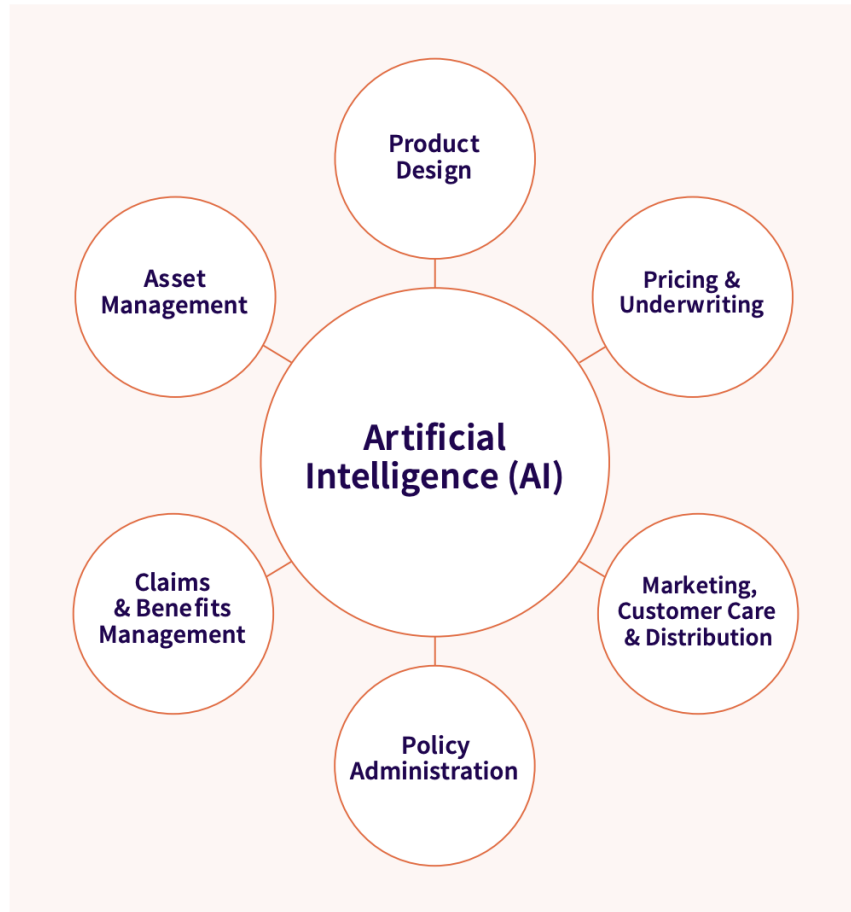# Implementing GenAI responsibly

–

An international perspective

Ethical AI Consulting I Aurelie Jacquet

AI & insurance

1. Safety, Security & Robustness

2. Appropriate Transparency & Explainability

3. Fairness

4. Accountability & Governance

5. Contestability & Redress

ABI's AI guide, practical ideas for getting started with responsible AI, February 2024

1 | Building bridges across international policy initiatives

What are ISO's international standards on AI?

25
Standards published

31
Standards under development

64
Countries involved

5
Working groups

>150
Experts involved

Liaisons with all major international initiatives (OECD, ITU, PAI,EC)

1
Pathway to certification: AI management system standard

NSW AI Assurance Framework

NSW gov AI projects to face new risk regime

By Justin Hendry
Dec 15 2021
12:18PM

0 Comments

From March 2022.

NSW government agencies will need to assess all projects that use bespoke artificial intelligence systems before deployment from March 2022 under a new framework aimed at identifying potential risks.

The AI assurance framework has been developed by the Department of Customer Service to assist agencies to design, build and use the technology "safely and securely".
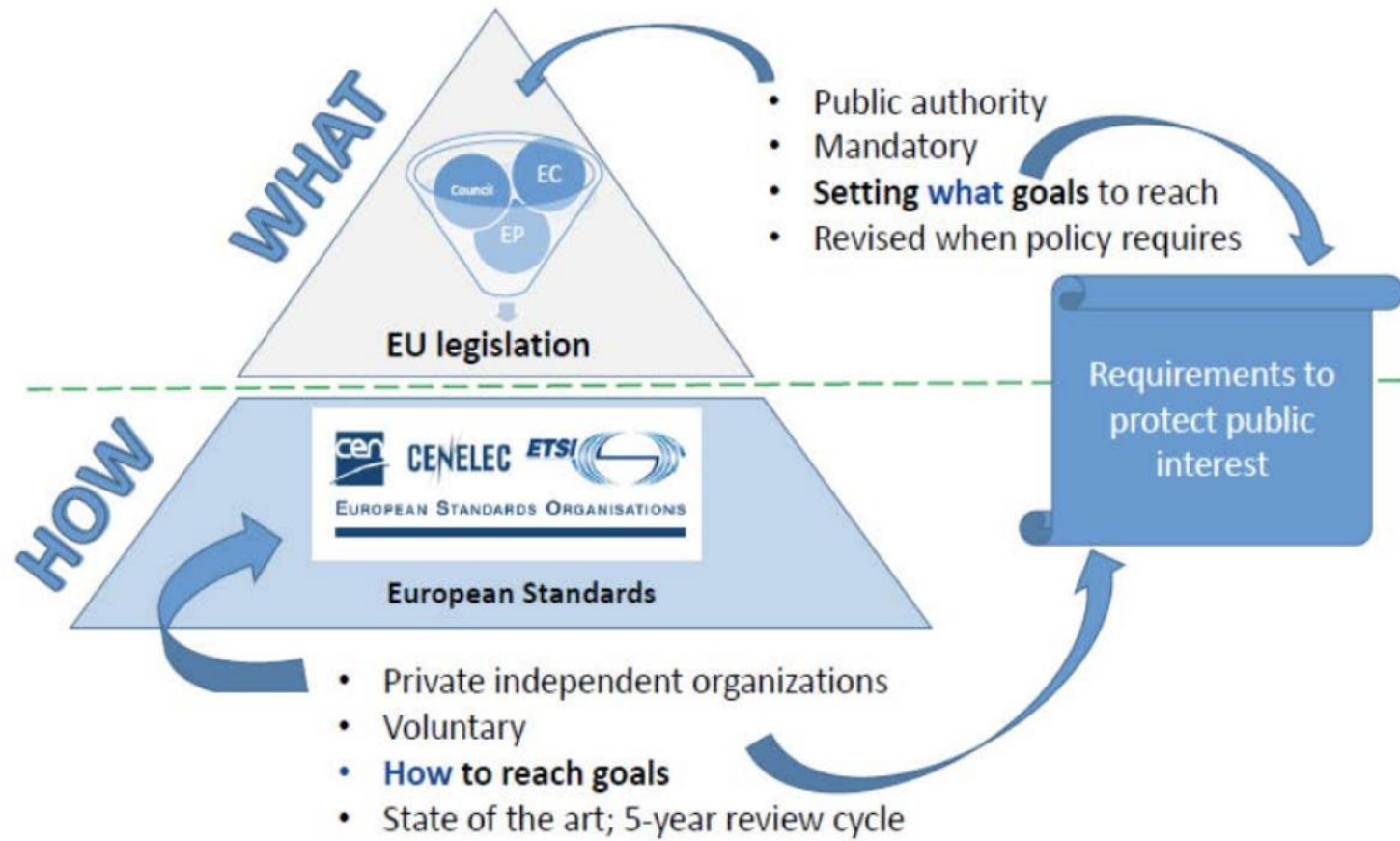
Mapping ISO to the work of NIST

The NIST roadmap for the AIMRMF cites alignment with international standards and production crosswalks to related standards as a top priority

NIST AI RMF to ISO/IEC FDIS 42001 AI Management system Crosswalk				
AI RMF		**ISO/IEC FDIS 42001**		
Govern 1.1	Legal and regulatory requirements involving AI are understood, managed, and documented.	4.1	Understanding the organization and its context	
		6.2	AI objectives and planning to achieve them	
		B.2.2	AI policy	
		B.2.4	Review of the AI policy	
Govern 1.2	The characteristics of trustworthy AI are integrated into organizational policies, processes, procedures, and practices.	B.9.3	Objectives for responsible use of AI system	
		B.6.1.2	Objectives for responsible development of AI system	
		B.6.1.3	Processes for responsible design and development of AI systems	
		B.10.3	Suppliers	
		B.2.2	AI policy	
		4.4	AI management system	
		5.2	AI Policy	
Govern 1.3	Processes, procedures, and practices are in place to determine the needed level of risk management activities based on the organization's risk tolerance.	6.1.2	AI risk assessment	
		6.1.1	General	
		6.1.3	AI risk treatment	

The EU AI ACT

- Public authority
- Mandatory
- **Setting what goals** to reach
- Revised when policy requires

WHAT

EU legislation

HOW

European Standards

- Private independent organizations
- Voluntary
- **How** to reach goals
- State of the art; 5-year review cycle

Requirements to protect public interest

2 | How Standards are helping

AI is not easy to develop

Helping getting started with AI

DATA	MODEL	ORGANISATION
What kind of data do I need to get started with AI?	*What do I need to know and what processes should I follow to develop and implement an AI model?*	*Where do I start for my organisation to better understand AI and consider AI initiatives?*
ISO/IEC 5259 series - Data quality for analytics and ML	ISO/IEC 23053:2022 – Framework for AI Systems using ML	ISO/IEC 22989:2022 – AI concepts and terminology
• 5259 -1 - Overview , terminology and examples	ISO/IEC 24372: 2021 - Overview of computational approaches	ISO/IEC DIS 5339 - Guidelines for AI applications
ISO/IEC FDIS 8183 - Data life cycle framework	ISO/IEC 5392 - Reference architecture of knowledge engineering	

AI can be dangerous

The Risks

Keeping AI systems on a tight leash using dynamic risk management:

- Identifying the acceptable range of variation for performance of AI systems
- monitoring
- Adapting controls as required

Managing Risk and ensuring appropriate oversight

DATA	MODEL	ORGANISATION
How do I know if my data is good enough to be processed by an AI system?	*How do I build safety functions?*	*What are the AI risks I should consider? How do I manage them?* *How can an organization have appropriate oversight of its AI systems?*

DATA

ISO/IEC 5259 series   Data quality for analytics and ML
5259 -2   Data quality measures
5259-3   Data quality requirements and guidelines
5259-4   Data quality process framework

MODEL

ISO/IEC TR 5469   Functional safety & AI systems
ISO/IEC TS 8200   Controllability of automated AI systems
ISO/IEC 24029   Assessment of the robustness of neural networks
ISO/IEC TS 29119-11   Testing of AI systems

ORGANISATION

ISO/IEC 23894:2023   Guidance on risk management
ISO/IEC 38507:2022   Governance implications of the use of AI by organisations
ISO/IEC 42001   Management system

The Ethics

From

"Can we do it?"

to

"Should we do it?"

Addressing Ethics

DATA	MODEL	ORGANISATION
How do I identify and manage algorithmic bias?	*How do I ensure AI systems are trustworthy?*	*What are the key ethical and societal concerns I need to consider? How can I address them?*
ISO/IEC TR 24027:2021 - Bias in AI systems and AI aided decision making	ISO/IEC TS 6254 - Objectives and approaches for explainability of ML models and AI systems	ISO/IEC TR 24368:2022 - Overview of ethical and societal concerns
ISO/IEC TS 12791 – Treatment of unwanted bias in classification and regression ML tasks	ISO/IEC 12792 – Transparency taxonomy of AI systems	ISO/IEC/ 42105 Guidance for human oversight of AI systems
	ISO/IEC TR 24028:2020 - Overview of Trustworthiness in AI	ISO/IEC 42005 – AI system impact assessment

Providing an AI Assurance Framework

AI Assurance Framework

DATA | MODEL | ORGANISATIONAL GOVERNANCE

AI Management system Standard

WG1: Foundational Standard

WG 2: Data

WG5: Computational Approaches | JWG1: Governance Implications

WG3: Trustworthiness | WG3: Trustworthiness

WG4: Use Cases & Specifications | WG4: Use Cases & Specifications

© Aurelie Jacquet 2021

© Ethical AI Consulting Pty Ltd, 2024
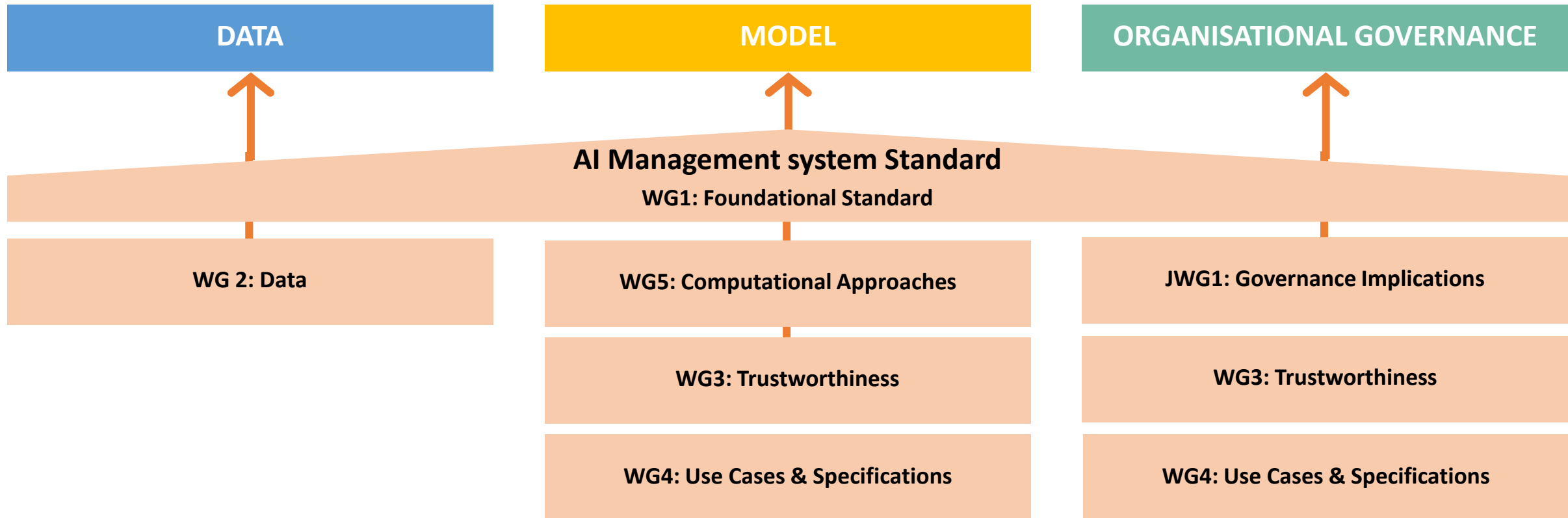
Standards' work relies on the 3Rs formula

Robust repeatable process + Recognised internationally = Scale AI Responsibly

Trust