Formal AI Standardization

Gavin Jones
BSI Lead Standards Development Manager
(AI and quantum)

Joy Tan
BSI Senior Standards Manager
(Digital & Digital Manufacturing)

What is AI?

- AI products and services in early stages of development

- Regulations and standards looking into new ways of working

AI Definition
"An AI system is a machine-based system designed to operate with varying levels of autonomy and that for, explicit or implicit objectives, generate outputs such as predictions, recommendations or decisions, that influence physical or virtual environments." (OECD)

AI History
Arthur Samuel develops program playing checkers, the first to learn independently (1952); "artificial intelligence" term first used by John McCarthy (1955)

bsi.

What are National Standards Bodies and "formal" standards?

- *BSI appointed by UK Government* as the national standards body (NSB) and holds a *Royal Charter* to *improve* the quality and safety of products, services and systems by enabling standards creation and encouraging their use

- *Represents* UK interests at the International Organization for Standardization (ISO), the International Electrotechnical Commission (IEC) and the European Standards Organizations (CEN, CENELEC and ETSI)

- *Purpose* to support government policy, UK business and industry, UK education, and UK consumer rights and safety

- *3,100* standards published each year written by *9,000* UK experts

By Royal Charter

Where do BSs fit?

- Convenor role facilitating consensus on good-practice organisational issues

- Frameworks, principles, codes, guidance, and technical specifications and reports

All ENs and most international standards adopted as British Standards (BSs)

BSI as NSB manages BS, CENELEC, EN, ISO and IEC standards for UK

Private and professional standards, codes and guidance

Int'l Standards (ISOs, IECs)

Regional Standards (ENs)

National Standards (BSs)

Sponsored standards (BSI PASs)

Private and Consortia Standards

Corporate Technical Specifications

Professional Codes, Guidance, Best Practice

bsi.

Consensus and building trust

"General agreement, characterized by the absence of sustained opposition to substantial issues by any important part of the concerned interests and by a process that involves seeking to take into account the views of all parties concerned and to reconcile any conflicting arguments.
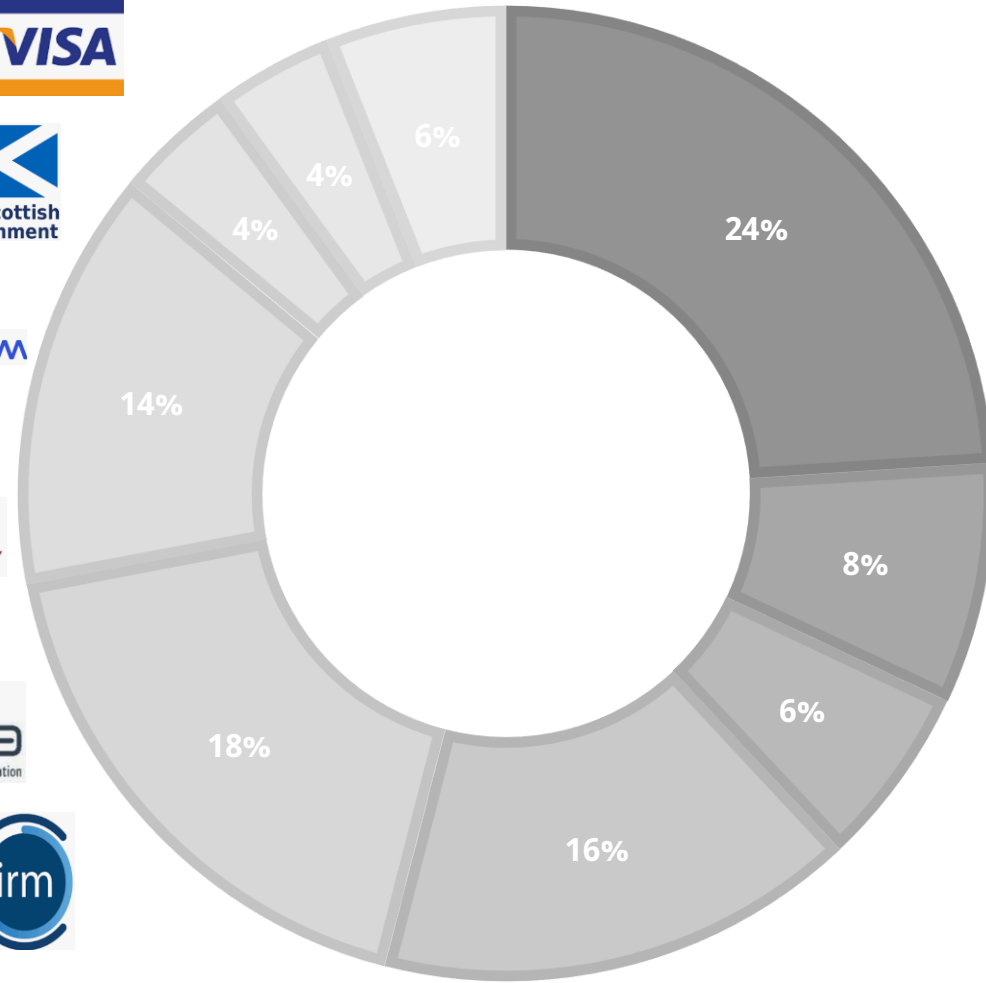
NOTE Consensus need not imply unanimity."

(ISO/IEC Guide 2:2004, definition 1.7)

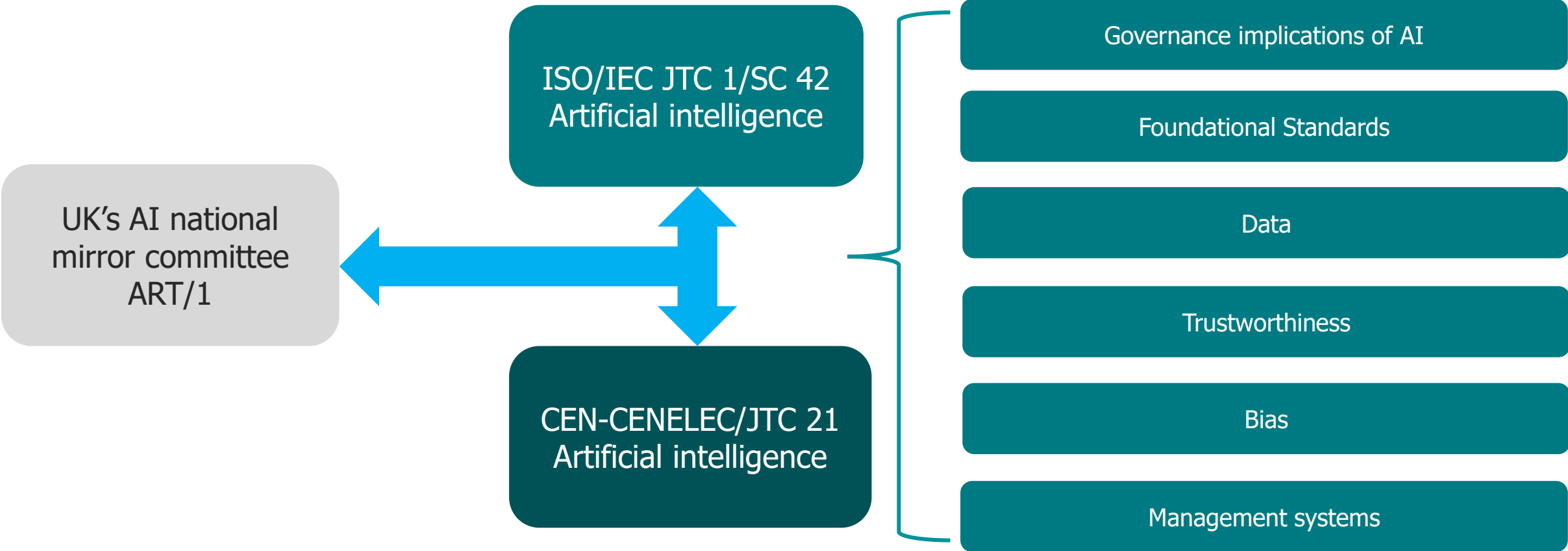# National Standards Bodies don't write standards - expert committees do!

ART/1 – Diverse growing membership (118 experts, 11 centuries' AI expertise)

Donut chart legend and values:
- Academic/Research Institution — 24%
- Consultancy/ Professional Services — 8%
- Government — 6%
- Individual Experts — 16%
- Large Enterprise — 18%
- SME — 14%
- Non-profit — 4%
- Trade Association/Professional Body — 4%
- Other — 6%

UK, international and regional engagement in AI standards

```
                    ┌──────────────────────────────┐
                    │  Governance implications of AI │
                    └──────────────────────────────┘
┌─────────────────────────┐   ┌──────────────────────────────┐
│  ISO/IEC JTC 1/SC 42    │   │    Foundational Standards    │
│  Artificial intelligence │   └──────────────────────────────┘
└─────────────────────────┘   ┌──────────────────────────────┐
                              │             Data             │
┌──────────────────┐          └──────────────────────────────┘
│ UK's AI national │          ┌──────────────────────────────┐
│ mirror committee │          │        Trustworthiness       │
│      ART/1       │          └──────────────────────────────┘
└──────────────────┘          ┌──────────────────────────────┐
┌─────────────────────────┐   │             Bias             │
│  CEN-CENELEC/JTC 21     │   └──────────────────────────────┘
│  Artificial intelligence │   ┌──────────────────────────────┐
└─────────────────────────┘   │      Management systems      │
                              └──────────────────────────────┘
```

bsi.

ISO/IEC/JTC 1/SC 42 AI activities

UK is active member with experts on various work streams, with successful proposals already won for new standards.

- *37* participating member countries

- *22* observing member countries

- *30* projects under development

- *20* published projects

- *10* working groups including *foundation standards, data trustworthiness, health informatics, testing, use cases*

bsi.

CEN-CENELEC/JTC 21 AI activities

- European-specific work programmes for regional needs

- Development of "harmonized" standards in support of the European AI Act

- Membership not unique to EU nations; UK leads on key development areas including conformity assessment (ISO 42006)

- Alignment and adoption of international standards

bsi.

ISO/IEC 42001 *AI Management Systems*

- World's *first* standard organisations can be *certified* to

- Requirements/guidance for establishing, implementing, maintaining and continually improving organizations' AI management systems.

Written to help organizations..

- develop and use AI responsibly in line with organizational objectives
- meet applicable regulatory requirements and obligations related to interested parties

Content requirements include...

- AI policy
- planning
- risk assessment (treatments and impact)
- competence
- internal auditing and programming (define objectives, criteria and scope; select auditors; ensure audits are appropriate)

Management System Standards (MSSs)

MSSs set requirements/guidance to help ensure policy and process management achieves organisational objectives across **diverse sectors**, **organization types** and **sizes**, and **diverse geographies**, **cultures** and **social conditions**

- *MSS **structure** allows **multiple** and **concurrent MSS use***

- World's most-used MSSs first created as British Standards:

ISO 9001 *quality management* **(BS 5750)**

ISO 14001 *environmental management* **(BS 7750)**

ISO 27001 *information security* **(BS 7799)**

bsi.

Overview of AI work programme

Flexible

Durable

Transversal & Scalable

Key Principles to create a responsible certifiable AI system

Data Ecosystem

Characteristics

Properties

Quality

Digital Infrastructure

Technologies & Applications
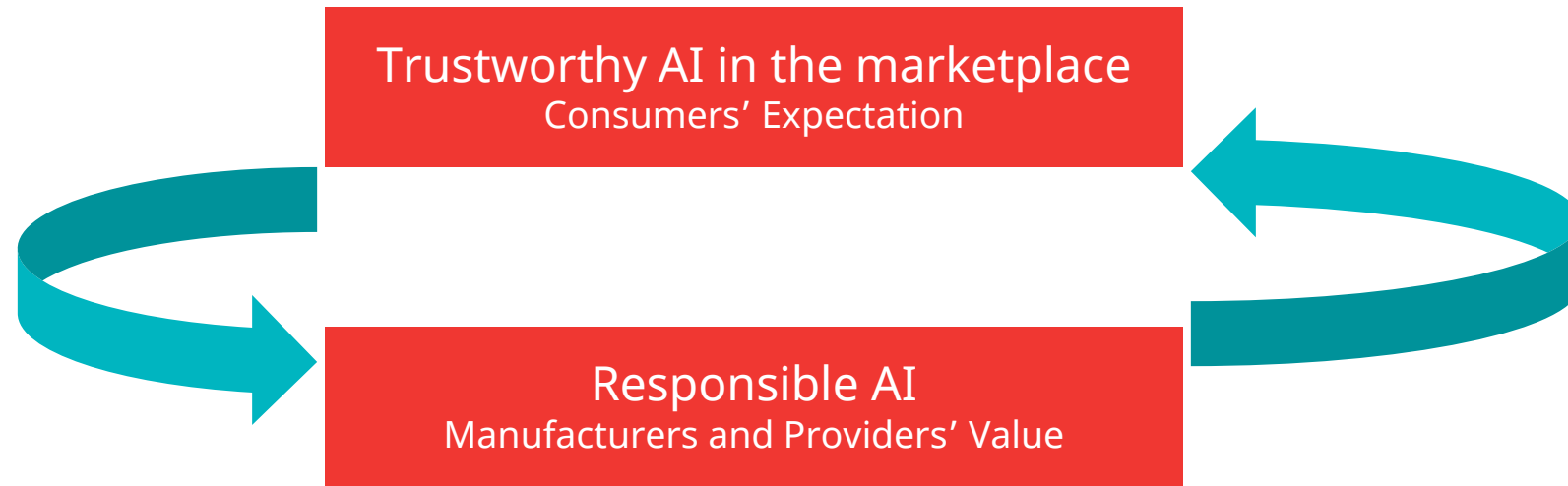
Artificial Intelligence

Big Data

Analytics

Building the ecosystem for AI systems to acquire, process and apply knowledge which relies on data, properties and quality.

- Management of extensive data sets (storage, manipulation, management and analysis)

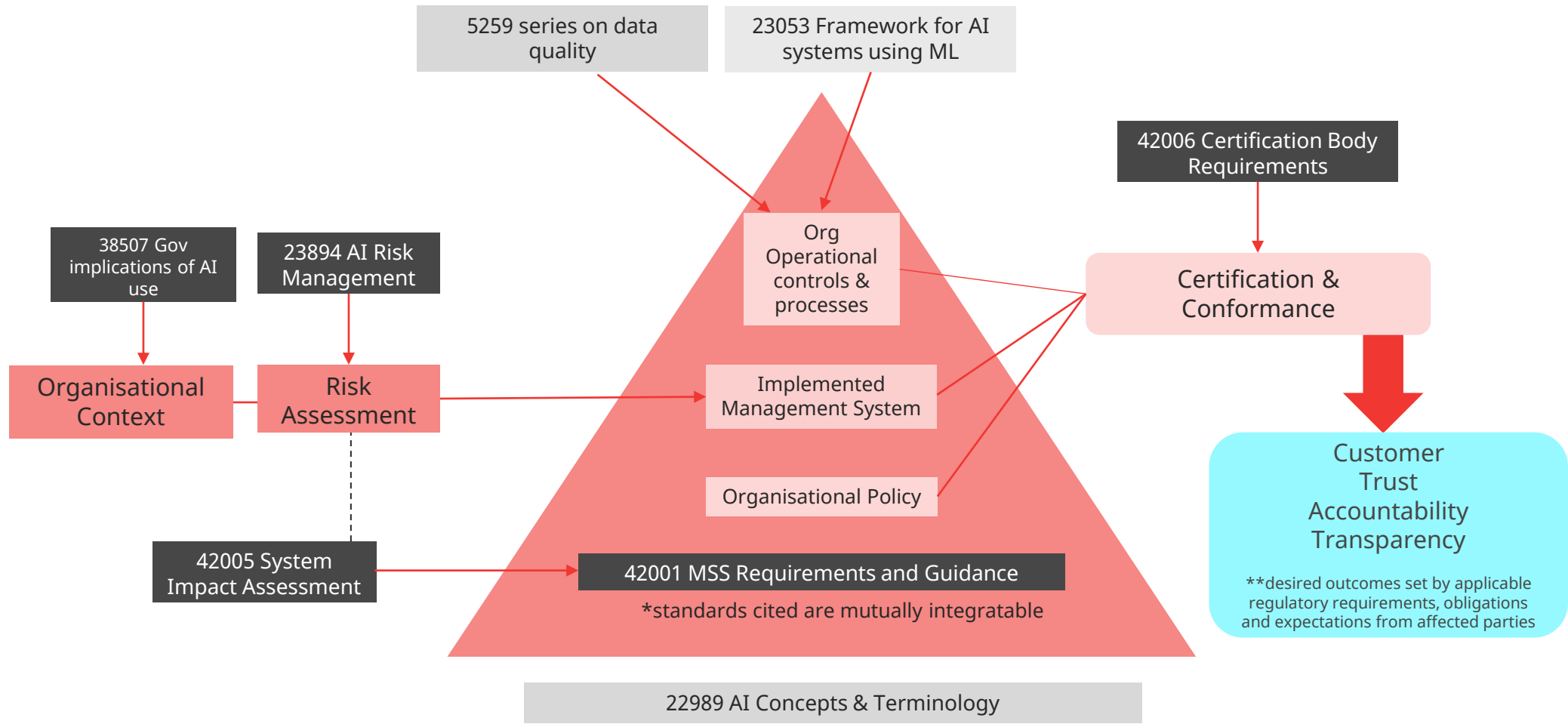- Volume, variety, velocity, variability (qualities that ensure scalability)

Derivation of data insights which requires acquisition, collection, validation and processing
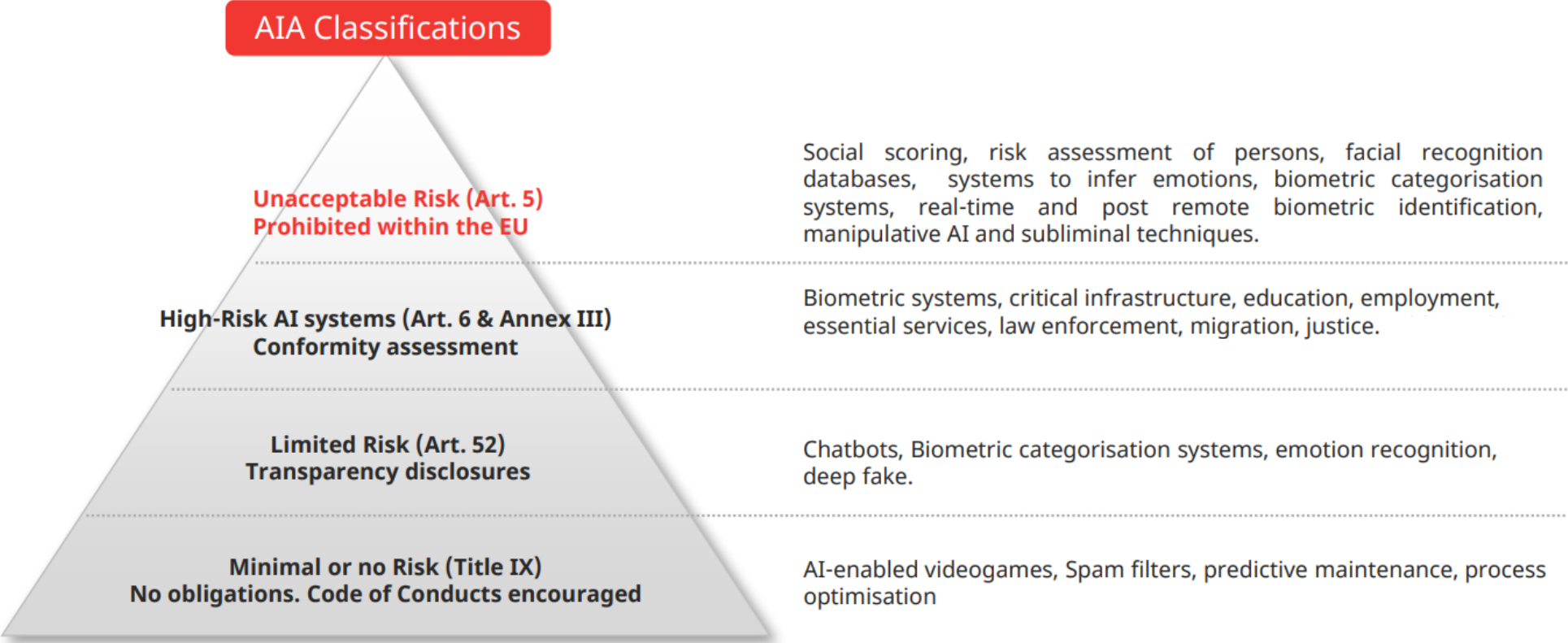
bsi.

Definition of a Beneficial AI Ecosystem

Trustworthy AI in the marketplace
Consumers' Expectation

Responsible AI
Manufacturers and Providers' Value

How should a responsible AI Standards Ecosystem look like?

5259 series on data quality

23053 Framework for AI systems using ML

42006 Certification Body Requirements

38507 Gov implications of AI use

23894 AI Risk Management

Organisational Context

Risk Assessment

42005 System Impact Assessment

Org Operational controls & processes

Implemented Management System

Organisational Policy

42001 MSS Requirements and Guidance

*standards cited are mutually integratable

22989 AI Concepts & Terminology

Certification & Conformance

Customer
Trust
Accountability
Transparency

**desired outcomes set by applicable regulatory requirements, obligations and expectations from affected parties

EU AI Act (coming into force 2024)

In development, one of the world's first AI regulations, risk-based delegation of requirements to harmonized ENs

AIA Classifications

Unacceptable Risk (Art. 5)
Prohibited within the EU
Social scoring, risk assessment of persons, facial recognition databases, systems to infer emotions, biometric categorisation systems, real-time and post remote biometric identification, manipulative AI and subliminal techniques.

High-Risk AI systems (Art. 6 & Annex III)
Conformity assessment
Biometric systems, critical infrastructure, education, employment, essential services, law enforcement, migration, justice.

Limited Risk (Art. 52)
Transparency disclosures
Chatbots, Biometric categorisation systems, emotion recognition, deep fake.

Minimal or no Risk (Title IX)
No obligations. Code of Conducts encouraged
AI-enabled videogames, Spam filters, predictive maintenance, process optimisation

bsi.

AI Regulatory Landscape: is global consensus possible?

Pro-Innovation
Existing regulators are expected to implement these 5 principles across sectors;
- Safety, security and robustness
- Appropriate transparency and explainability
- Fairness
- Accountability and governance
- Contestability and redress

Layered approach
Layer 1: consistent common foundations across regulatory remits
Layer 2: Adapted to specific risk
Layer 3: Encourage the adoption of sector-specific technical standards to support compliance and performance

Stringent
Classification of AI systems by risk and mandating development and user requirements

- Unacceptable
E.g. Social scoring or biometric identification
- High
E.g. AV, Med Devices, critical infrastructure
- Limited
- Minimal
E.g. Filters and video games

Adherence to Regulations = Rigorous testing, proper documentation of data quality and accountability framework

Est publication: JAN 2024

Aus: No comprehensive AI legislation, reliance on existing law. Some binding state guidelines e.g. NSW AI Assurance framework

US: No comprehensive AI legislation. Some guidelines e.g. White House 2020 Guidance for Reg AI Applications, AI Bill of rights, NIST AI RMF

China: Provincial & Local level law, regulation and guidelines.

Singapore: Reliance on existing laws e.g. MAS FEAT, Use of AI and Data Analytics in Financial Sector (2018) and Nat AI Strategy (2019) and the Veritas Initiative (2021)

Japan: Reliance on existing laws. Governance guidelines for AI principles, Contract guidelines on AI utilisation and Data, ML Quality Management Guidelines.

UK and AI in 2023 so far...

- *Pro-innovation Approach to AI Regulation* UK Government whitepaper issued March 2023

- Supporting technical standards are in consultation and drafting phases

- Large language model (LLM) systems made AI **red hot** in 2023

- BSI engaging and influencing *nationally, regionally* and *internationally* in AI

Take part and become a standards maker!

Why?

- Help improve life quality and safety for millions of people

- Create a better world

- Influence standards that affect your business and network

- Gain credibility with customers and stakeholders

How?

- Search "become a standards maker" (see Annex 3)

- Identify the committee you wish to join

- Complete the survey

- Upload your CV

For international/regional participation please contact your NSB for the country your reside in.

bsi.

Thank you

gavin.jones@bsigroup.com

joy.tan@bsigroup.com

By Royal Charter

bsi.

Annex 1 – ART/1 published standards

BS ISO/IEC 20546:2019	Information technology. Big data. Overview and vocabulary
BS 10102-1:2020	Big data. Guidance on data-driven organizations
BS 10102-2:2020	Big data. Guidance on data-intensive projects
BS ISO/IEC 20547-3:2020	Information technology. Big data reference architecture. Reference architecture
PD ISO/IEC TR 24028:2020	Information technology. Artificial intelligence. Overview of trustworthiness in artificial intelligence
PD ISO/IEC TR 20547-1:2020	Information technology. Big data reference architecture. Framework and application process
PD ISO/IEC TR 24029-1:2021	Artificial Intelligence (AI). Assessment of the robustness of neural networks. Overview
PD ISO/IEC TR 24027:2021	Information technology. Artificial intelligence (AI). Bias in AI systems and AI aided decision making
PD ISO/IEC TR 24372:2021	Information technology. Artificial intelligence (AI). Overview of computational approaches for AI systems
BS ISO/IEC 38507:2022	Information technology. Governance of IT. Governance implications of the use of artificial intelligence by organizations
PD ISO/IEC TR 24368:2022	Information technology. Artificial intelligence. Overview of ethical and societal concerns
PD ISO/IEC/TS 4213:2022	Information technology. Artificial Intelligence. Assessment of machine learning classification performance
BS ISO/IEC 24668:2022	Information technology. Artificial intelligence. Process management framework for big data analytics
BS ISO/IEC 23894:2023	Information technology. Artificial intelligence. Guidance on risk management
BS ISO/IEC 25059:2023	Software engineering. Systems and software Quality Requirements and Evaluation (SQuaRE). Quality model for AI systems
BS EN ISO/IEC 22989:2023	Information technology - Artificial intelligence - Artificial intelligence concepts and terminology
BS EN ISO/IEC 23053:2023	Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)
BS ISO/IEC 5469	Artificial intelligence — Functional safety and AI systems

bsi.

Due	Standard	Standard title
14/12/2023	ISO/IEC TR 24030:2021	Information technology — Artificial intelligence (AI) — Use cases
9/2023	EN ISO/IEC 5338	Information technology — Artificial intelligence — AI system life cycle processes
9/2023	EN ISO/IEC 5339	Information technology — Artificial intelligence — Guidance for AI applications
21/9/2023	EN ISO/IEC 24029-2:2023	Artificial intelligence (AI) — Assessment of the robustness of neural networks — Part 2: Methodology for the use of formal methods
	ISO/IEC 8183:2023	Information technology – Artificial Intelligence – Data life cycle framework
11/2023	**EN ISO/IEC 42001**	**Information technology — Artificial intelligence — Management system**
10/31/2023	ISO/IEC 24030	Information technology — Artificial intelligence (AI) — Use cases
11/28/2023	ISO/IEC 25058	Software and systems engineering — Systems and software Quality Requirements and Evaluation — Guidance for quality evaluation of AI systems
12/20/2023	EN ISO/IEC 5392	Information technology — Artificial intelligence — Reference architecture of knowledge engineering
1/18/2024	EN ISO/IEC 5259-2	Artificial intelligence — Data quality for analytics and machine learning (ML) — Part 2: Data quality measures
2/28/2024	ISO/IECTR 5259-6	Artificial intelligence — Data quality for analytics and machine learning (ML) — Part 6: Visualization framework for data quality
3/7/2024	EN ISO /IEC/TR 24027	Information technology — Artificial Intelligence (AI) — Bias in AI systems and AI aided decision making
	EN ISO /IEC/TR 24029-1	Artificial Intelligence (AI) — Assessment of the robustness of neural networks — Part 1: Overview
3/14/2024	EN ISO/IEC CDTS 12791	Information technology — Artificial intelligence — Treatment of unwanted bias in classification and regression machine learning tasks

bsi.

Annex 2 – ART/1 standards pipeline (Part 2)

Due	Standard	Standard title
4/26/2024	ISO/IECCDTS 8200	Information technology - Artificial intelligence - Controllability of automated artificial intelligence systems
5/13/2024	ISO/IECCDTR 17903	Information technology — Artificial intelligence — Overview of machine learning computing devices
5/29/2024	ISO/IEC 6254	Information technology — Artificial intelligence — Objectives and approaches for explainability of ML models and AI systems
6/18/2024	ISO/IEC/TR 20226	Information technology — Artificial intelligence — Environmental sustainability aspects of AI systems
7/8/2024	**EN ISO/IEC 42006**	**Information technology — Artificial intelligence — Requirements for bodies providing audit and certification of artificial intelligence management systems**
8/29/2024	ISO/IEC AWI TS 17847	Information technology — Artificial intelligence — Verification and validation analysis of AI systems
9/17/2024	EN ISO/IEC 5259-1	Artificial intelligence — Data quality for analytics and machine learning (ML) — Part 1: Overview, terminology, and examples
	EN ISO/IEC 5259-3	Artificial intelligence — Data quality for analytics and machine learning (ML) — Part 3: Data quality management requirements and guidelines
	EN ISO/IEC 5259-4	Artificial intelligence — Data quality for analytics and machine learning (ML) — Part 4: Data quality process framework
11/18/2024	ISO/IEC/TR 42103	Information technology — Artificial intelligence — Overview of synthetic data in the context of AI systems
1/2/2025	EN ISO/IEC 5259-5	Artificial intelligence — Data quality for analytics and machine learning (ML) — Part 5: Data quality governance
1/6/2025	ISO/IEC/TR 18988	Artificial intelligence — Application of AI technologies in health informatics
3/5/2025	EN ISO/IEC 12792	Information technology — Artificial intelligence — Transparency taxonomy of AI systems
3/19/2025	ISO/IEC/TR 42106	Information technology — Artificial intelligence — Overview of differentiated benchmarking of AI system quality characteristics
5/7/2025	ISO/IEC/TR 21221	Information technology – Artificial intelligence – Beneficial AI systems
8/19/2025	EN ISO/IEC 42005	Information technology — Artificial intelligence — AI system impact assessment
3/11/2026	EN ISO/IEC 42105	Information technology — Artificial intelligence — Guidance for human oversight
3/18/2026	EN ISO/IEC 42102	Information technology — Artificial intelligence — Taxonomy of AI system methods and capabilities

bsi.

Annex 3 – useful links

Register for updates on the ISO/IEC 42001 launch!
www.bsigroup.com/en-GB/standards/isoiec-42001

ISO/JTC 1/SC 42 committee
www.iso.org/committee/6794475.html

CEN/JTC 21 committee
www.cencenelec.eu/areas-of-work/cen-cenelec-topics/artificial-intelligence

BSI's ART/1 committee
standardsdevelopment.bsigroup.com/committees/50281655

Become a standards maker
www.bsigroup.com/en-GB/about-bsi/uk-national-standards-body/how-to-get-involved-with-standards/become-a-standards-maker

UK Government AI whitepaper
www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper

AI Standards Hub
aistandardshub.org