October 23, 2024

Deploying AI Systems I: Obligations and Governance

Julian Flamant
Data Protection Counsel PLLC

Privacy+
Security
Forum

Speaker

Privacy+ Security Forum

Julian Flamant

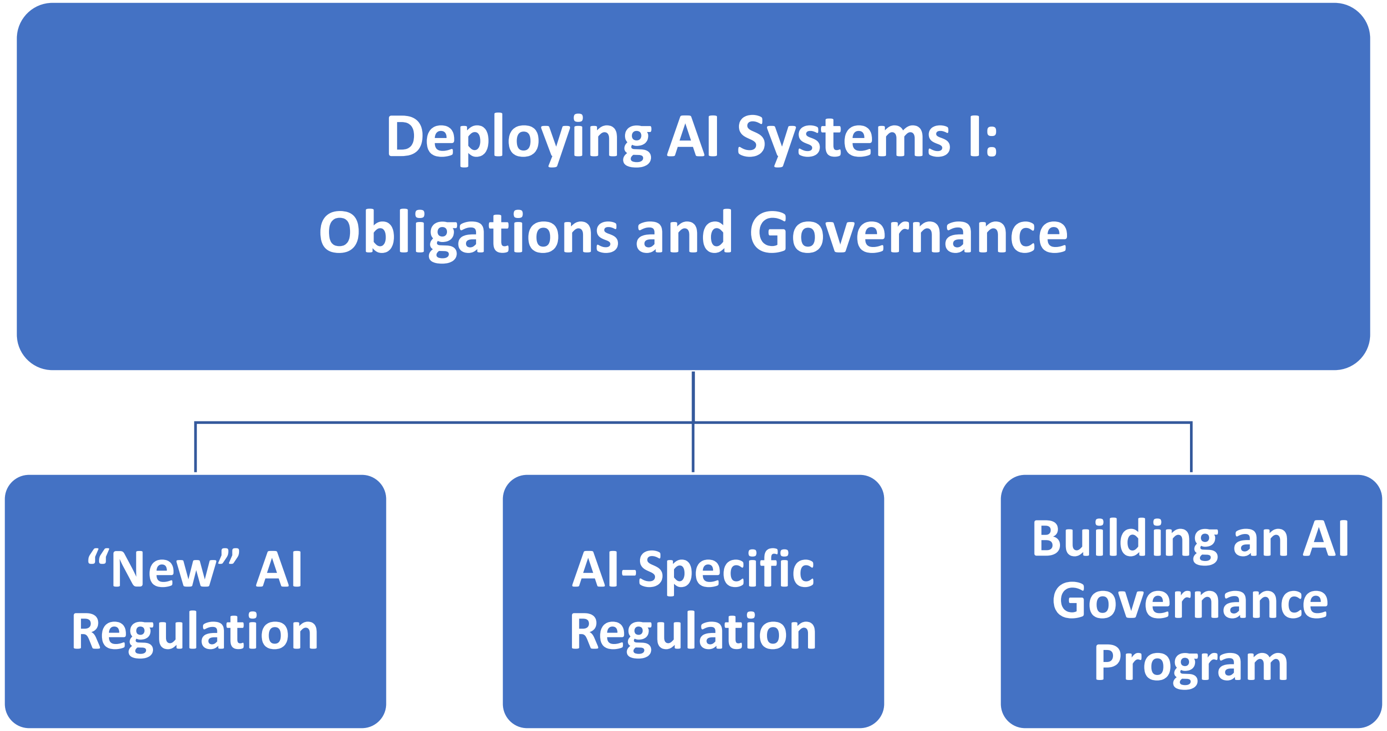
Principal

Data Protection Counsel PLLC

Julian is an experienced privacy, security, and AI attorney who provides strategic advising and compliance management services as principal of Data Protection Counsel PLLC. Julian previously was a senior associate in the Privacy and Cybersecurity group at Hogan Lovells and was one of the first policy fellows at the Future of Privacy Forum. Julian's law practice covers a broad range of data protection issues in the U.S. and internationally. He works with clients to develop practical solutions for addressing reputational, legal, and business risks.

Workshop Agenda

1:30pm - 2:45pm (presentation)



Deploying AI Systems I:
Obligations and Governance

"New" AI Regulation

AI-Specific Regulation

Building an AI Governance Program

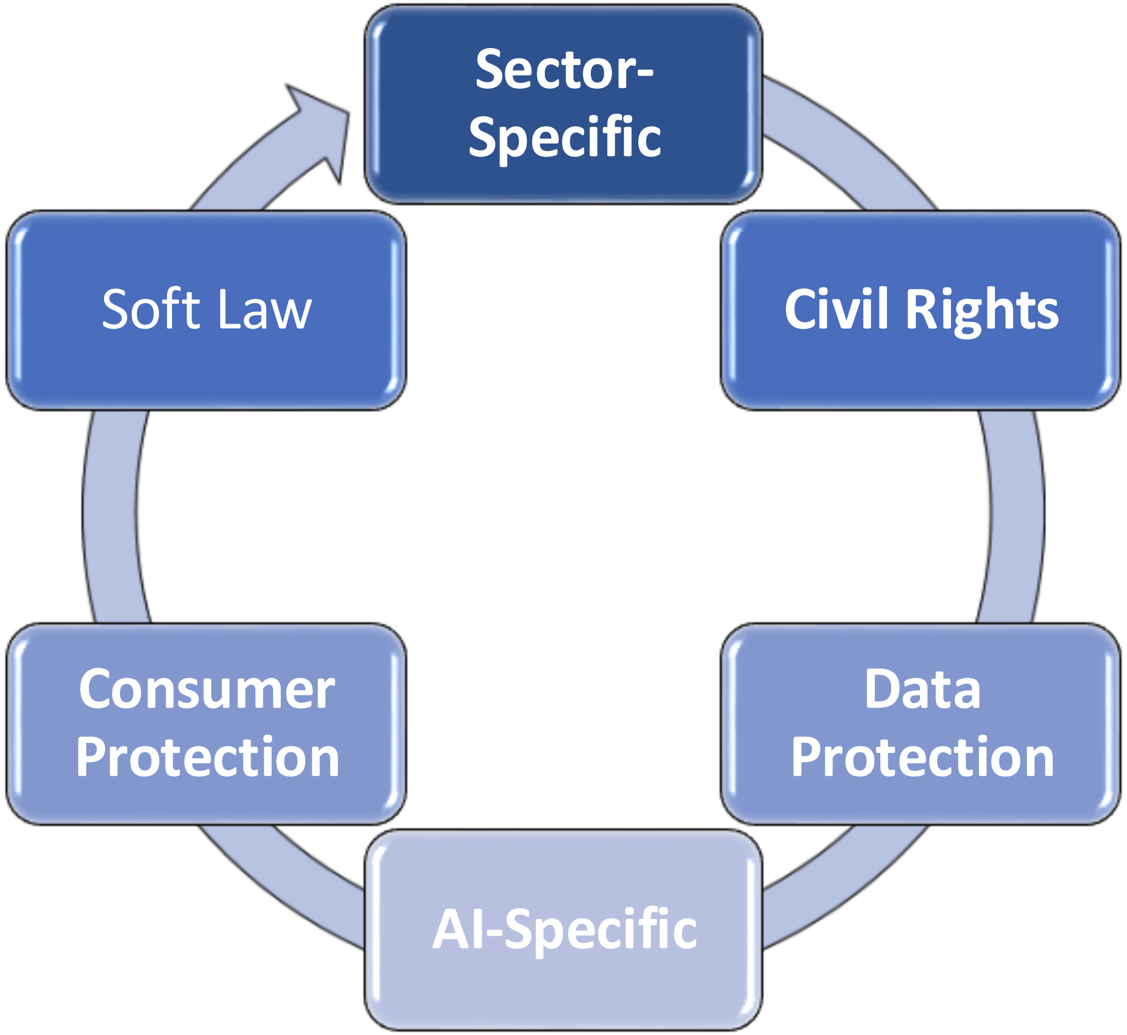
3:15pm – 4:30pm (panel discussion)

Deploying AI Systems II:
Practical Perspectives

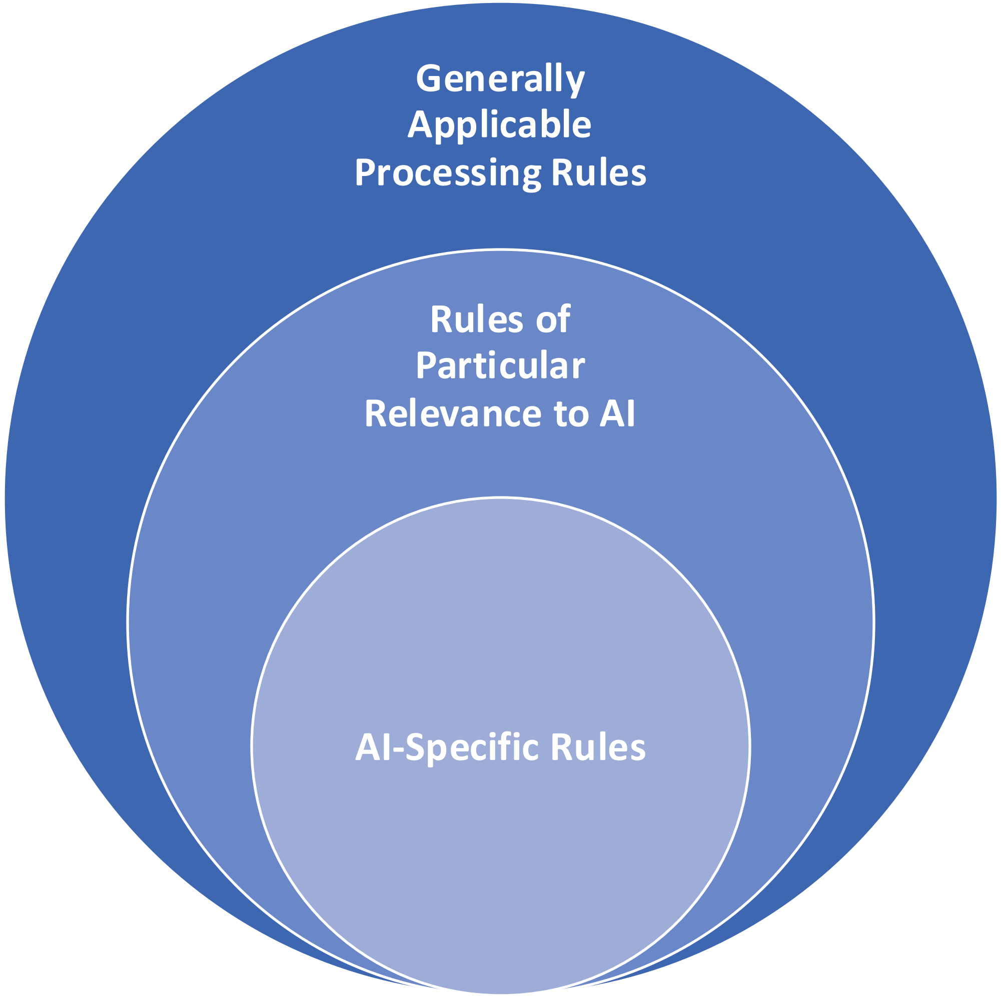
“New” AI Regulation

Privacy+
Security
Forum

Sources of Law Impacting AI Use



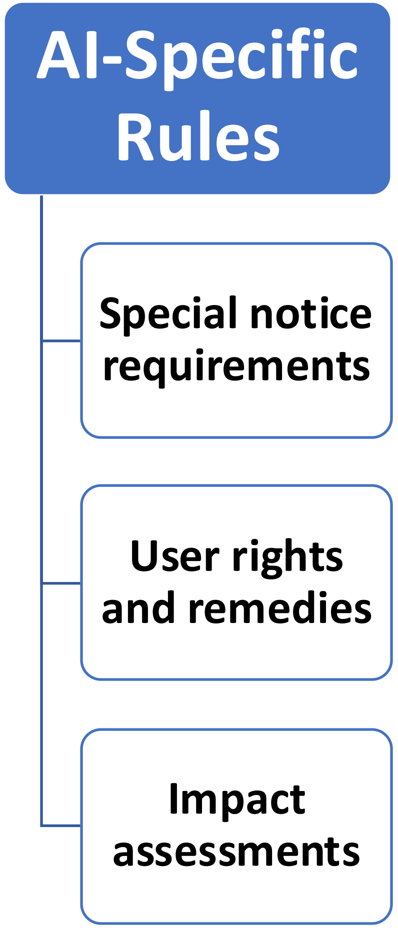
Data Protection Rules Impacting AI Use



Generally Applicable Processing Rules

Rules of Particular Relevance to AI

AI-Specific Rules

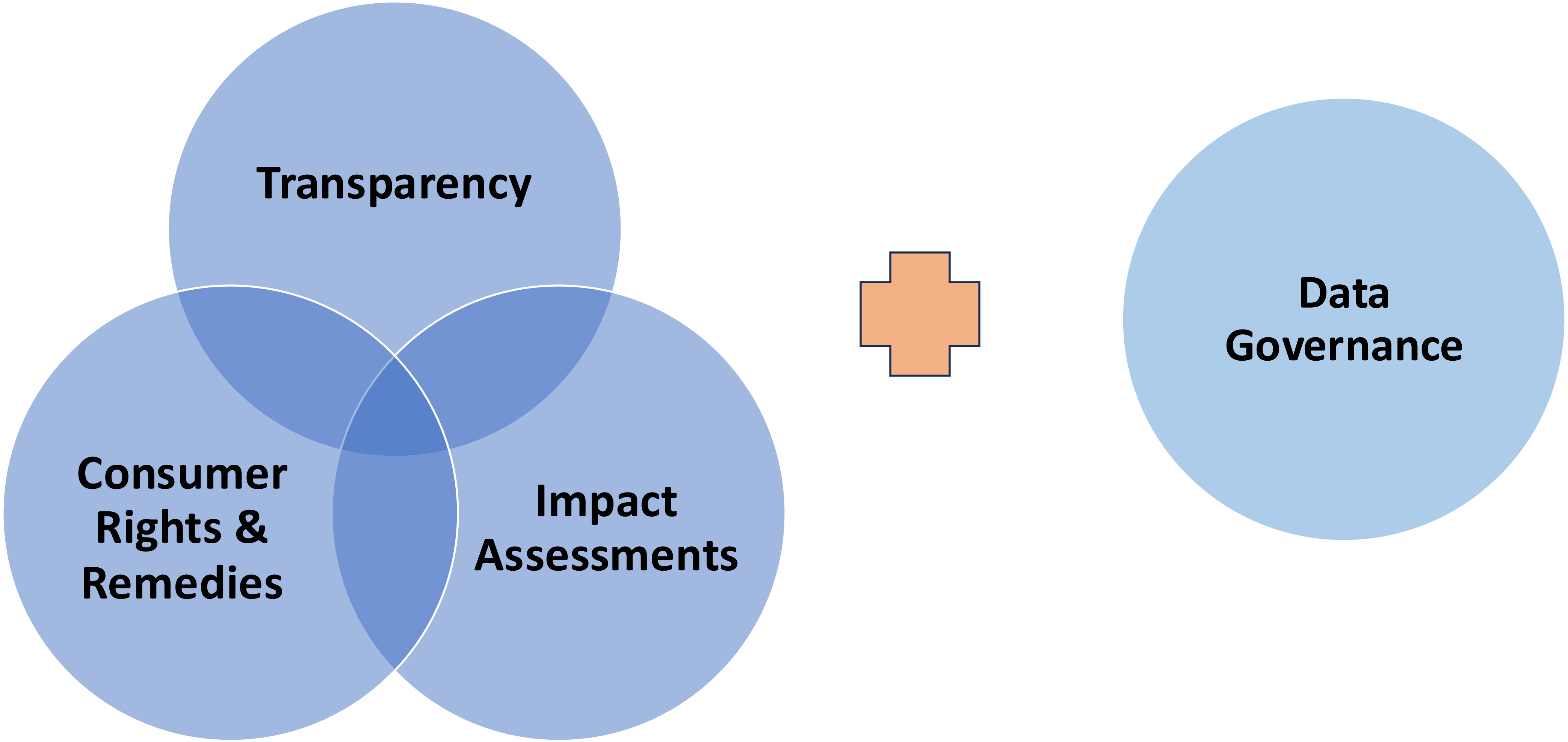


AI-Specific Rules

- Special notice requirements
- User rights and remedies
- Impact assessments

Overlap Between AI and Data Protection

Privacy+ Security Forum



AI-Specific Legislation

Colorado AI Act

- Effective: Feb. 2026
- First state comprehensive AI regulation*
- Focus on “high risk” AI systems
- Establishes duty of care to prevent algorithmic discrimination
- Establishes detailed documentation, monitoring, notice, and reporting obligations for deployers of high-risk AI
- Limited territorial and material scope

EU AI Act

- Tiered effective dates:
  - Prohibition of unacceptable risk AI (Feb. 2025)
  - Obligations for general purpose AI (Aug. 2025)
  - All rules become applicable (Aug. 2026)
- Risk-based approach
- Creates obligations according to role with more burden-sharing between AI actors
- Establishes documentation, monitoring, notice, and reporting obligations for deployers of high-risk AI
- Expansive territorial and material scope

AI-Specific Regulation

Deployer obligations for
high-risk AI

Privacy+
Security
Forum

Deployer Obligations (CAIA) 1/2

Duty of Care

Use reasonable care to protect consumer from any known or reasonably foreseeable risks of algorithmic discrimination.

Risk Management Policy

Internal policy governing high-risk AI uses and describing processes and personnel used to govern algorithmic discrimination.

Impact Assessment

Annual assessment detailing purpose, intended use, risk of algorithmic discrimination, steps to mitigate such risks, description of data used and produced, performance, transparency measures, and post-deployment monitoring.

Pre-Deployment Statement of Use

Consumer notice disclosing purpose of system, nature of the consequential decision, description of how the system assesses information to reach a decision, and sources of personal data processed.

Deployer Obligations (CAIA) 2/2

Privacy+ Security Forum

Publicly Available Statement

Public statement regarding the use of a high-risk AI system.

Data Privacy Rights

Right to opt-out of processing of certain profiling, right to correct inaccurate information.

Adverse Decision Rights

If adverse decision is made using high-risk AI system:

→ Explanation.

→ Opportunity to correct data.

→ Appeals procedure.

Attorney General Disclosures

Report to Attorney General any known or reasonably foreseeable risks of algorithmic discrimination (within 90 days).

Deployer Obligations (EU AI Act) 1/2

Avoid Misuse

Take appropriate technical and organizational measures, including human oversight, to ensure use of high-risk AI system is in accordance with the instructions for use.

Monitor the operation of the high-risk AI system on the basis of deployer instructions.

Prevent Harm

If reason to believe that using high-risk AI system in accordance with instructions may adversely affect individuals' health, safety or fundamental rights:

→ inform, without undue delay, the provider or distributor and the relevant market surveillance authority.

→ suspend the use of the high-risk AI system.

Report Serious Incidents

Immediately (within 15 days) report any "serious incident" to provider, importer, then market surveillance authority.

Fundamental Rights Impact Assessment

Assessment must consider the processes in which the high-risk AI system will be employed, the duration and frequency of its usage, the categories of individuals affected, the specific risks of harm, the measures for human oversight, and the actions to be taken if risks materialize.

Required prior to "first use" of certain high-risk AI systems.

Deployer Obligations (EU AI Act) 2/2

Privacy+ Security Forum

Input Data

Ensure that input data is relevant and sufficiently representative in view of the intended purpose of the system.

Retain logs of information generated by high-risk AI

Retain the logs automatically generated by the high-risk AI system, to the extent that such logs are within deployer's control, for a duration appropriate to the system's intended purpose but of at least six months, unless provided otherwise in applicable EU or national law.

Transparency

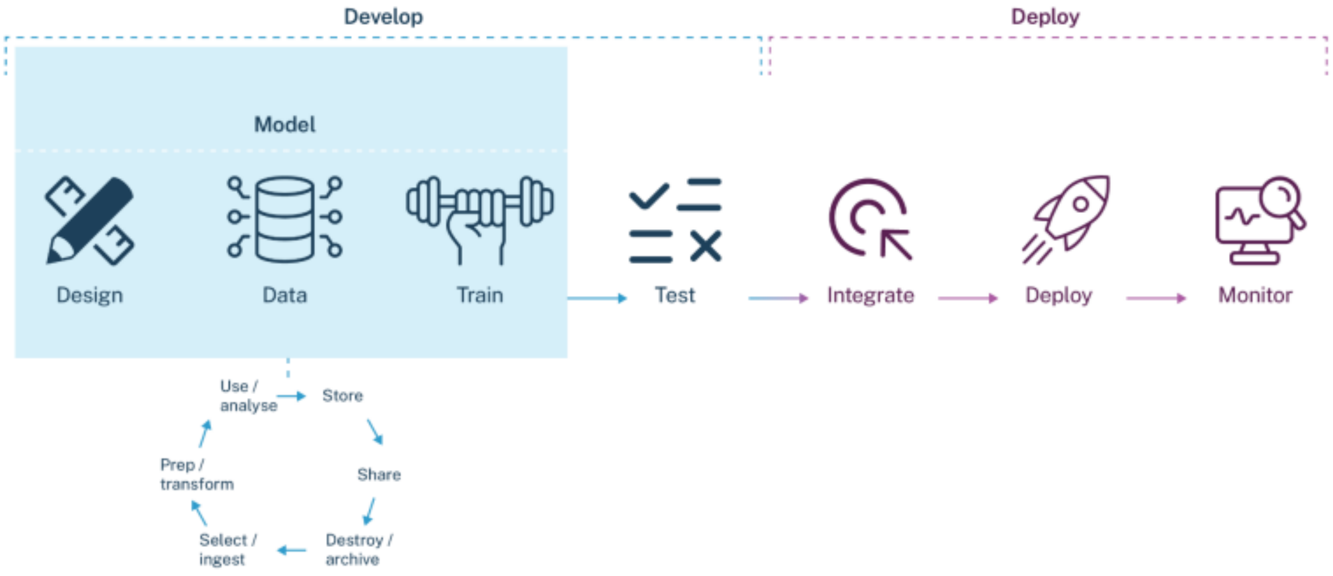
In certain contexts, inform persons that they are subject to the use of a high-risk AI system. (critical infrastructure, education and vocational training, employment, worker management, and access to self-employment)

Inform workers' representatives and affected workers that they will be subject to the use of a high-risk AI system.

Building an AI Governance Program

Privacy+ Security Forum

AI Lifecycle



Design

Define the problem, design specifications and success measures.

Data

Prepare data for training.

Train

Pre-train
Select an algorithm for training.

Fine tune
Select an existing model for fine-tuning.

Test

Test and evaluate model against success measures.

Integrate

Integrate model into IT systems to end-users.

Deploy

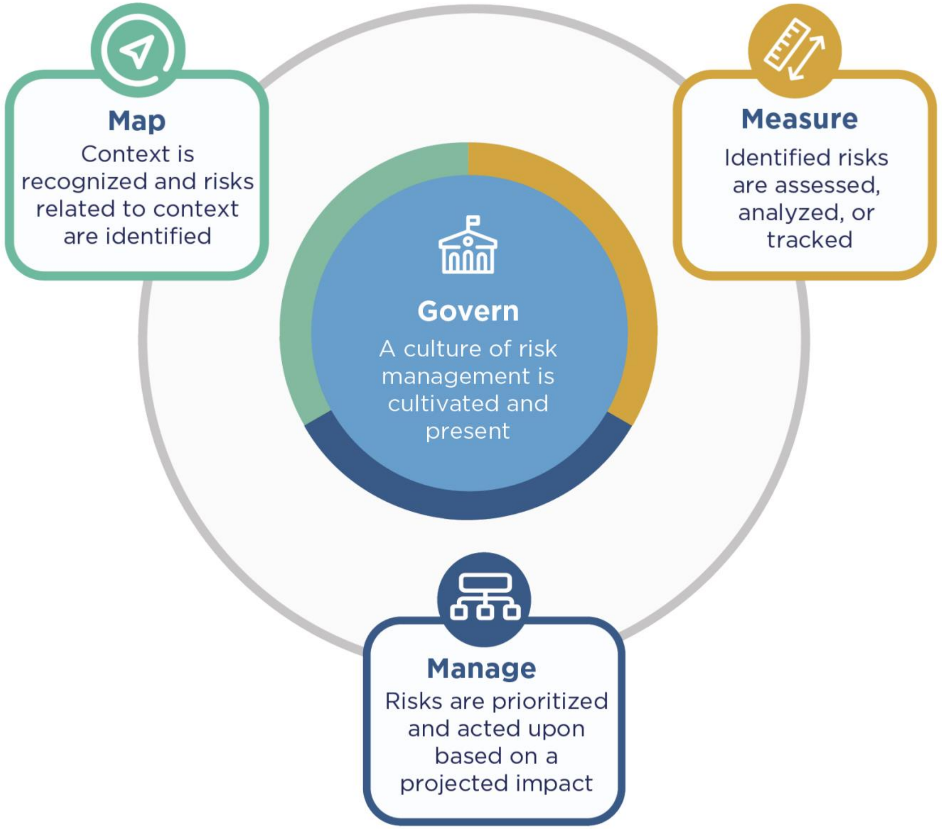
Deploy systems to end-users.

Monitor

Perform ongoing testing and monitoring.

Source: Department of Industry, Science and Resources, Safe and responsible AI in Australia: Proposals paper for introducing mandatory guardrails for AI in high-risk settings, September 2024.

NIST AI RMF – “Core” Overview



Map
Context is recognized and risks related to context are identified

Measure
Identified risks are assessed, analyzed, or tracked

Govern
A culture of risk management is cultivated and present

Manage
Risks are prioritized and acted upon based on a projected impact

Source: NIST AI 100-1, Artificial Intelligence Risk Management Framework (AI RMF 1.0)

NIST AI RMF – "Core" Functions 1/2

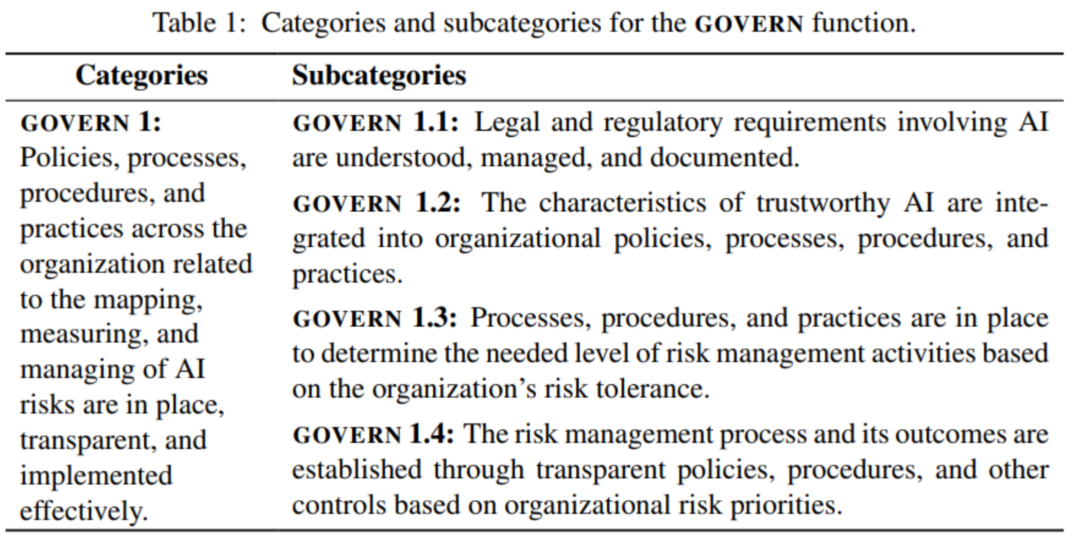


Table 1: Categories and subcategories for the GOVERN function.

Categories	Subcategories
GOVERN 1: Policies, processes, procedures, and practices across the organization related to the mapping, measuring, and managing of AI risks are in place, transparent, and implemented effectively.	**GOVERN 1.1:** Legal and regulatory requirements involving AI are understood, managed, and documented. **GOVERN 1.2:** The characteristics of trustworthy AI are integrated into organizational policies, processes, procedures, and practices. **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. **GOVERN 1.4:** The risk management process and its outcomes are established through transparent policies, procedures, and other controls based on organizational risk priorities.

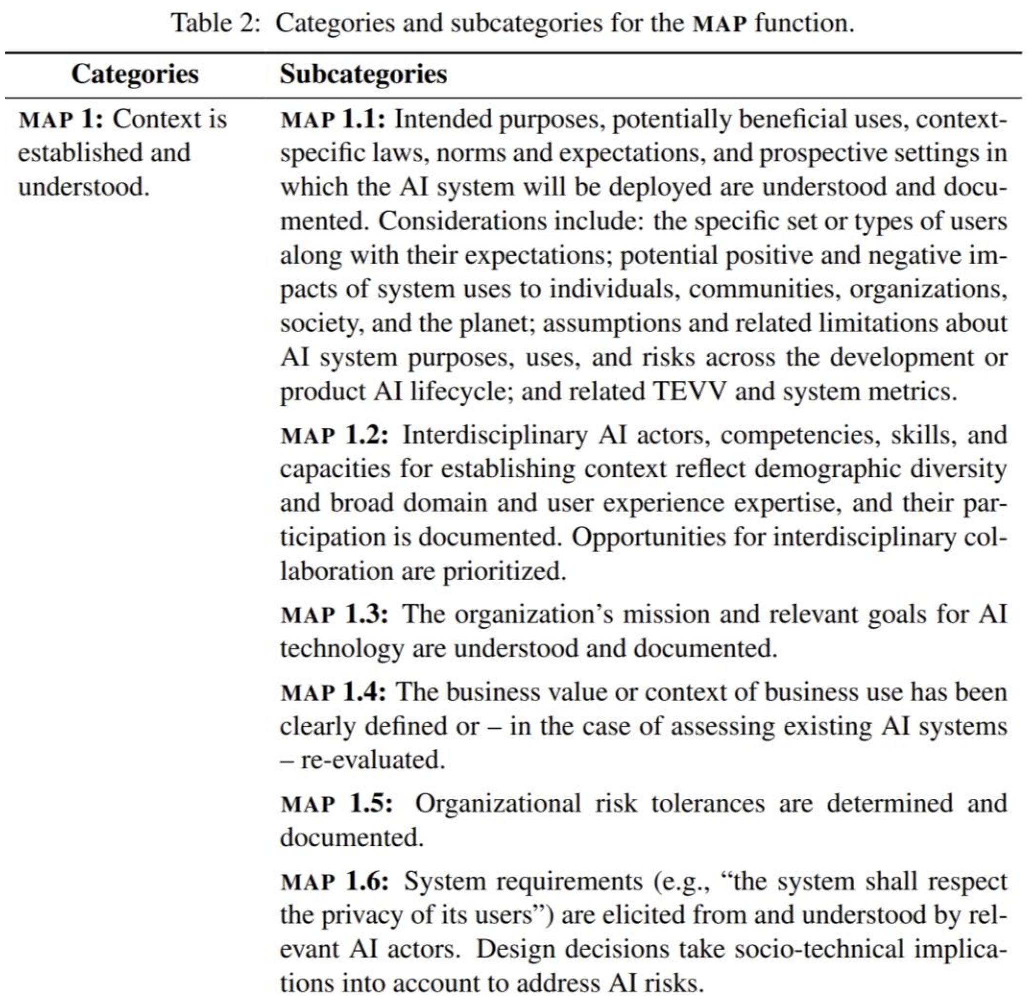


Table 2: Categories and subcategories for the MAP function.

Categories	Subcategories
MAP 1: Context is established and understood.	**MAP 1.1:** Intended purposes, potentially beneficial uses, context-specific laws, norms and expectations, and prospective settings in which the AI system will be deployed are understood and documented. Considerations include: the specific set or types of users along with their expectations; potential positive and negative impacts of system uses to individuals, communities, organizations, society, and the planet; assumptions and related limitations about AI system purposes, uses, and risks across the development or product AI lifecycle; and related TEVV and system metrics. **MAP 1.2:** Interdisciplinary AI actors, competencies, skills, and capacities for establishing context reflect demographic diversity and broad domain and user experience expertise, and their participation is documented. Opportunities for interdisciplinary collaboration are prioritized. **MAP 1.3:** The organization's mission and relevant goals for AI technology are understood and documented. **MAP 1.4:** The business value or context of business use has been clearly defined or – in the case of assessing existing AI systems – re-evaluated. **MAP 1.5:** Organizational risk tolerances are determined and documented. **MAP 1.6:** System requirements (e.g., "the system shall respect the privacy of its users") are elicited from and understood by relevant AI actors. Design decisions take socio-technical implications into account to address AI risks.

Source: NIST AI 100-1, Artificial Intelligence Risk Management Framework (AI RMF 1.0)

NIST AI RMF – "Core" Functions 2/2

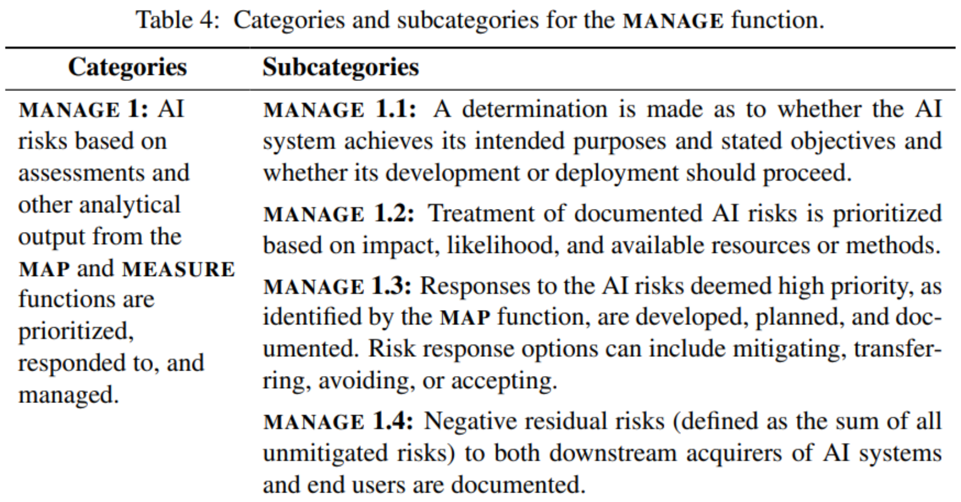


Table 4: Categories and subcategories for the **MANAGE** function.

Categories	Subcategories
MANAGE 1: AI risks based on assessments and other analytical output from the **MAP** and **MEASURE** functions are prioritized, responded to, and managed.	**MANAGE 1.1:** A determination is made as to whether the AI system achieves its intended purposes and stated objectives and whether its development or deployment should proceed. **MANAGE 1.2:** Treatment of documented AI risks is prioritized based on impact, likelihood, and available resources or methods. **MANAGE 1.3:** Responses to the AI risks deemed high priority, as identified by the **MAP** function, are developed, planned, and documented. Risk response options can include mitigating, transferring, avoiding, or accepting. **MANAGE 1.4:** Negative residual risks (defined as the sum of all unmitigated risks) to both downstream acquirers of AI systems and end users are documented.

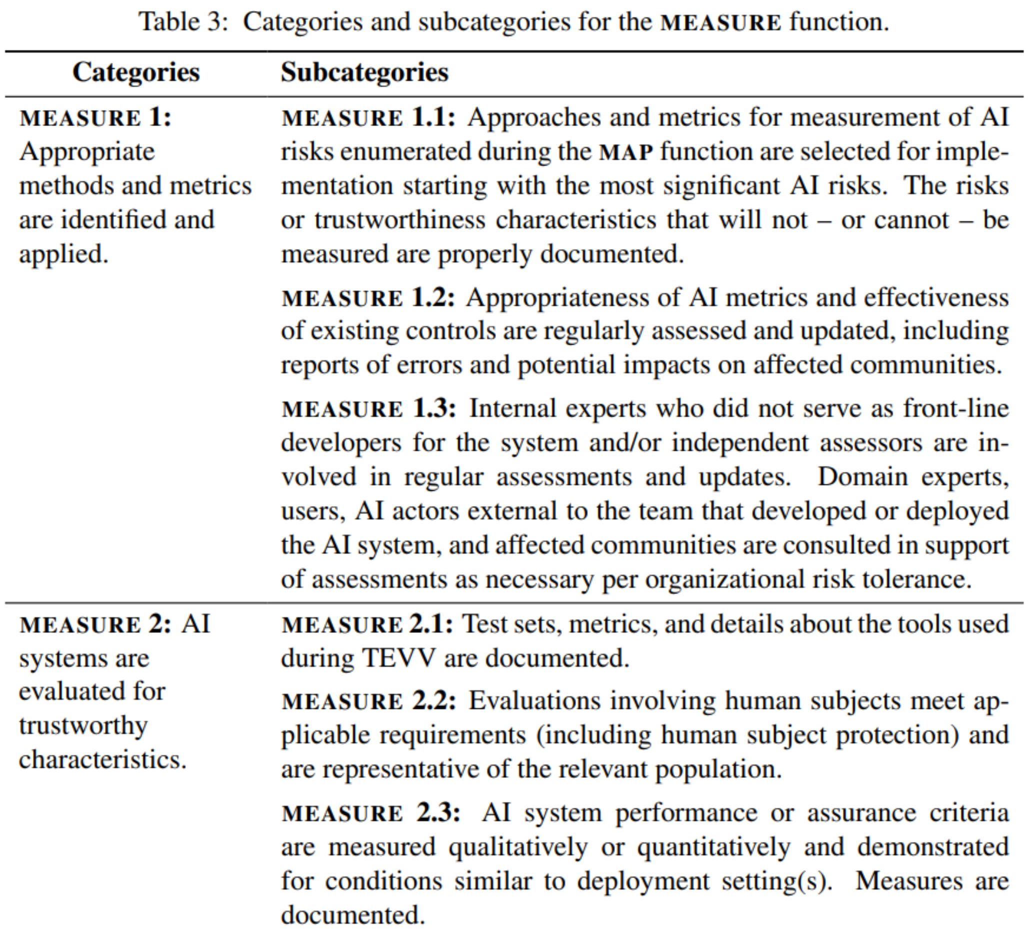


Table 3: Categories and subcategories for the **MEASURE** function.

Categories	Subcategories
MEASURE 1: Appropriate methods and metrics are identified and applied.	**MEASURE 1.1:** Approaches and metrics for measurement of AI risks enumerated during the **MAP** function are selected for implementation starting with the most significant AI risks. The risks or trustworthiness characteristics that will not – or cannot – be measured are properly documented. **MEASURE 1.2:** Appropriateness of AI metrics and effectiveness of existing controls are regularly assessed and updated, including reports of errors and potential impacts on affected communities. **MEASURE 1.3:** Internal experts who did not serve as front-line developers for the system and/or independent assessors are involved in regular assessments and updates. Domain experts, users, AI actors external to the team that developed or deployed the AI system, and affected communities are consulted in support of assessments as necessary per organizational risk tolerance.
MEASURE 2: AI systems are evaluated for trustworthy characteristics.	**MEASURE 2.1:** Test sets, metrics, and details about the tools used during TEVV are documented. **MEASURE 2.2:** Evaluations involving human subjects meet applicable requirements (including human subject protection) and are representative of the relevant population. **MEASURE 2.3:** AI system performance or assurance criteria are measured qualitatively or quantitatively and demonstrated for conditions similar to deployment setting(s). Measures are documented.

Source: NIST AI 100-1, Artificial Intelligence Risk Management Framework (AI RMF 1.0)

End of Session I

Privacy+ Security Forum

Questions?

Please return by 3:15pm for

Deploying AI System II: Practical Perspective

October 23, 2024

Deploying AI Systems II: Practical Perspectives

Privacy+ Security Forum

Speakers

Privacy+ Security Forum

Julian Flamant

Principal
Data Protection Counsel PLLC

Tatiana Rice

Deputy Director of U.S. Legislation
Future of Privacy Forum

Pollyanna Sanderson

Regulatory Compliance Lead for Privacy & AI
IBM

Lindsay Vogel

Lead Privacy Counsel
Bumble