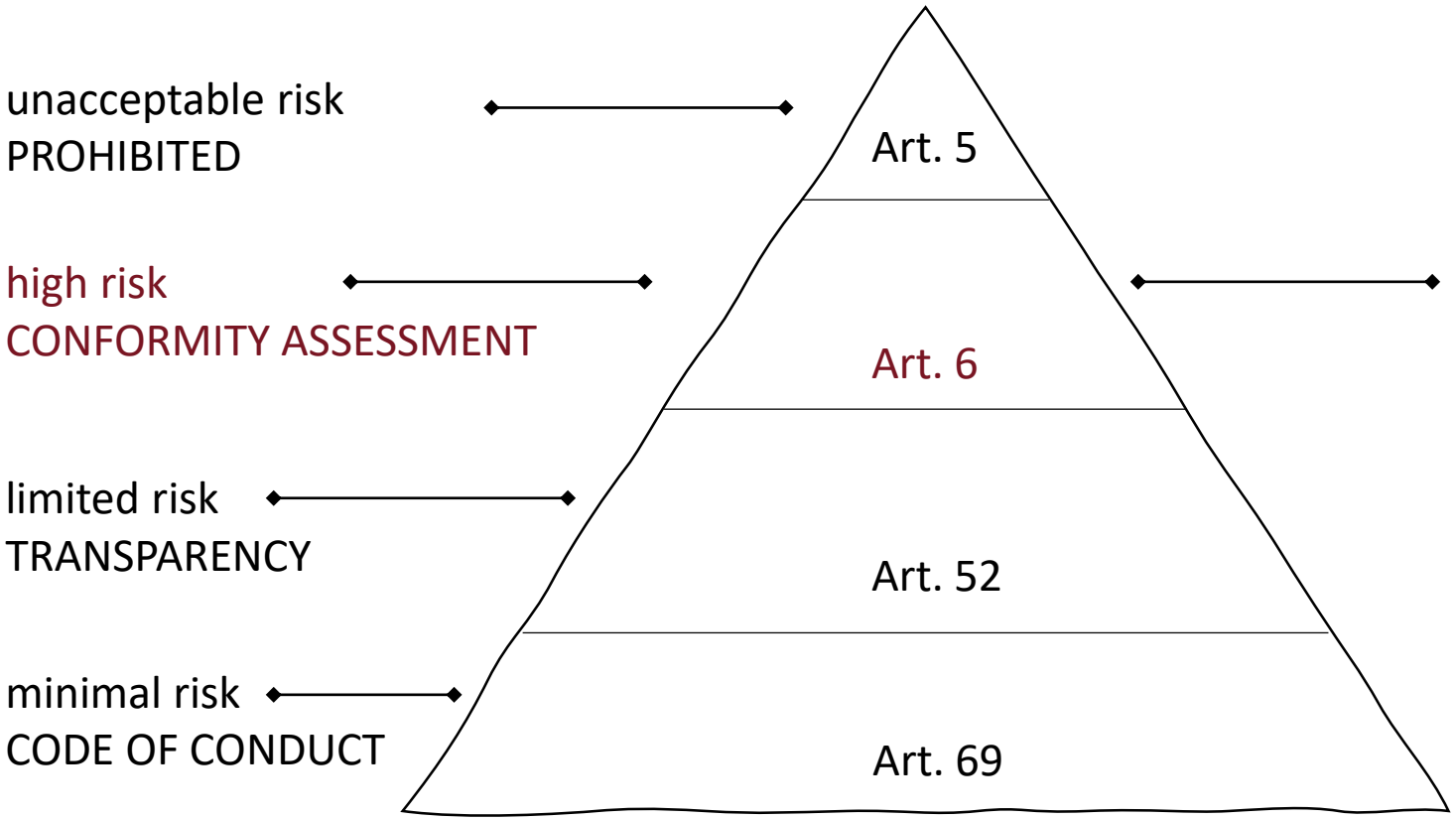
Example: Complexity of the value of **Dignity**

Lernsieg –
Austrian Teacher Rating App



Case study presented by Prof Sarah Spiekermann at the CPDP 2022 conference

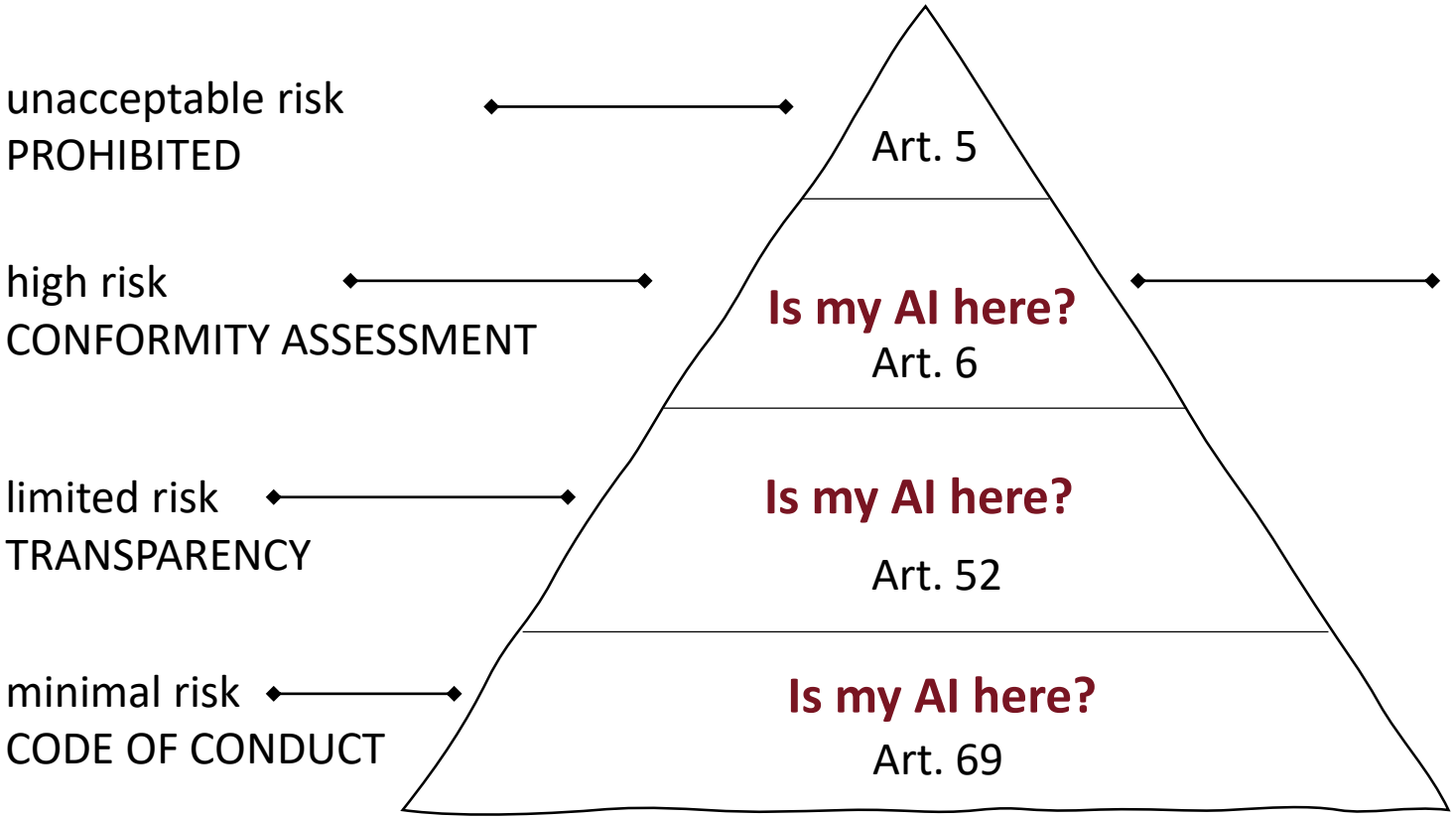
The draft AI Act in Europe has developed a first view on what high risk systems are.*



- Biometric identification and categorization of natural persons
- Management and operation of critical infrastructure (such as supply of water, gas, heating and electricity)
- Education and vocational training
- Employment, workers' management and access to self-employment
- Access to/enjoyment of essential private services and public services and benefits (like credit and emergency first response services)
- Law enforcement
- Migration, asylum and border control management
- Administration of justice and democratic processes

*illustration taken form LAWFARE, "AI Act: What is the European Approach for AI?", Eve Gaumond, June 4th 2021

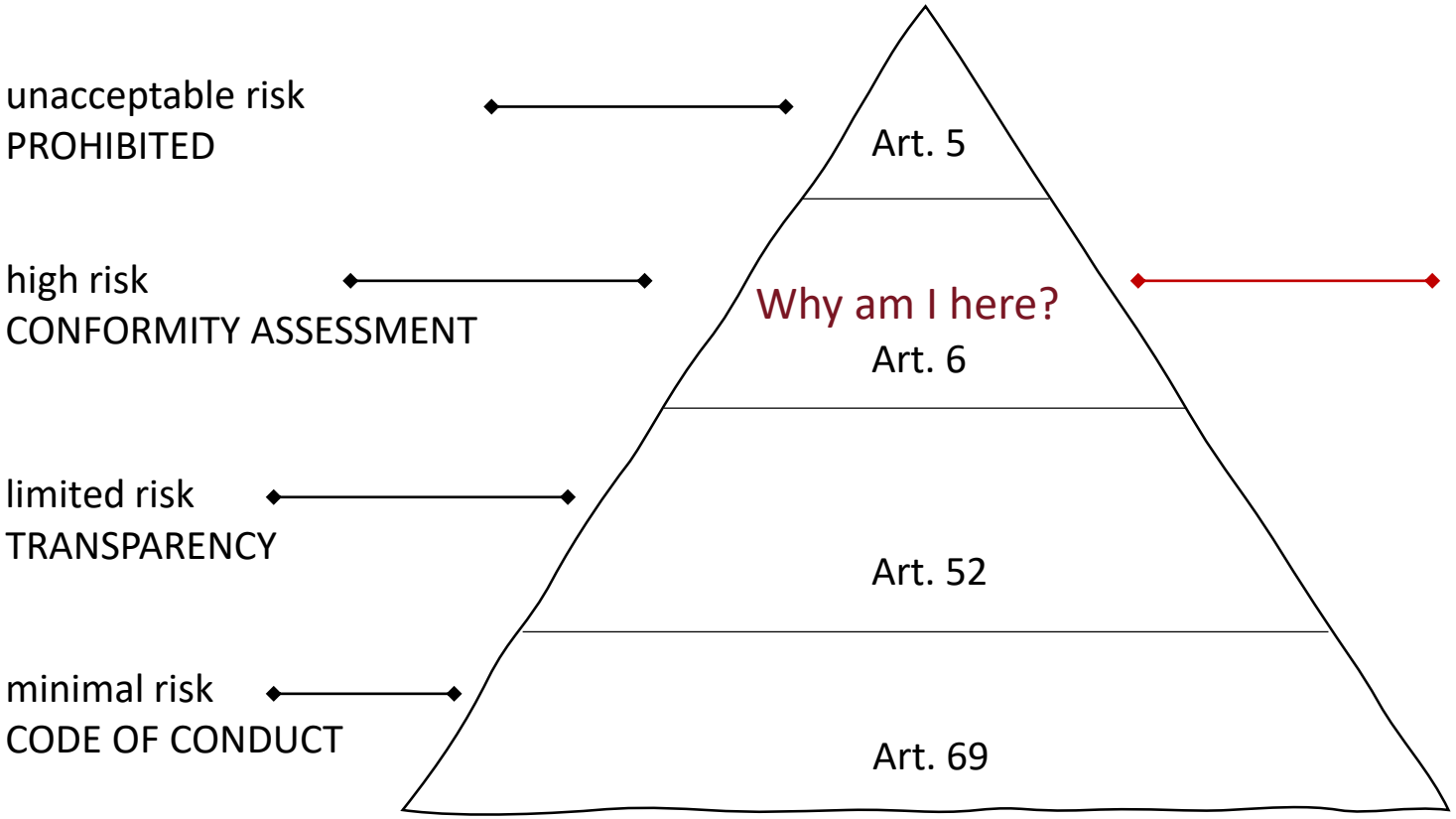
Prior to launch, operators need to understand whether their applications may bear high risks.



- NOT ALL APPLICATIONS IN THESES SECTORS ARE HIGH RISK
- HIGH RISK TO FUNDAMENTAL RIGHTS CAN OCCUR IN OTHER AI OPERATING ENVIRONMENTS AS WELL

*illustration taken form LAWFARE, “AI Act: What is the European Approach for AI?”, Eve Gaumond, June 4th 2021

Prior to launch IEEE 7000 provides clarity as to the risk level they are potentially in and why

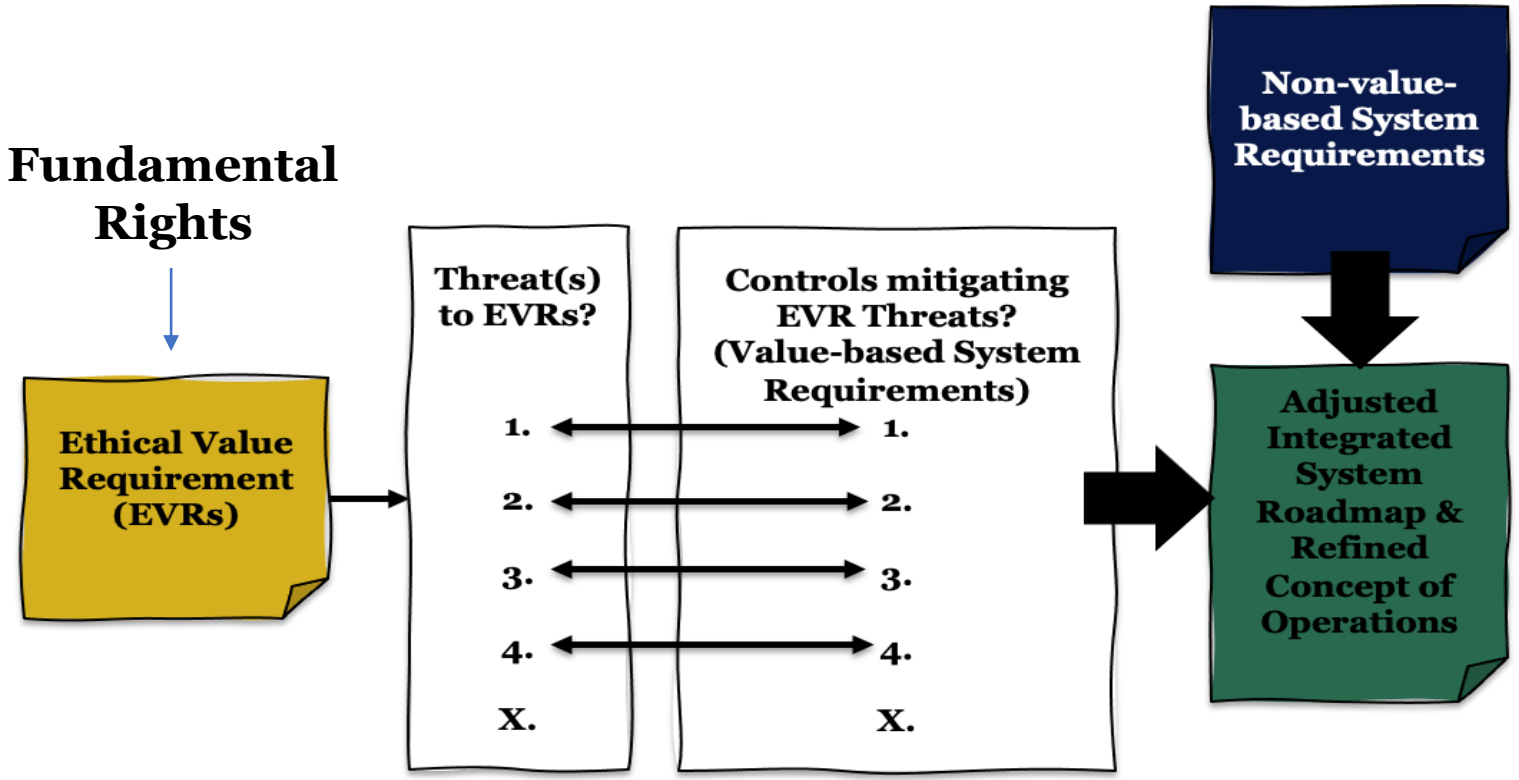


- IEEE 7000 runs companies through an ethical assessement that allows them to understand:

  >> what values (and fundamental rights!) are potentially impacted by their system
  >> why and how the rights might be impacted

*illustration taken form LAWFARE, “AI Act: What is the European Approach for AI?”, Eve Gaumond, June 4th 2021

IEEE 7000 (ISO/IEC/IEEE 24748) gives companies a process for assessing and addressing these risks



In 2020/2021 Value-based Engineering with IEEE 7000TM was tested with UNICEF's Yoma

The Yoma Story

Yoma went from an AI-driven talent calculation machine to a community platform for mutual and local support of African youth

Before (very first idea November 2019):

AI focus

- Young peoples' data is combined and aggregated to calculate individuals' 'talent scores' with an AI engine hosted in Germany
- Young people are represented through its AI-based talent score and homogenized, comparable profiles are created
- All data providers can pull talent scores from young people.
- "Diamonds in the rough" are contacted to be channeled into innovation hubs to support African business

After (summer 2021):

Bottom-up Youth support

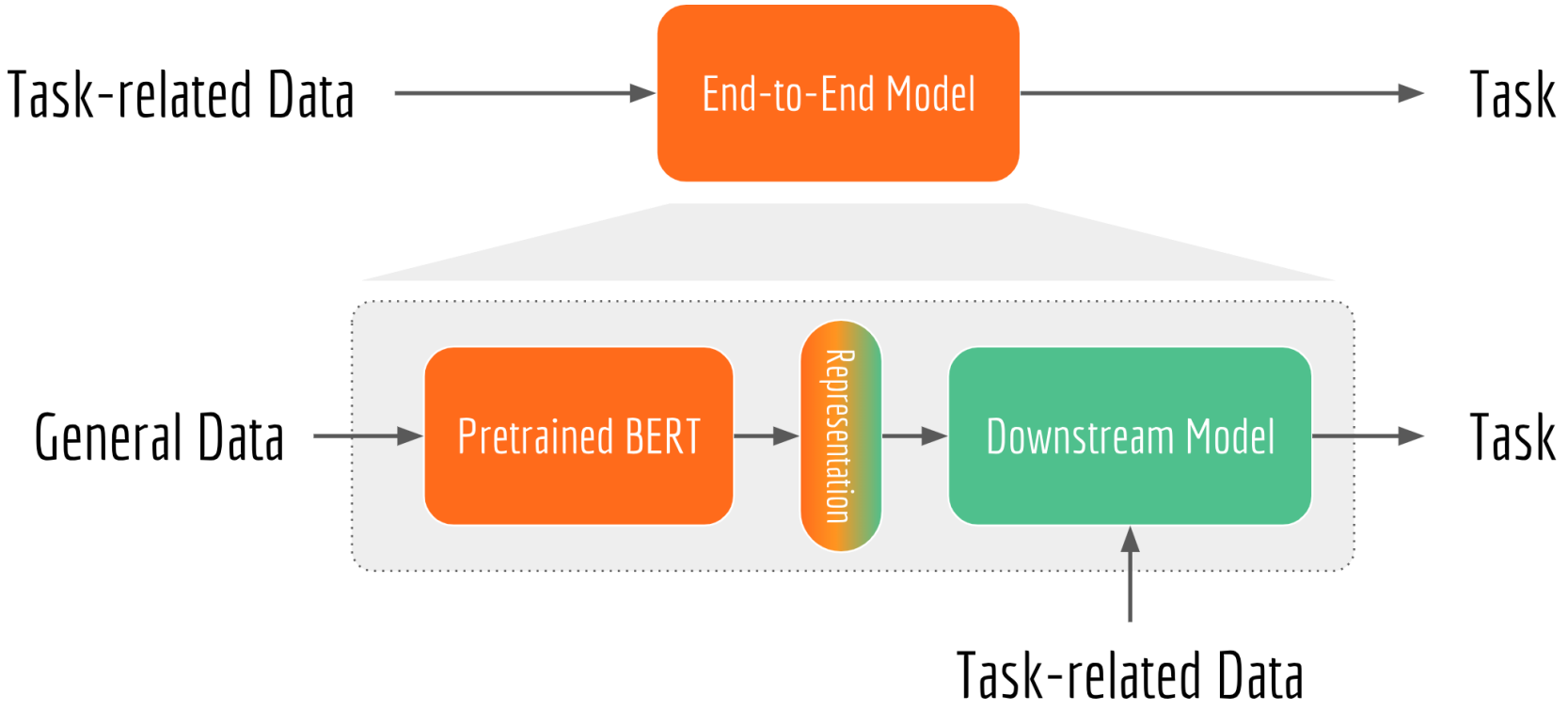
- Online marketplace with opportunities to develop skills, engage with a community and access employment
- Achievements and personal growth are verified using blockchain and added to a digital CV, which they can share with peers and employers.
- This allows youth to build an alternative trust profile, which enhances employability and allows for more informed matching with the labour market.
- Young people have maximum privacy and control and agency
- "Diamonds in the rough" become mentors for other young people to provide local community support AND can respond to African businesses if they want

Dr. Sarah Spiekermann & Kathrin Bednar, CC BY-NC-ND 4.0

OTHER EXAMPLES OF SOCIO-TECHNICAL STANDARDS

- Model Process for Addressing Ethical Concerns During System Design
- Transparency of Autonomous Systems
- Algorithmic Bias Considerations
- Ontological Standard for Ethically Driven Robotics
- Architectural Framework and Application of Federated Machine Learning
- Standard for the Performance and Safety Evaluation of Artificial Intelligence Based Medical Device: Terminology
- Recommended Practice for Organizational Governance of AI
- Standard for Responsible AI Licensing
- Draft Standard for Age Appropriate Digital Services Framework
- Synthetic Data Quality

Ecosystem Control



Find the right partners

- An acknowledged or directed form of partnership is recommended
- For an external AI service to be integrable, the IEEE 7000™ standard outlines that the necessary control over an AI system should only be presumed where there is control
  - The quality of the data used in the AI system;
  - The selection processes feeding the AI;
  - The algorithm design;
  - The evolution of the AI's logic; and
  - The best available techniques (BATs) for a sufficient level of transparency of how the AI is learning and reaching its conclusions.

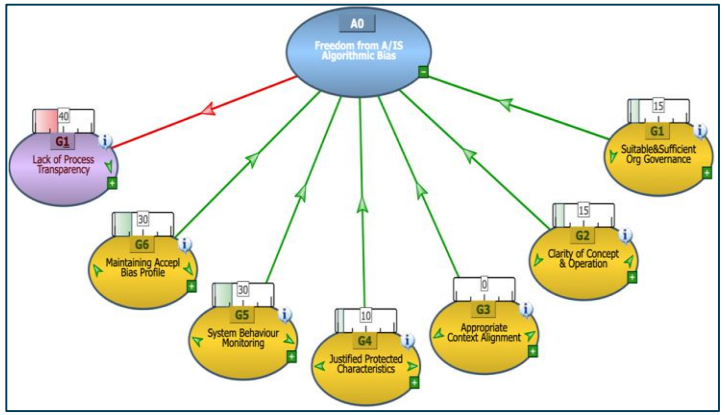
Ethical aligned design is necessary

But not sufficient for AI systems

IEEE CertifAIEd

The Mark of AI Ethics

- **Certification is a basis for trust in products and services**
- **Independent Scrutiny**
  - ✓ Identifies Strengths and Shortfalls
  - ✓ Demonstrates Duty of Care
- **Adaptive and agile**
  - ✓ Assessment effort proportionate to the risk profile
  - ✓ Rapid tailoring to the needs of a sector



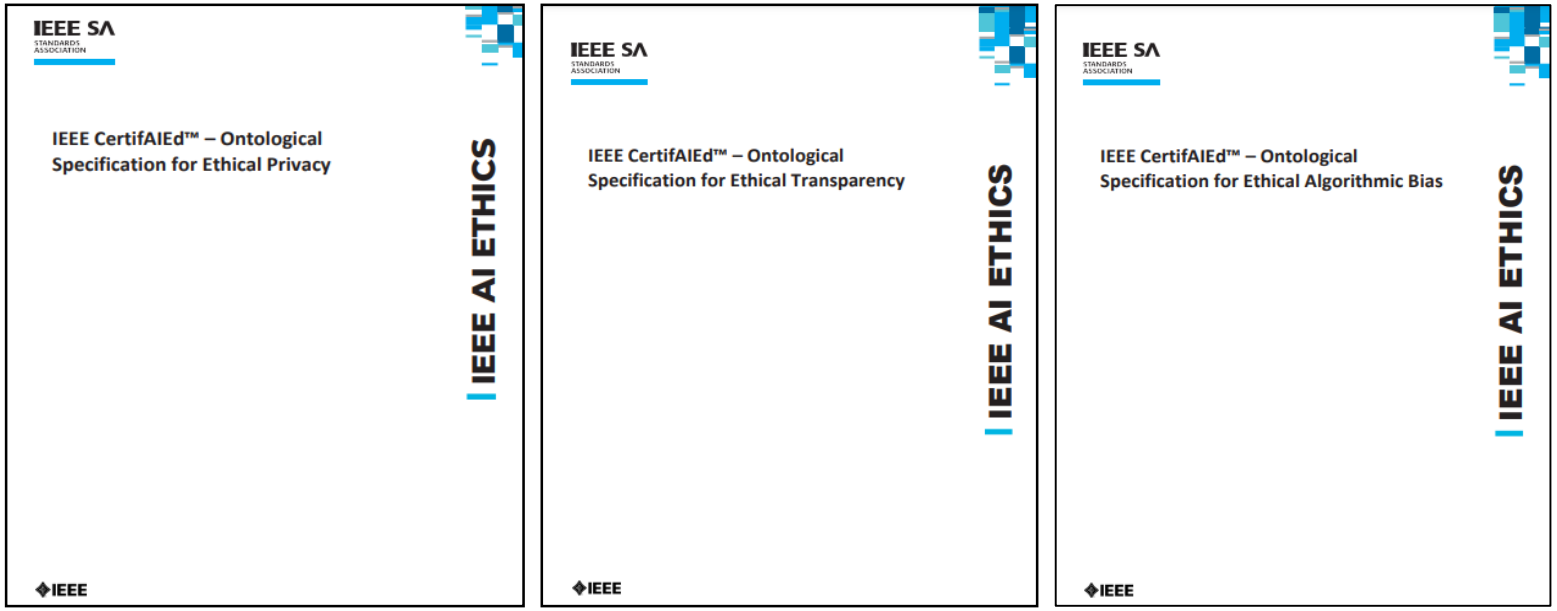
- **Demonstrate implementation towards the proposed EU AI regulation**

City of Vienna Earns IEEE AI Ethics Certification Mark

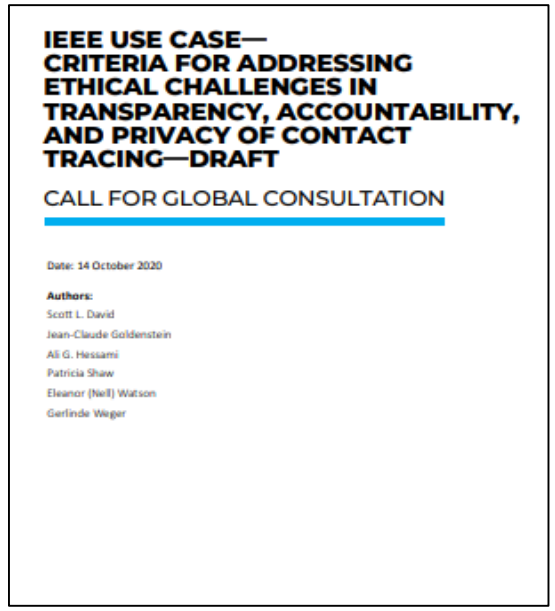
IEEE CertifAIEd

Released base level criteria subsets available for public access and improved understanding. Below are snapshots of several of the referenced level releases available on our site.

Adaptive



Contextualized



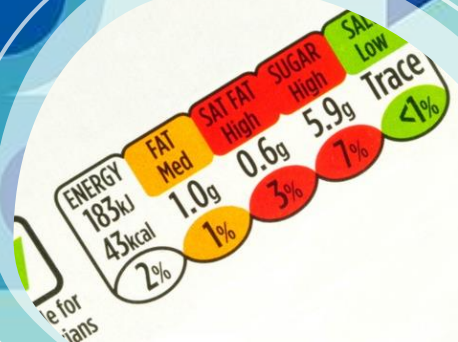
Conformity assessment according to EU AI Act proposal

Ethical Assurance Requires A Global Ecosystem

- Principles and Criteria Developers & Publishers
- Training and Information Entities
- Oversight, Advisory and Assessment Entities
- Certification Entities
- Regulatory Entities

- Civil Society Entities
- Public and Private Enterprises
- Product, System Developers and Service Providers
- User Communities

Holistic Approach from Design to Deployment

- Socio-technical Context
- Risk and Value Proposition
- Data quality
- Standards and Certifications
- Transparency
- Ecosystem Control
- Skills

Towards a Positive Vision for Humanity