UC BERKELEY
CENTER FOR LONG-TERM CYBERSECURITY

MAPPING OF THE AI Risk-Management Standards Profile for General-Purpose AI (GPAI) and Foundation Models V1.1

GUIDANCE TO KEY STANDARDS AND REGULATIONS

ANTHONY M. BARRETT | JESSICA NEWMAN | BRANDIE NONNECKE | NADA MADKOUR
DAN HENDRYCKS | EVAN R. MURPHY | KRYSTAL JACKSON | DEEPIKA RAMAN

Cover art: The cover image is an adaptation of a photograph titled, Steam Engine near the Grand Transept, Crystal Palace, taken by the photographer Philip Henry Delamotte in 1851. The impact of artificial intelligence and especially general purpose artificial intelligence is often compared to the impact of the steam engine during the Industrial Revolution, which brought enormous economic gains, but also dangerous workplaces and horrible living conditions for many. The Crystal Palace housed the Great Exhibition of 1851, where examples of technology developed in the Industrial Revolution were put on display for thousands of people to see. While enjoyed by many, the Crystal Palace was also critiqued for representing a false utopia. Similarly, the rise of general purpose AI is often discussed with utopian visions, but such positive visions will not be possible without the establishment of meaningful risk management strategies. The image is a reminder of the entanglement of people and machines, and the profound and lasting impact of general purpose technologies on society.

MAPPING OF THE

AI Risk-Management Standards Profile for General-Purpose AI (GPAI) and Foundation Models V1.1

GUIDANCE TO KEY STANDARDS AND REGULATIONS

ANTHONY M. BARRETT[†] • **JESSICA NEWMAN**[†] • **BRANDIE NONNECKE**[††] • **NADA MADKOUR**[†]
DAN HENDRYCKS[†††] • **EVAN R. MURPHY**[†] • **KRYSTAL JACKSON**[†] • **DEEPIKA RAMAN**[†]

† AI Security Initiative, Center for Long-Term Cybersecurity, UC Berkeley
†† CITRIS Policy Lab, CITRIS and the Banatao Institute; Goldman School of Public Policy, UC Berkeley
††† Berkeley AI Research Lab, UC Berkeley
All affiliations listed are either current, or were during main contributions to this work or a previous version.

Adapting material in the full Profile (Barrett et al. 2025).
For the full AI Risk-Management Standards Profile for General-Purpose AI (GPAI) and Foundation Models V1.1, see:
https://cltc.berkeley.edu/publication/ai-risk-management-standards-profile-v1-1

CLTC Center for Long-Term Cybersecurity UC Berkeley
AI SECURITY INITIATIVE
CITRIS and the BANATAO INSTITUTE | CITRIS POLICY LAB

Contents

1 The Profile V1.1 and its supporting documents were drafted and finalized prior to the recession of Executive Order 14110 on the Safe, Secure, and Trustworthy AI on January 20, 2025.

Introduction

For users of the AI Risk-Management Standards Profile for General-Purpose AI (GPAI) and Foundation Models (Barrett et al. 2025) who are working with AI risk management-related standards, codes of conduct, and regulations other than the NIST AI RMF, this document provides mappings, or "crosswalks," showing how guidance in the main Profile relates to clauses in those other standards and regulations. Using the best-practices guidance and resources in this document can help users of the Profile achieve conformity with those standards or regulations.

1. Mapping to ISO/IEC 23894

In this section, we provide mapping of Profile guidance to key clauses in ISO/IEC 23894:2023, "Information technology — Artificial intelligence — Guidance on risk management." This is based in part on the NIST draft crosswalk between the AI RMF 1.0 and ISO/IEC 23894 draft international standard (NIST 2023).

Table 1: Mapping to ISO/IEC 23894 Clauses

ISO/IEC 23894 Clause	NIST AI RMF Functions, Categories, or Subcategories with the most relevant guidance in this Profile
5.2 Leadership and commitment	Govern 1, 4
5.3 Integration	Govern
5.4 Design	Govern
5.4.1 Understanding the organization and its context	Map 1 Govern Measure
5.4.2 Articulating risk management commitment	Govern
5.4.3 Assigning organizational roles, authorities, responsibilities, and accountabilities	Govern 2
5.4.4 Allocating resources	Govern 1, 2
5.4.5 Establishing communication and consultation	Govern
5.5 Implementation	Manage
5.6 Evaluation	Measure 2.13, 3, 4
5.7 Improvement	Govern Measure Manage

ISO/IEC 23894 Clause	NIST AI RMF Functions, Categories, or Subcategories with the most relevant guidance in this Profile
6.2 Communication and consultation	Govern 2, 4, 5 Map 5.2
6.3.2 Defining the scope	Map 1
6.3.3 External and internal context	Map 1
6.3.4 Defining risk criteria	Map 1.5, 5 Measure Manage 1.1
6.4.2 Risk identification	Map 1.1, 5
6.4.2.3 Identification of risk sources	Map
6.4.2.4 Identification of potential events and outcomes	Map 5.1
6.4.2.5 Identification of controls	Map Measure Manage
6.4.2.6 Identification of consequences	Map 5.1
6.4.3 Risk analysis	Map Measure
6.4.3.2 Assessment of consequences	Map 5.1 Measure
6.4.3.3 Assessment of likelihood	Map 5.1 Measure
6.4.4 Risk evaluation	Map Measure Manage
6.5 Risk treatment	Manage
6.5.2 Selection of risk treatment options	Map 1.5 Manage 1
6.5.3 Preparing and implementing risk treatment plans	Manage 2
6.6 Monitoring and review	Measure Manage 4
6.7 Recording and reporting	Govern 4 Map Measure Manage 4

2. Mapping to White House AI Commitments

In this section, we provide mapping of Profile guidance to the code of conduct represented by the commitments announced with the White House (2023) by several frontier model developers. These commitments apply to developing and releasing foundation models more capable than the July 2023 industry frontier.

Table 2: Mapping to White House AI Commitments

White House AI Commitments	NIST AI RMF Functions, Categories, or Subcategories with the most relevant guidance in this Profile
1) Commit to internal and external red teaming of models or systems in areas including misuse, societal risks, and national security concerns, such as bio, cyber, and other safety areas.	Govern 1.5, 5.1 Map 2.3, 5.1 Measure 1.1, 1.3, 2 Manage 1.3, 2.4
3) Invest in cybersecurity and insider threat safeguards to protect proprietary and unreleased model weights.	Measure 2.7 Manage 1.3
4) Incent third-party discovery and reporting of issues and vulnerabilities.	Govern 4, 5 Map 5.2 Measure 3.3 Manage 4
5) Develop and deploy mechanisms that enable users to understand if audio or visual content is AI-generated, including robust provenance, watermarking, or both, for AI-generated audio or visual content.	Measure 2.7, 2.8 Manage 1.3, 4
6) Publicly report model or system capabilities, limitations, and domains of appropriate and inappropriate use, including discussion of societal risks, such as effects on fairness and bias.	Govern 4 Map 1.5 Manage 1.3
7) Prioritize research on societal risks posed by AI systems, including on avoiding harmful bias and discrimination, and protecting privacy.	Govern 2.3 Measure 1

3. Mapping to Executive Order 14110 Section 4.2

In this section, we provide mapping of Profile guidance to portions of Section 4.2 of Executive Order 14110[2] (Biden 2023) that seem most relevant to GPAI/foundation model developers, especially to developers of dual-use foundation models (i.e., frontier models) in the United States. Section 4.2 of the Executive Order includes reporting requirements for developers of dual-use foundation models, including reporting for cybersecurity protections and results of red-team evaluations of dual-use model capabilities (e.g., in CBRN and cyber domains).

Most other sections of the Executive Order do not directly apply to GPAI/foundation model developers. Instead, they provide directives to US federal agencies (e.g., NIST), including directives to develop detailed guidance for GPAI/foundation model developers. We list neither those other sections nor those directives to federal agencies. However, we have added resulting guidance as resources in Section 3 of the Profile as available, e.g., we listed NIST AI 800-1 ipd (NIST 2024) as a resource for red-teaming under Measure 1.1.

Table 3: Mapping to Executive Order 14110 Section 4.2

Executive Order 14110 Sections	NIST AI RMF Functions, Categories, or Subcategories with the most relevant guidance in this Profile
§4.2(a)(i)(B) on measures to protect model weights	Measure 2.7
§4.2(a)(i)(C) on red-team evaluations of dual-use model capabilities (e.g., in CBRN and cyber domains), potential for influence, and possibility for self-replication or propagation; and associated measures to meet safety objectives	Measure 1.1, 1.3, 2.7 Manage

2 The Profile V1.1 and its supporting documents were drafted and finalized prior to the recession of Executive Order 14110 on the Safe, Secure, and Trustworthy AI on January 20, 2025.

4. Mapping to Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems

In this section, we provide mapping of Profile guidance to the International Code of Conduct for Organizations Developing Advanced AI Systems, which was developed by the G7 Digital and Tech Ministers (G7 2023). This guidance aims to promote safe, secure, and trustworthy AI worldwide and provides voluntary guidance for organizations developing the most advanced AI systems, including foundation models and generative AI systems, and builds upon the existing OECD AI Principles (G7 2023).

Table 4: Mapping to the International Code of Conduct for Organizations Developing Advanced AI Systems

International Code of Conduct for Organizations Developing Advanced AI Systems Actions	NIST AI RMF Functions, Categories, or Subcategories with the most relevant guidance in this Profile
1) Take appropriate measures throughout the development of advanced AI systems, including prior to and throughout their deployment and placement on the market, to identify, evaluate, and mitigate risks across the AI lifecycle.	Govern
2) Identify and mitigate vulnerabilities, and, where appropriate, incidents and patterns of misuse, after deployment, including placement on the market.	Govern 6.2 Map 2.3, 3.2, 3.4, 4.2, 5.1, 5.2 Measure 1.1, 2.7 Manage 2.2, 2.3, 2.4, 4.1, 4.2, 4.3
3) Publicly report advanced AI systems' capabilities, limitations, and domains of appropriate and inappropriate use, to support ensuring sufficient transparency, thereby contributing to increased accountability.	Govern 1.2, 2.1, 4.2, 4.3 Map 1.1, 1.3, 1.5, 2.1, 2.2, 2.3, 3, 4.1, 4.2, 5.1 Measure 1.1, 1.2, 2.1, 2.3, 2.5, 2.7-2.13, 3.1, 4.2, 4.3 Manage 2.3, 2.4
4) Work towards responsible information sharing and reporting of incidents among organizations developing advanced AI systems, including with industry, governments, civil society, and academia.	Govern 1.2, 2.1, 4.2, 4.3, 5 Map 1.1, 1.3, 1.5, 1.6, 2.1, 2.2, 2.3, 3.1, 3.2, 4.2 Measure 1.1, 1.2, 1.3, 2.7, 2.8, 2.9, 3.1, 4.1, 4.2, 4.3 Manage 2.4
5) Develop, implement and disclose AI governance and risk management policies, grounded in a risk-based approach – including privacy policies, and mitigation measures.	Govern Map 1 Measure 1 Manage 1, 2, 4
6) Invest in and implement robust security controls, including physical security, cybersecurity and insider threat safeguards across the AI lifecycle.	Govern 1.1, 1.2, 1.5, 1.7, 2.1, 4.1, 6.2 Map 1.5, 4.2, 5.1 Measure 1.1, 2.7, 2.10 Manage 1.1, 1.3, 4

International Code of Conduct for Organizations Developing Advanced AI Systems Actions	NIST AI RMF Functions, Categories, or Subcategories with the most relevant guidance in this Profile
7) Develop and deploy reliable content authentication and provenance mechanisms, where technically feasible, such as watermarking or other techniques to enable users to identify AI-generated content.	Govern 1.2, 4.2 Map 2.1 Measure 2.7, 2.9
8) Prioritize research to mitigate societal, safety and security risks and prioritize investment in effective mitigation measures.	Govern 3.1
9) Prioritize the development of advanced AI systems to address the world's greatest challenges, notably but not limited to the climate crisis, global health and education.	Govern 1.2, 5.2 Map 3.1, 5

5. Mapping to EU AI Act

In this section, we provide mapping of the Profile guidance to relevant provisions represented by the EU AI Act articles (EP 2024). The EU AI Act includes regulatory requirements for GPAI models, GPAI models with systemic risk, and high-risk AI systems. At a minimum, GPAI models (or at least one or more subsets of GPAI models identified to have high-impact capabilities) are subject to requirements for transparency, and for assessing, mitigating, and documenting several types of reasonably foreseeable risks.

Table 5: Mapping to EU AI Act Provisions

EU AI Act Provisions (with relevance to GPAI/foundation models)	NIST AI RMF Functions, Categories, or Subcategories with the most relevant guidance in this Profile
Chapter III: High-Risk AI System	
Section 1: Classification of AI Systems as High-Risk	
Article 6: Classification Rules for High-Risk AI Systems	Map 1.1, 1.3, 2, 3.1, 3, 4, 5.1
Section 2: Requirements for High-Risk AI Systems	
Article 8: Compliance with Requirements	Govern 1.1, 1.2, 1.3, 1.4 Map 1.1, 1.3 Measure 2.1 Manage 2.3
Article 9: Risk Management System	Map Measure Manage
Article 10: Data and Governance	Map 1.1,1.6, 2.1, 2.3, 3.3, 4.1 Measure 2.1, 2.2, 2.5, 2.11, 2.13, 3.1, 4.2
Article 11: Technical Documentation	Map 1.1, 1.5, 1.6, 2, 3, 4 Measure 1, 2.1, 2.2, 2.3, 2.4, 2.6, 2.7, 2.8, 2.9, 2.13, 3, 4 Manage 2.2, 3.1, 4
Article 13: Transparency and Provision of Information Deployers	Map 1.1, 1.3, 1.6, 2, 3 Measure 2.1, 2.3, 2.5, 2.7, 2.8, 2.9, 2.10, 2.11, 2.12, 2.13, 4 Manage 1.3, 1.4, 4.3
Article 14: Human Oversight	Govern 2.1, 2.2, 3, 5 Map 1.2, 2.2, 3.5, 5.2 Measure 1.3, 2, 3.1, 3.3, 4.1, 4.2 Manage 2.4, 4.1, 4.3
Article 15: Accuracy Robustness and Cybersecurity	Govern 1.2, 4.1 Map 2.3 Measure 1.1, 2, 4.2 Manage 4.1

EU AI Act Provisions (with relevance to GPAI/foundation models)	NIST AI RMF Functions, Categories, or Subcategories with the most relevant guidance in this Profile
Section 3: Obligations of Providers and Deployers of High-Risk AI Systems and Other Parties	
Article 16: Obligations of Providers of High-Risk AI Systems	Govern 1 Map 2, Measure 1, 2, 3, 4 Manage 1, 2.2, 2.4, 4
Article 17: Quality Management System	Govern 1.5, 1.6, 1.7, 3.1, 3.2, 4.2, 4.3, 5.1, 5.2 Map 1.1, 2.1, 2.2, 2.3, 3.4, 3.5, 4.1, 4.2 Measure 1.1, 1.2, 1.3, 2.1, 2.2, 2.3, 2.4, 2.6, 2.7, 2.9, 2.13, 3.1, 3.2, 4.1, 4.2, 4.3 Manage 1.1, 1.2, 1.3, 2.2, 2.2, 2.4, 3.2, 4.1, 4.2, 4.3
Article 18: Documentation Keeping	Map 1.1, 1.5, 1.6, 2, 3, 4 Measure 1, 2.1, 2.2, 2.3, 2.4, 2.6, 2.7, 2.8, 2.9, 2.13, 3, 4 Manage 2.2, 3.1, 4
Article 20: Corrective Actions and Duty of Information	Manage 2
Article 24: Obligations of Distributors	Govern 1.1, 1.2, 1.3, 1.4 Map 1.1, 1.3 Measure 2.1 Manage 2.3
Article 25: Responsibilities along the AI Value Chain	Govern 2.1, 6 Map 2.2, 3.4, 4.1, 4.2 Manage 1.4
Article 26: Obligations of Deployers of High-Risk AI Systems	Govern 1, 2.2, 4.2, 4.3 Map 2.2, 2.3, 3.4, 3.5 Measure 2.7, 2.10, 2.13 Manage 2.4, 4.3
Article 27: Fundamental Rights Impact Assessment for High-Risk AI Systems	Govern 3.2 Map 1.1, 1.3, 1.6, 2.2, 2.3, 3.3, 3.5, 4.1, 5 Measure 2, 3.3, 4.1 Manage 1, 2.4
Section 5: Standards, Conformity Assessment, Certificates, Registration	
Article 43: Conformity Assessment	Govern 1.1, 4.1
Chapter IV: Transparency Obligations for Providers and Deployers of Certain AI Systems and GPAI Models	
Article 50: Transparency Obligations for Providers and Users of Certain AI Systems and GPAI Models	Govern 1.2 Map 2.1, 2.3 Measure 2.8, 2.7 Manage 1.3, 4.3

EU AI Act Provisions (with relevance to GPAI/foundation models)	NIST AI RMF Functions, Categories, or Subcategories with the most relevant guidance in this Profile
Chapter V: General Purpose AI Models	
Section 1: Classification Rules	
Article 51: Classification of General-Purpose AI Models as General Purpose AI Models with Systemic Risk	Map 1.1, 2, 3.3 Measure 2.1, 2.9
Section 2: Obligations for Providers of General Purpose AI Models	
Article 53: Obligations for Providers of General Purpose AI Models	Govern 1.1, 1.3, 1.4, 1.6, 4.2, 4.3, Map 1.1, 1.5, 1.6, 2, 3.1, 3.2, 3.5, 4.2, 5.2 Measure 1.1, 1.2, 1.3, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.8, 2.9, 2.10, 2.12, 2.13, 3.1, 3.2, 4.1, 4.2, 4.3 Manage 1.1, 1.2, 1.3, 1.4, 2.2, 2.3, 2.4, 4.1, 4.2, 4.3
Section 3: Obligations for Providers of General Purpose AI Models with Systemic Risk	
Article 55: Obligations for Providers of General-Purpose AI Models with Systemic Risk	Govern 1.1, 1.3, 1.4, 1.5, 1.6, 1.7, 3.1, 4.2, 4.3, 5.2 Map 1.1, 2, 3.4, 3.5, 4, 5.1 Measure Manage 1, 2.2, 2.3, 2.4, 4
Article 56: Codes of Practice	Govern 2, 5.2 Map 1, 2, 3, 4, 5.1 Measure 1, 2, 4.1 Manage (See also our more detailed mapping in the next subsection's table, specifically for Article 56.)
Chapter IX: Post-Market Monitoring, Information Sharing, Market Surveillance	
Section 1: Post-Market Monitoring	
Article 72: Post-Market Monitoring by Providers and Post-Market Monitoring Plan for High-Risk AI Systems	Manage 2, 3, 4
Article 73: Reporting of Serious Incidents	Manage 2.3, 4.1, 4.3
Article 78: Confidentiality	Govern 6.1 Map 4.1 Measure 2.10

6. Mapping to EU AI Act Article 56 on Codes of Practice for Providers of General Purpose AI Models with Systemic Risk

In this section, we provide mapping of the Profile to the EU AI Act Article 56 Codes of Practice for Providers of General Purpose AI Models with Systemic Risk. Articles 53 and 55 were also included in the mapping.

As an organizing approach, we structure this mapping as an extension of mapping to the International Code of Conduct for Organizations Developing Advanced AI Systems, which was developed by the G7 Digital and Tech Ministers (G7 2023).

Table 6: Mapping to International Code of Conduct for Organizations Developing Advanced AI Systems and the EU AI Act Article 56

International Code of Conduct for Organizations Developing Advanced AI Systems Actions	NIST AI RMF Functions, Categories, or Subcategories with the most relevant guidance in this Profile	EU AI Act Article 56 Codes of Practice for Providers of General Purpose AI Models with Systemic Risk
1) Take appropriate measures throughout the development of advanced AI systems, including prior to and throughout their deployment and placement on the market, to identify, evaluate, and mitigate risks across the AI lifecycle.	Govern	Article 56: 2(c), 2(d), Article 55: 1(a), 1(b), 1(d)
2) Identify and mitigate vulnerabilities — and, where appropriate, incidents and patterns of misuse — after deployment, including placement on the market.	Govern 6.2 Map 2.3, 3.2, 3.4, 4.2, 5.1, 5.2 Measure 1.1, 2.7 Manage 2.2, 2.3, 2.4, 4.1, 4.2, 4.3	Article 56: 2(c), 2(d) Article 55 1(a), 1(b), 1(d)
3) Publicly report advanced AI systems' capabilities, limitations, and domains of appropriate and inappropriate use, to support ensuring sufficient transparency, thereby contributing to increased accountability.	Govern 1.2, 2.1, 4.2, 4.3 Map 1.1, 1.3, 1.5, 2.1, 2.2, 2.3, 3, 4.1, 4.2, 5.1 Measure 1.1, 1.2, 2.1, 2.3, 2.5, 2.7-2.13, 3.1, 4.2, 4.3 Manage 2.3, 2.4	Article 53 1(b), 1(d)

International Code of Conduct for Organizations Developing Advanced AI Systems Actions	NIST AI RMF Functions, Categories, or Subcategories with the most relevant guidance in this Profile	EU AI Act Article 56 Codes of Practice for Providers of General Purpose AI Models with Systemic Risk
4) Work towards responsible information sharing and reporting of incidents among organizations developing advanced AI systems, including industry, governments, civil society, and academia.	Govern 1.2, 2.1, 4.2, 4.3, 5 Map 1.1, 1.3, 1.5, 1.6, 2.1, 2.2, 2.3, 3.1, 3.2, 4.2 Measure 1.1, 1.2, 1.3, 2.7, 2.8, 2.9, 3.1, 4.1, 4.2, 4.3 Manage 2.4	Article 53 1(a), 1(b), 1(d) Article 55 1(c) Article 56 5
5) Develop, implement, and disclose AI governance and risk management policies that are grounded in a risk-based approach – including privacy policies and mitigation measures.	Govern Map 1 Measure 1 Manage 1, 2, 4	Article 53 1(c) Article 55 1 Article 56 2(c), 2(d)
6) Invest in and implement robust security controls, including physical security, cybersecurity, and insider threat safeguards across the AI lifecycle.	Govern 1.1, 1.2, 1.5, 1.7, 2.1, 4.1, 6.2 Map 1.5, 4.2, 5.1 Measure 1.1, 2.7, 2.10 Manage 1.1, 1.3, 4	Article 55 1(a), 1(d) Article 56 2(d)
7) Develop and deploy reliable content authentication and provenance mechanisms where technically feasible, such as watermarking or other techniques to enable users to identify AI-generated content.	Govern 1.2, 4.2 Map 2.1 Measure 2.7, 2.9	-
8) Prioritize research to mitigate societal, safety, and security risks, and prioritize investment in effective mitigation measures.	Govern 3.1	-
9) Prioritize the development of advanced AI systems to address the world's greatest challenges, notably but not limited to the climate crisis, global health, and education.	Govern 1.2, 5.2 Map 3.1, 5	-
11) Implement appropriate data input measures and protections for personal data and intellectual property.	-	Article 53 1(c), 1(d)

7. Mapping to Frontier AI Safety Commitments

In this section, we provide mapping of the Profile guidance to the Frontier AI Safety Commitments (DSIT 2024).

Table 7: Mapping to Frontier AI Safety Commitments

Frontier AI Safety Commitments	NIST AI RMF Functions, Categories, or Subcategories with the most relevant guidance in this Profile
Outcome 1. Organisations effectively identify, assess, and manage risks when developing and deploying their frontier AI models and systems. They will:	
I. Assess the risks posed by their frontier models or systems across the AI lifecycle, including before deploying that model or system, and, as appropriate, before and during training. Risk assessments should consider model capabilities and the context in which they are developed and deployed, as well as the efficacy of implemented mitigations to reduce the risks associated with their foreseeable use and misuse. They should also consider results from internal and external evaluations as appropriate, such as by independent third-party evaluators, their home governments, and other bodies their governments deem appropriate.	Govern 4.2, 5 Map 1.1, 1.3, 1.4, 1.5, 1.6, 2, 3, 4, 5 Measure
II. Set out thresholds at which severe risks posed by a model or system, unless adequately mitigated, would be deemed intolerable. Assess whether these thresholds have been breached, including monitoring how close a model or system is to such a breach. These thresholds should be defined with input from trusted actors, including organisations' respective home governments as appropriate. They should align with relevant international agreements to which their home governments are party. They should also be accompanied by an explanation of how thresholds were decided upon, and by specific examples of situations where the models or systems would pose intolerable risk.	Govern 1.1, 1.3, 1.4 Map 1.5, 1.6, Measure 1, 2
III. Articulate how risk mitigations will be identified and implemented to keep risks within defined thresholds, including safety- and security-related risk mitigations such as modifying system behaviours and implementing robust security controls for unreleased model weights.	Map 2.1, Measure 1, 2, Manage 1, 2, 3, 4
IV. Set out explicit processes they intend to follow if their model or system poses risks that meet or exceed the pre-defined thresholds. This includes processes to further develop and deploy their systems and models only if they assess that residual risks would stay below the thresholds. In the extreme, organisations commit not to develop or deploy a model or system at all if mitigations cannot be applied to keep risks below the thresholds.	Govern 1.7 Manage 1.1, 1.3, 2.4

Frontier AI Safety Commitments	NIST AI RMF Functions, Categories, or Subcategories with the most relevant guidance in this Profile
V. Continually invest in advancing their ability to implement commitments I-IV, including risk assessment and identification, thresholds definition, and mitigation effectiveness. This should include processes to assess and monitor the adequacy of mitigations, and identify additional mitigations as needed to ensure risks remain below the pre-defined thresholds. They will contribute to and take into account emerging best practice, international standards, and science on AI risk identification, assessment, and mitigation.	Measure 3 Manage 4
Outcome 2. Organisations are accountable for safely developing and deploying their frontier AI models and systems. They will:	
VI. Adhere to the commitments outlined in I-V, including by developing and continuously reviewing internal accountability and governance frameworks and assigning roles, responsibilities, and sufficient resources to do so.	Govern 1.5, 1.6, 2, 3.2
Outcome 3. Organisations' approaches to frontier AI safety are appropriately transparent to external actors, including governments. They will:	
VII. Provide public transparency on the implementation of the above (I-VI), except insofar as doing so would increase risk or divulge sensitive commercial information to a degree disproportionate to the societal benefit. They should still share more detailed information which cannot be shared publicly with trusted actors, including their respective home governments or appointed body, as appropriate.	Govern 1.4, 4.3
VIII. Explain how, if at all, external actors, such as governments, civil society, academics, and the public are involved in the process of assessing the risks of their AI models and systems, the adequacy of their safety framework (as described under I-VI), and their adherence to that framework.	Govern 5 Map 5 Measure 3.3, 4 Manage 4

8. Mapping to ISO/IEC 42001

In this section, we provide mapping of Profile guidance to key clauses in ISO/IEC 42001:2023, "Information technology — Artificial intelligence — Management system." This is based in part on the NIST AI RMF to ISO/IEC FDIS 42001 AI Management System Crosswalk, which was provided to NIST by Microsoft (2023).

Table 8: Mapping to ISO/IEC 42001:2023 Clauses

ISO/IEC 42001 Clause	NIST AI RMF Functions, Categories, or Subcategories with the most relevant guidance in this Profile
4.1 Understanding the organization and its context	Govern 1 Map 1
4.3 Determining the scope of the AI management system	Map 3 Map 4 Map 5 Measure 1
4.4 AI management system	Manage
5.1 Leadership and commitment	Govern 1 Govern 2 Govern 4
5.2 AI policy	Govern 1
5.3 Roles, responsibilities, and authorities	Govern 2
6.1.2 AI risk assessment	Map 2 Map 3 Map 4 Map 5
6.1.3 AI risk treatment	Measure 1 Measure 2 Manage 1, 2
6.1.4 AI system impact assessment	Map 5
7.2 Competence	Govern 2.2
9.1 Monitoring, measurement, analysis, and evaluation	Measure Manage
10.1 Continual improvement	Manage 2 Manage 4
10.2 Nonconformity and corrective action	Govern 1.7 Manage 2.3

References

Anthony M. Barrett, Jessica Newman, Brandie Nonnecke, Nada Madkour, Dan Hendrycks, Evan R. Murphy, Krystal Jackson, and Deepika Raman (2025) AI Risk-Management Standards Profile for General-Purpose AI (GPAI) and Foundation Models, Version 1.1. UC Berkeley Center for Long-Term Cybersecurity. https://cltc.berkeley.edu/wp-content/uploads/2025/01/Berkeley-AI-Risk-Management-Standards-Profile-for-General-Purpose-AI-and-Foundation-Models-v1-1.pdf

Joseph R. Biden (2023) Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. Executive Order 14110 of October 30, 2023. 88 FR 75191, 75191–75226, November 1, 2023. https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence

DSIT (2024) Frontier AI Safety Commitments. In AI Seoul Summit 2024. UK Department for Science, Innovation & Technology, https://www.gov.uk/government/publications/frontier-ai-safety-commitments-ai-seoul-summit-2024/frontier-ai-safety-commitments-ai-seoul-summit-2024

EP (2024) Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance). European Parliament, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689

G7 (2023) Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems. G7 2023 Hiroshima Summit, https://www.mofa.go.jp/files/100573473.pdf

Microsoft (2023) NIST AI Risk Management Framework to ISO-IEC-42001 Crosswalk. Microsoft, https://airc.nist.gov/docs/NIST_AI_RMF_to_ISO_IEC_42001_Crosswalk.pdf

NIST (2023) Crosswalk AI RMF (1.0) and ISO/IEC FDIS 23894 Information technology – Artificial intelligence – Guidance on risk management. National Institute of Standards and Technology, https://www.nist.gov/document/ai-rmf-crosswalk-iso

NIST (2024) Managing Misuse Risk for Dual-Use Foundation Models, Initial Public Draft. NIST AI 800-1 ipd. National Institute of Standards and Technology, https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.800-1.ipd.pdf

White House (2023) Ensuring Safe, Secure, and Trustworthy AI. White House, https://www.whitehouse.gov/wp-content/uploads/2023/07/Ensuring-Safe-Secure-and-Trustworthy-AI.pdf


CLTC
Center for Long-Term
Cybersecurity
UC Berkeley