IEEE SA
STANDARDS ASSOCIATION

EMPOWERING GLOBAL AI COLLABORATION TOWARDS AI FOR GOOD

Standardization for Ethical and Responsible AI

Yu Yuan, PhD

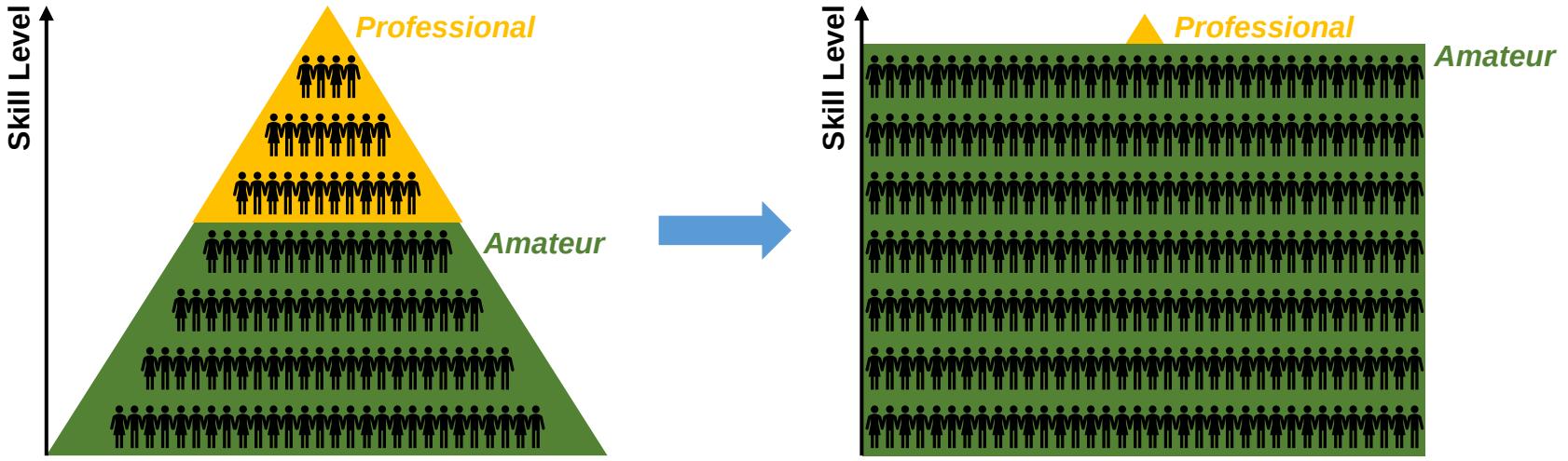
President, IEEE Standards Association

November 3, 2023

IEEE AI IT Artificial Intelligence Symposium

IEEE

Global Impacts of AI: *Labor Market*



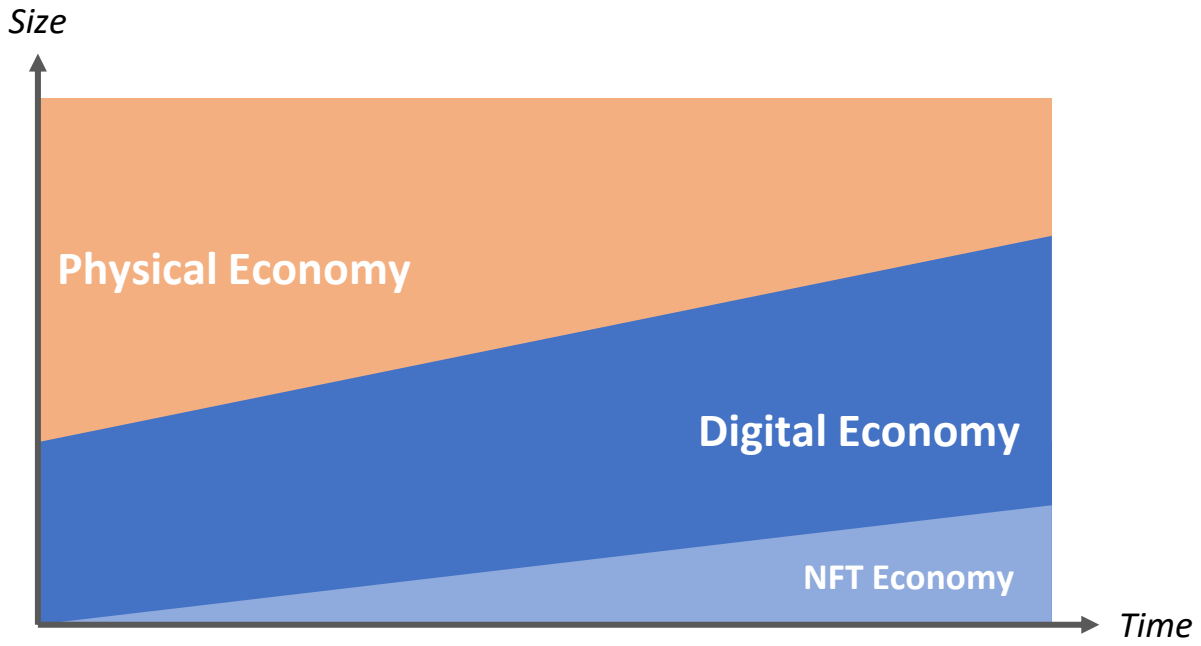
Labor Market Transformation Driven by Generative AI

1. In many professions, generative AI will significantly improve the skills of amateurs.
2. As a result, only a very small number of people can continue to be called professionals and make a living doing it as a job.
3. Many people will have to look for new job opportunities or enjoy fewer working hours and more leisure time.

NEW DEMANDS:

- New Job Opportunities
- Fewer Working Hours, More Leisure Time
- And more…

Global Impacts of AI: *Digital Economy*



- Generative AI will force the physical economy to shift towards the digital economy.
- Massive amounts of native digital assets will drive up the demand for ownership and transactions in the form of NFTs.

Global Impacts of AI: *Decentralization*

Image created with Midjourney

End Users	Decentralized Rights
End Users	Decentralized Ownership
Service Providers	Decentralized Governance
Service Providers	Decentralized Operations
Service Providers	Decentralized Infrastructure

Generative AI has lowered the barriers to production. Tasks that required a group of people to form a company can now be done by a single individual. This will drive a transformation of production relations towards decentralization.

Global Impacts of AI: *Risks*

Generated with AI · October 30, 2023 at 7:35 PM

AI has consciousness or not

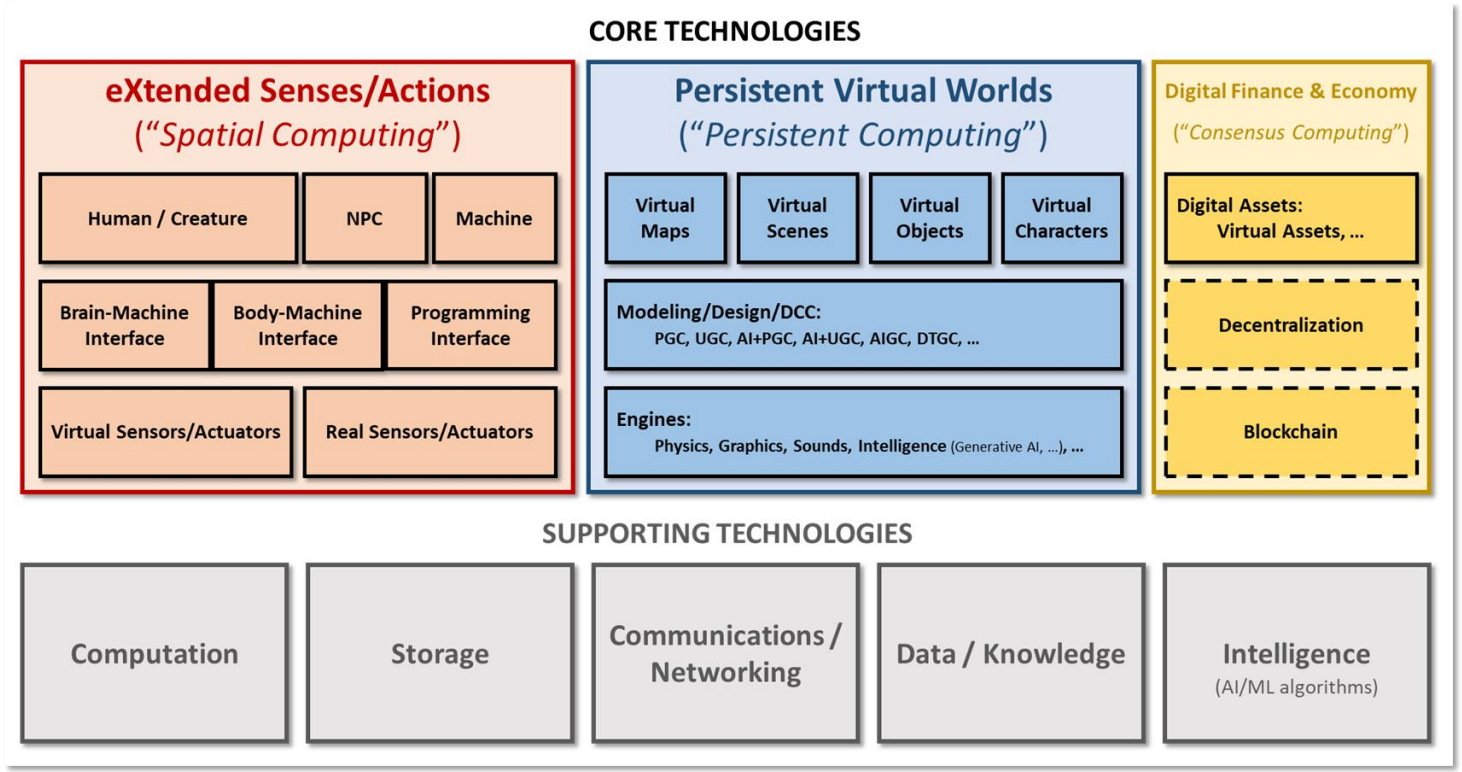
≠

AI can be dangerous or not

Future Directions of AI: *To Support Metaverse*

Large-scale and fine-grained virtual worlds, including massive and diverse virtual scenes and virtual objects, will be created at a hundred or even a thousand times faster with the aid of Generative AI. Hundreds of millions of virtual characters, driven by Generative AI, will provide a variety of deeply personalized interactive experiences and services.

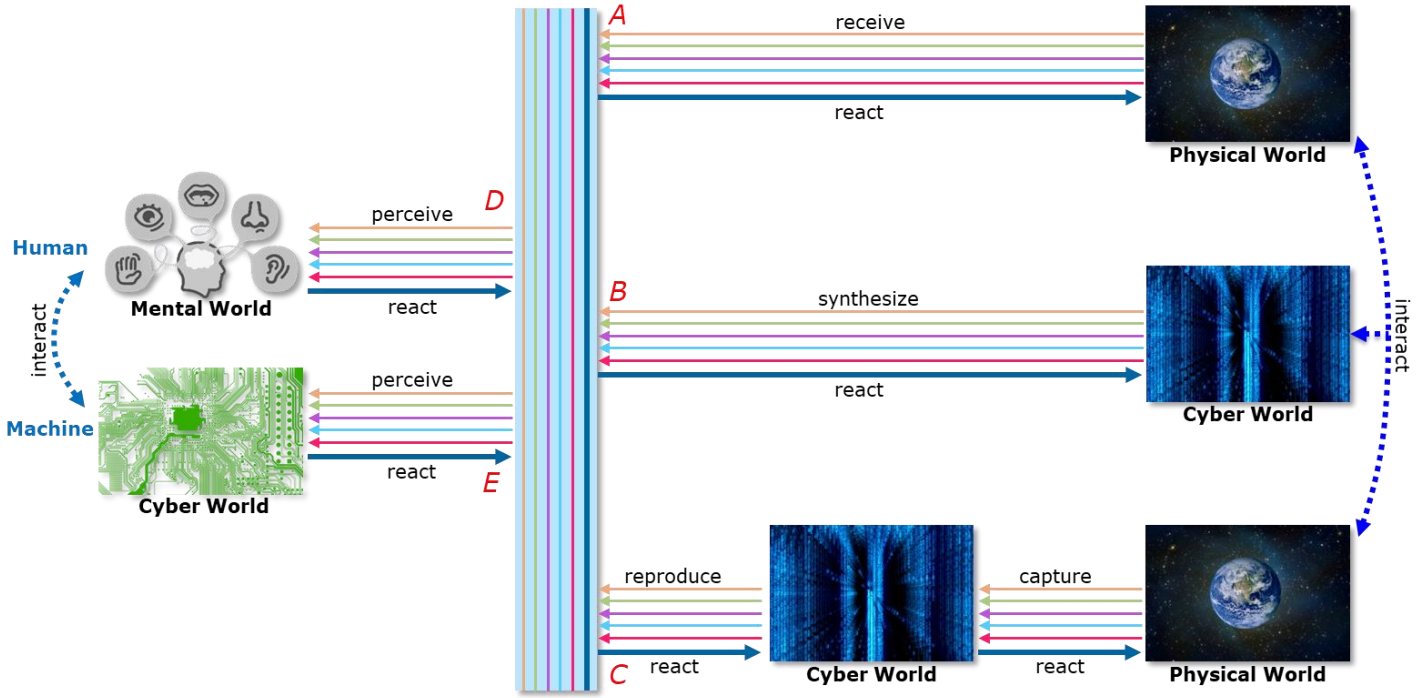
	"VR" Metaverse	"AR" Metaverse	"DT" Metaverse
Enriching the content	★★★	★★	★
Driving intelligent NPCs	★★★	★★	★
Understanding the worlds	★★	★★★	★★
Understanding the users	★★★	★★★	★



Metaverse Technology Landscape

Future Directions of AI: *To Be Supported by Metaverse*

The further development of AI, including the generalization of capabilities of large models, the rise of embodied AI, and the emergence of Artificial General Intelligence (AGI), requires a "real world" training environment to obtain massive multimodal training data. Metaverse can provide such a training environment with the lowest cost and risk, and the highest efficiency and diversity, thereby helping boost the development of AI. **Training AI in metaverses will become the new paradigm for AI research and development.** For example, testing autonomous driving under various complex road conditions in a virtual world is one of the first sprouts of this new paradigm and has already become a common practice in the industry.



Digital Senses – Conceptual Diagram

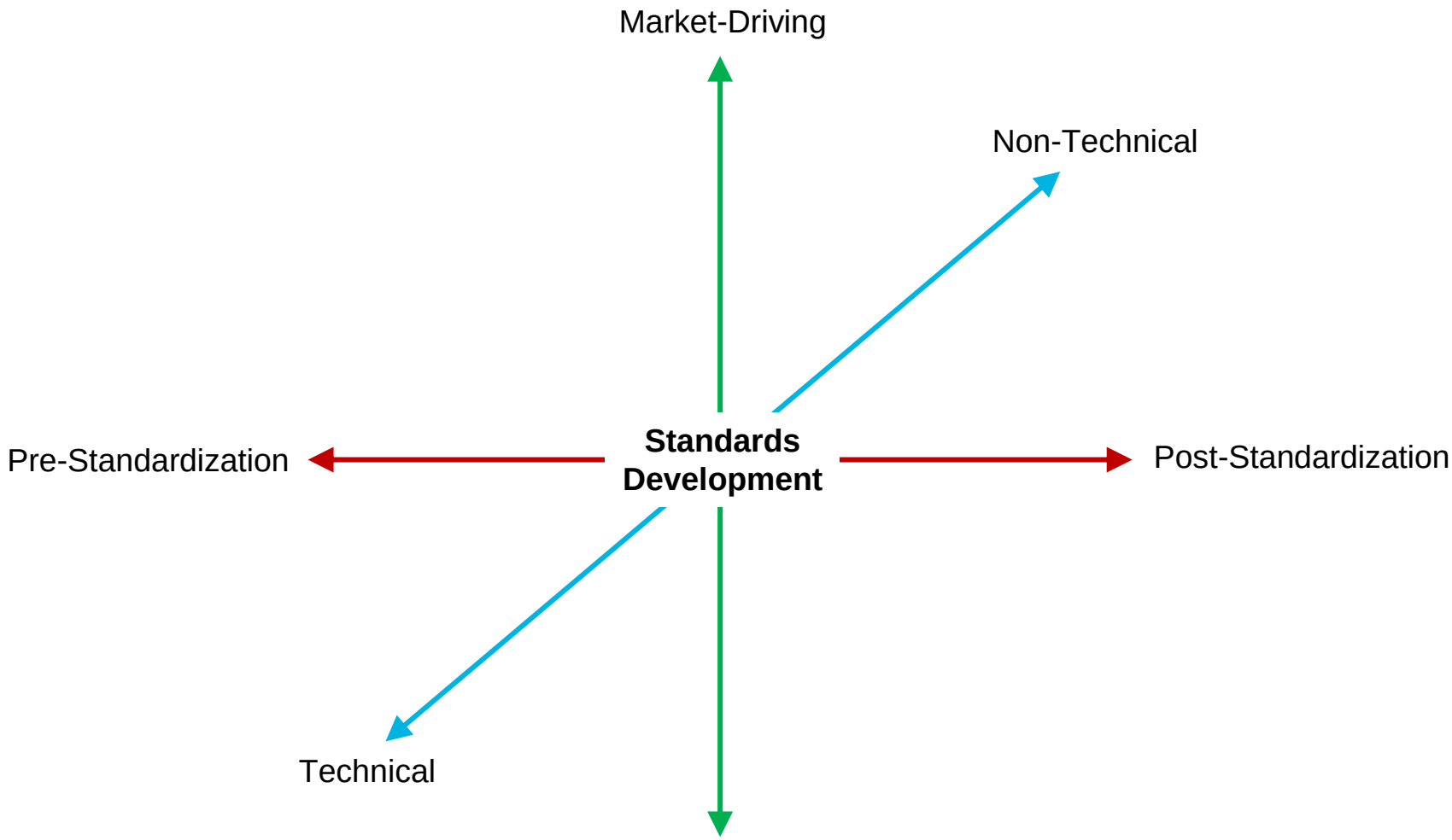
Source: Yu Yuan, “IEEE Digital Senses Initiative - overview and master plan”, 23 Sep 2016

AI Is Driving Digitalization

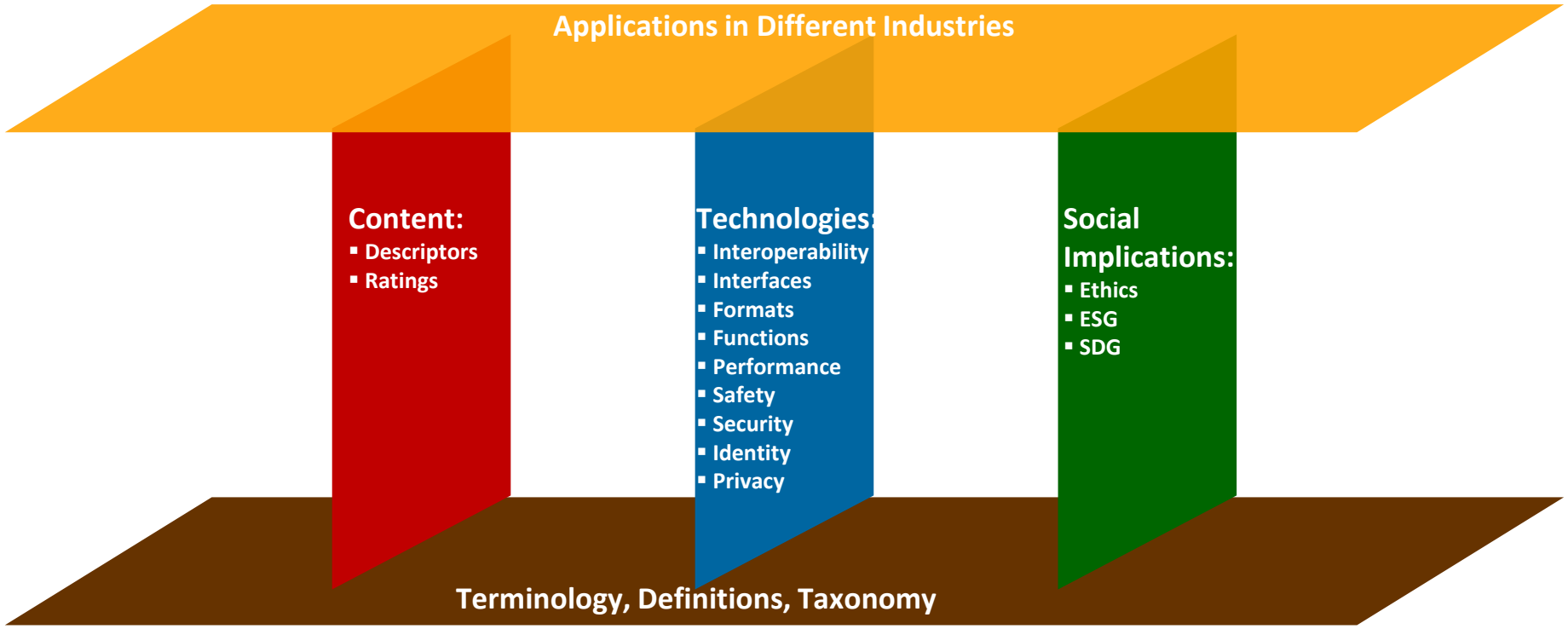
We are currently between two stages of Digitalization:
the exploding *Intelligentization* (AI) and
the upcoming *Realitization* (Metaverse).

Digitalization 1.0 – 5.0

The Expanding Scope of Standards Development



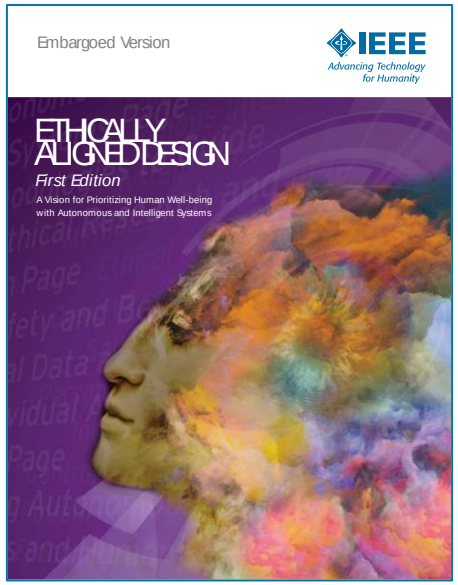
AI Standardization and Ecosystem

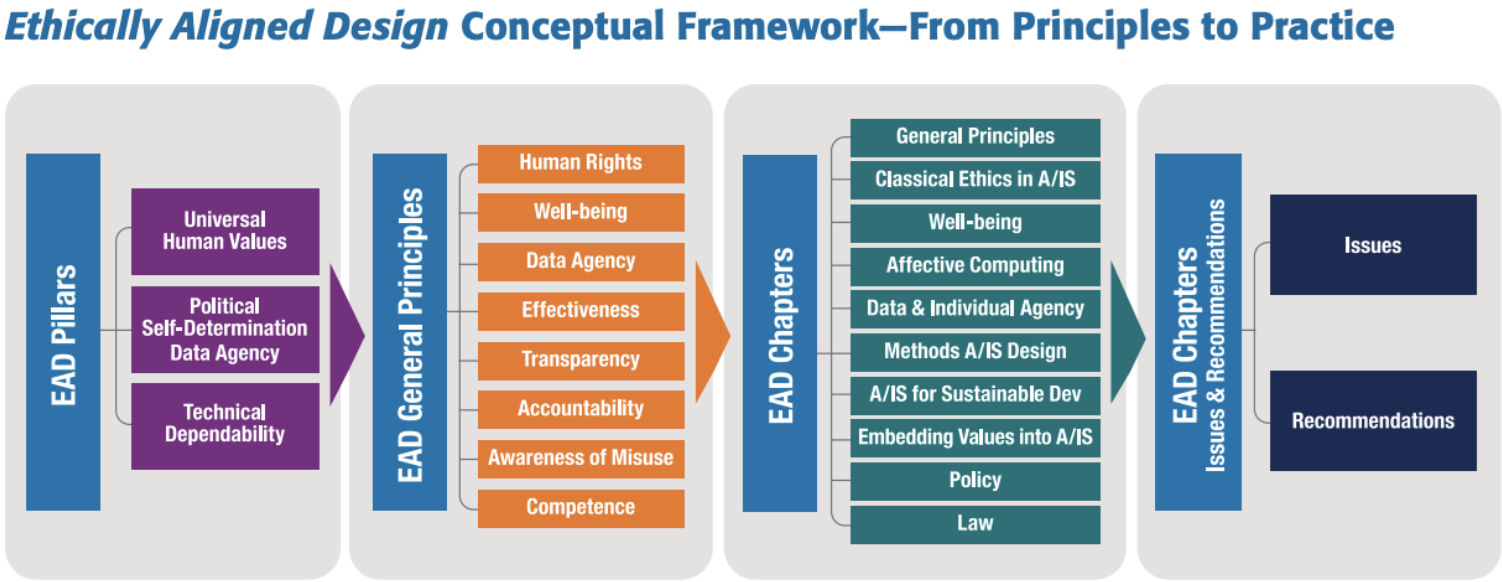


IEEE SA
STANDARDS ASSOCIATION

AIS ETHICS IN ACTION SINCE 2015

ETHICALLY ALIGNED DESIGN





- First version released in 2016.
- Latest version released March, 2019.
- Received over 500 pages of feedback.
- Written by more than 700 global thought leaders.
- Over 300 pages long – not just a list of AI Principles.
- Features hundreds of evergreen, pragmatic recommendations.
- Inspired IEEE's AIS Ethics Certification work.

- EAD,v2 used by OECD for their AI principles.
- EAD, v2 used by IBM for "Everyday Ethics for AI."
- EAD, v2 used by FLI for their AI Principles.
- EAD, v2 used by UNICEF for their Children's Data principles.
- EAD, v2 used by UNESCO for their AI Principles.
- EAD in all versions mentioned in more than three dozen academic journals, AI Principles and media since 2016

IEEE

IEEE SA
STANDARDS ASSOCIATION

PROGRAMS

IEEE ACTIVITIES IN AUTONOMOUS AND INTELLIGENT SYSTEMS

STANDARDS

A growing portfolio of over 40 technical and impact standards that promote innovation, foster interoperability and recognize human values.

https://standards.ieee.org/initiatives/artificial-intelligence-systems/index1.html

CERTIFICATION

Criteria and processes for certification methodology and marks in A/IS addressing (ECPAIS):

- Transparency
- Accountability
- Reduction of Algorithmic System Bias

standards.ieee.org/industry-connections/ecpais.html

EDUCATION & LEARNING

Empowering professionals in business and academia on how to practically address and include ethical considerations when developing and using A/IS products and services.

IEEE

IEEE SA
STANDARDS ASSOCIATION

IMPACT

IEEE ACTIVITIES IN AUTONOMOUS AND INTELLIGENT SYSTEMS

ETHICALLY ALIGNED DESIGN

A comprehensive report, created by the Global Initiative and endorsed by IEEE's Board of Directors, that combines a conceptual framework addressing universal human values, data agency and technical dependability with a set of principles to guide A/IS creators and users through a comprehensive set of recommendations.

https://ethicsinaction.ieee.org/

EAD FOR SPECIFIC SECTORS

A series of action-based papers targeted to specific audiences to help them implement Ethically Aligned Design. Current issues:

- Business
- Parenting
- Health
- Arts
- Advertising

PUBLIC AFFAIRS

Engaging with governments, municipalities and intergovernmental fora to contribute our technical and standardization expertise and open process skills to bridge the gap between technology, standards and policy making.

IEEE

IEEE SA
STANDARDS ASSOCIATION

STANDARDS

IEEE SA IMPACT STANDARDS – CONTEXTUA

IEEE P7000™ -
MODEL PROCESS FOR ADDRESSING ETHICAL CONCERNS DURING SYSTEM DESIGN

IEEE P7001™ -
TRANSPARENCY OF AUTONOMOUS SYSTEMS

IEEE P7002™ -
DATA PRIVACY PROCESS

IEEE P7003™ -
ALGORITHMIC BIAS CONSIDERATIONS

IEEE P7004™ -
STANDARD ON CHILD AND STUDENT DATA GOVERNANCE

IEEE P7005™ -
STANDARD ON EMPLOYER DATA GOVERNANCE

IEEE P7006™ -
STANDARD ON PERSONAL DATA AI AGENT

IEEE P7007™ -
ONTOLOGICAL STANDARD FOR ETHICALLY DRIVEN ROBOTICS AND AUTOMATION SYSTEMS

IEEE P7008™ -
STANDARD FOR ETHICALLY DRIVEN NUDGING FOR ROBOTIC, INTELLIGENT AND AUTONOMOUS SYSTEMS

IEEE P7009™ -
STANDARD FOR FAIL-SAFE DESIGN OF AUTONOMOUS AND SEMI-AUTONOMOUS SYSTEMS

IEEE P7010™ -
WELLBEING METRICS STANDARD FOR ETHICAL ARTIFICIAL INTELLIGENCE AND AUTONOMOUS SYSTEMS

IEEE P7011™ -
STANDARD FOR THE PROCESS OF IDENTIFYING & RATING THE TRUST-WORTHINESS OF NEWS SOURCES

IEEE P7012™ -
STANDARD FOR MACHINE READABLE PERSONAL PRIVACY TERMS

IEEE P7014™ -
STANDARD FOR EMULATED EMPATHY IN AUTONOMOUS AND INTELLIGENT SYSTEMS

IEEE

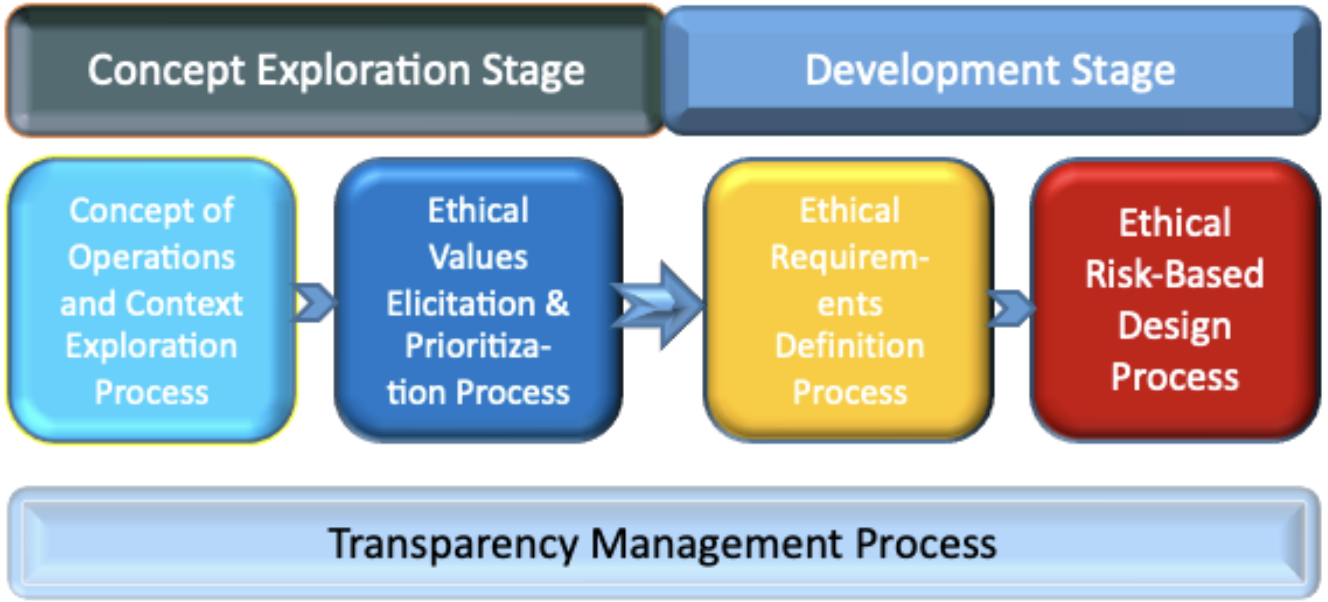
IEEE SA
STANDARDS ASSOCIATION

IEEE 7000-2021™

VALUES ORIENTED, METRICS-DRIVEN TECHNOLOGY AT THE OUTSET OF DESIGN

IEEE 7000™ offers concrete processes, with activities, tasks and mandatory outcomes that ensure an ethically aligned design

Value-based Engineering with IEEE 7000™ Processes



IEEE

IEEE SA
STANDARDS ASSOCIATION

AI PROCUREMENT

IEEE P3119™ ***Standard for the Procurement of Artificial Intelligence and Automated Decision Systems*** aims to address the needs of government workers, policymakers, and technologists to make meaningful and accountable choices that are transparent about the socio-technical considerations and impact of AI products, services, and or systems on the public.

IEEE

IEEE SA
STANDARDS ASSOCIATION

STANDARDS

IEEE SA TECHNICAL STANDARDS / PROJECTS

IEEE 3652.1-2020™ - GUIDE FOR ARCHITECTURAL FRAMEWORK AND APPLICATION OF FEDERATED MACHINE LEARNING

P2807™, P2807.1™ - KNOWLEDGE GRAPHS (FRAMEWORK, EVALUATION)

P1872.2™ - STANDARD FOR AUTONOMOUS ROBOTICS (AUR) ONTOLOGY

P2040/P2040.1™ - STANDARD FOR CONNECTED, AUTOMATED AND INTELLIGENT VEHICLES: OVERVIEW AND ARCHITECTURE (Taxonomy, Definitions)

P2801™ - IEEE Draft Recommended Practice for the Quality Management of Datasets for Medical Artificial Intelligence

P2802™ - Standard for the Performance and Safety Evaluation of Artificial Intelligence Based Medical Device: Terminology

P2418.4™ - STANDARD FOR THE FRAMEWORK OF DISTRIBUTED LEDGER TECHNOLOGY (DLT) USE IN CONNECTED AND AUTONOMOUS VEHICLES (CAVS)

P2751™ - 3D MAP DATA REPRESENTATION FOR ROBOTICS AND AUTOMATION

PC37.249™ - GUIDE FOR CATEGORIZING SECURITY NEEDS FOR PROTECTION AND AUTOMATION RELATED DATA FILES

P2671™ - IEEE Draft Standard for General Requirements of Online Detection Based on Machine Vision in Intelligent Manufacturing

P2672™ - GUIDE FOR GENERAL REQUIREMENTS OF MASS CUSTOMIZATION

P2975™ - Standard for Industrial Artificial Intelligence (AI) Data Attributes

IEEE 1589-2020™ - STANDARD FOR AN AUGMENTED REALITY LEARNING EXPERIENCE MODEL

P2247.1™, P2247.2™, P2247.3™ - ADAPTIVE INSTRUCTOINAL SYSTEMS (CLASSIFICATION, INTEROPERABILITY, AND EVALUATION)

P2830™ - STANDARD FOR TECHNICAL FRAMEWORK AND REQUIREMENTS OF SHARED MACHINE LEARNING

IEEE 3333.1.3-2022™ - STANDARD FOR THE DEEP LEARNING-BASED ASSESSMENT OF VISUAL EXPERIENCE BASED ON HUMAN FACTORS

P3123™ - Standard for Artificial Intelligence and Machine Learning (AI/ML) Terminology and Data Formats

P3127™ - Guide for an Architectural Framework for Blockchain-based Federated Machine Learning

P2986™ - Recommended Practice for Privacy and Security for Federated Machine Learning

P3142™ - Recommended Practice on Distributed Training and Inference for Large-scale Deep Learning Models

P3152™ - Description of the Natural or Artificial Character of Intelligent Communicators

P3154™ - Recommended Practice for the Application of Knowledge Graphs for Talent Services

P3156™ - Standard for Requirements of Privacy-preserving Computation Integrated Platform

P3157™ - Recommended Practice for Vulnerability Test for Machine Learning Models for Computer Vision Applications

P2959™ - Standard for Technical Requirements of Standard-Oriented Knowledge Graphs

P3110™ - Standard for Technical **Standard for Computer Vision (CV) - Algorithms, Application Programming Interfaces (API), and Technical Requirements for Deep Learning Framework**

P2660.1™ - RECOMMENDED PRACTICES ON INDUSTRIAL AGENTS: INTEGRATION OF SOFTWARE AGENTS AND LOW LEVEL AUTOMATION FUNCTIONS

P2841™ - IEEE Draft Framework and Process for Deep Learning Evaluation

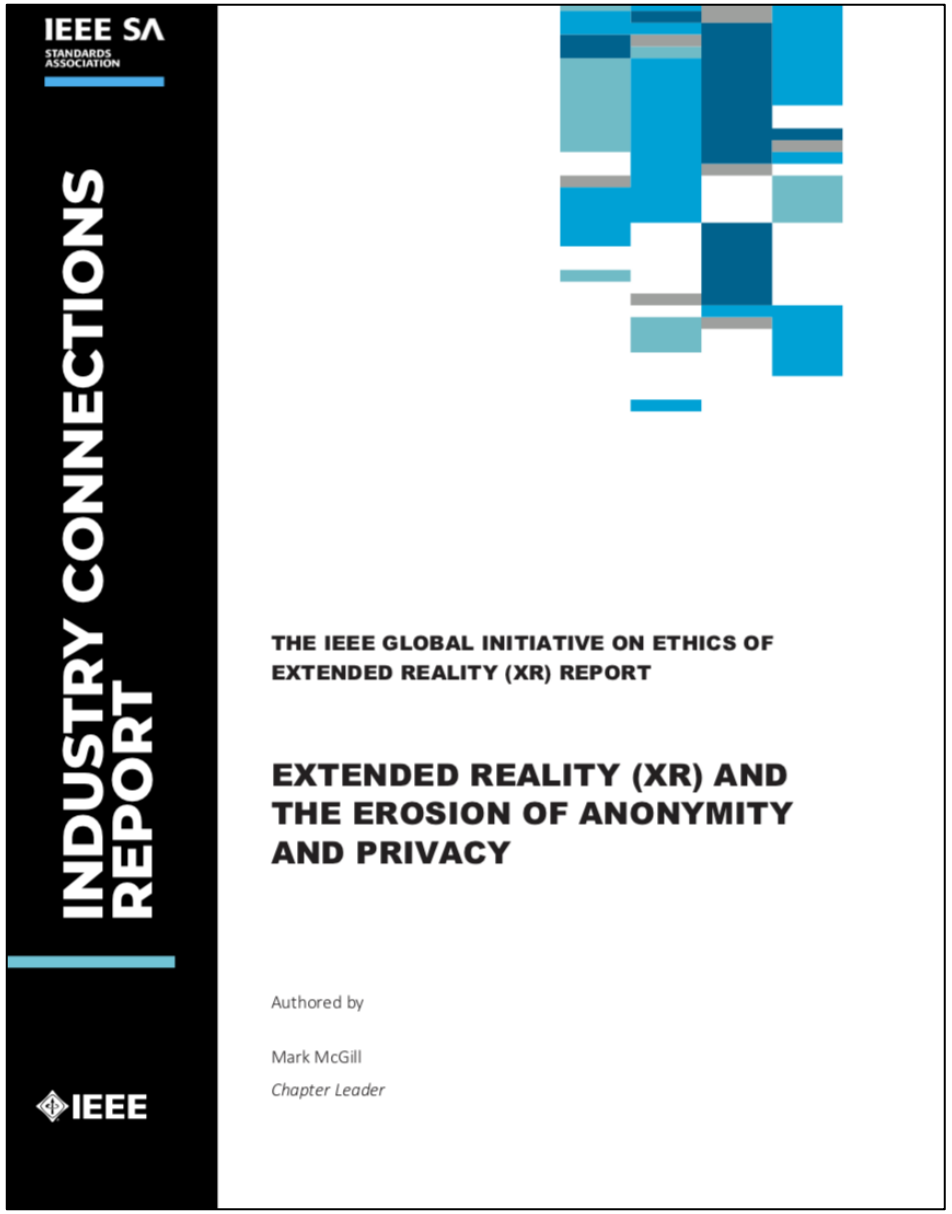
P2863™ - **Recommended Practice for Organizational Governance of Artificial Intelligence**

P2976™ - **Standard for XAI – eXplainable Artificial Intelligence - for Achieving Clarity and Interoperability of AI Systems Design**

IEEE

IEEE SA
STANDARDS ASSOCIATION

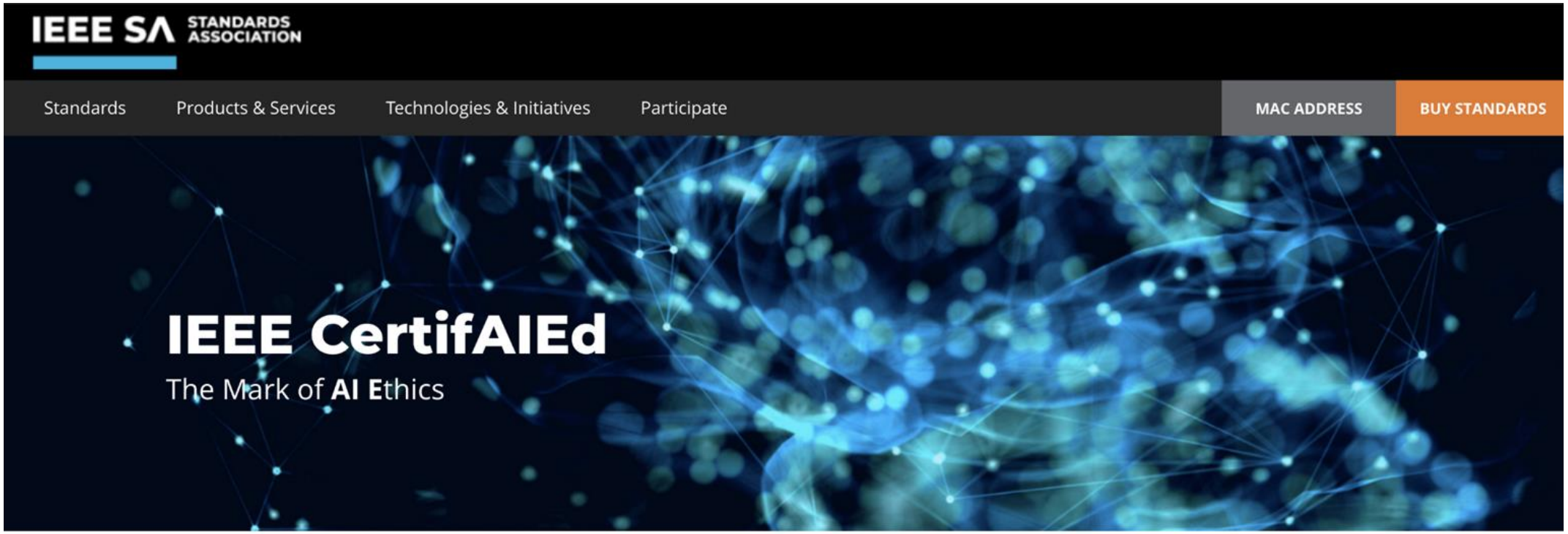
ETHICS OF EXTENDED REALITY



IEEE

IEEE SA
STANDARDS ASSOCIATION

IEEE CertifAIEd



IEEE SA is proud to announce IEEE CertifAIEd; a mark designed to inspire trust and responsible innovation in AI systems. IEEE CertifAIEd offers a risk-based framework with a suite of AI ethical criteria that can be contextualized to fit organizations' needs– helping them deliver a greater trustworthy experience for their users.

The IEEE CertifAIEd mark recognizes that your product, service, or system has been verified to meet relevant ethical criteria, contributing towards a greater level of confidence and demonstrating a proactive approach to building public trust in your AI system. It sets the standard that AI products, services and systems should meet in order to deliver authentic and practical value and trust.

The IEEE CertifAIEd Mark

Get Involved

IEEE

IEEE Industrial AI – Intersection of Industrial Automation & AI

- **PRE-STANDARDIZATION ACTIVITIES ON INDUSTRIAL AI**
  - Industrial AI differs from consumer AI applications in terms of data quality and privacy aspects, information content, and impact of AI on various stakeholders.
  - Identify requirements toward standardization, through an use case-driven
- **THE IEEE APPLIED ARTIFICIAL INTELLIGENCE SYSTEMS (AIS) RISK AND IMPACT FRAMEWORK INITIATIVE**
  - Understand and propose an applied risk framework or assessment aligning with the current regulations
  - Consider existing risk approaches in the fields of finance, cybersecurity, and more, identify gaps introduced by AI
- **STANDARDS UNDER DEVELOPMENT (OR PROPOSED)**
  - IEEE P2975: Standard for Industrial Artificial Intelligence (AI) Data Attributes
  - Verification and Validation of AI Models
  - AI Maturity Frameworks
  - Recommended Software framework at the Industrial Edge

IEEE SA STANDARDS ASSOCIATION

RAISING THE WORLD'S STANDARDS

NEW APPROACHES: IEEE SA Centers of Competence

Health & Life Sciences
- Pharma
- Clinical health
- Wellness

Mobility
- Safe autonomous driving
- Zero emissions
- MaaS

Energy
- Smart energy
- Energy efficiency
- Energy infrastructure

New Frontiers & Emerging Spaces
- Sustainability
- AIS
- Emerging technology areas

Foundational Technologies
- Industrial IoT
- Consumer Applications
- Infrastructure for Emerging Technologies

Telecommunications & Connectivity
- Rural
- Next generation telecom
- Wired and hybrid communications

IEEE SA
STANDARDS ASSOCIATION

RAISING THE WORLD'S STANDARDS

STANDARDS & SOCIAL IMPACT

Our global community is developing sustainable, consensus-based technical standards and solutions for societal issues.

- **Ethical AI Systems**
- **Data Governance**
- **Fairness in the Trade of Data**
- **Child Online Rights**
- **Digital Intelligence**
- **Dignity & Agency in Identity**
- **Clean & Sustainable Energy**
- **Connectivity & Mobility**

IEEE

IEEE SA
STANDARDS ASSOCIATION

RAISING THE WORLD'S STANDARDS

ABOUT IEEE SA

Developing market relevant open standards and solutions:

- Advancing global technologies and technology platforms
- Promoting innovation
- Protecting public safety, health & wellbeing
- Contributing to a sustainable future

IEEE

Thank you

- **Yu Yuan, PhD**

  President, IEEE Standards Association
  y.yuan@ieee.org

- IEEE Standards Home Page: standards.ieee.org