Global Index for

AI Safety

AGILE Index on Global AI Safety Readiness

Feb 2025

- Center for Long-term Artificial Intelligence (CLAI)
- Beijing Institute of AI Safety and Governance (Beijing-AISI)
- Beijing Key Laboratory of Safe AI and Superalignment
- Institute of Automation, Chinese Academy of Sciences

Towards A Global AI Safety Readiness Assessment

As artificial intelligence (AI) technologies experience explosive growth and proliferate across global industries, their transformative potential is increasingly accompanied by complex safety and security risks. From malicious exploitation and deceptive applications to privacy breaches, unintended consequences, and existential risks, the dual-use nature of AI and its global impact demand comprehensive safeguards and strengthened international cooperation. Against this backdrop, understanding how countries navigate AI safety challenges—through policy innovation, technical safeguards, and multilateral cooperation—has become critical to shaping a safe and sustainable future.

Developed under the theoretical framework of *AI Governance InternationaL Evaluation (AGILE) Index*, this ***Global Index for AI Safety (GIAIS)*** provides a systematic assessment of national capabilities, current status and preparedness in addressing AI safety challenges. The evaluation of the index covers six pillars: Governance Environment for AI Safety, National Institutions Targeting AI Safety, Governance Instruments for AI Safety, Research Status on AI Safety, International Participation on AI Safety, and Existential Safety Preemption. It currently includes 12 dimensions to depict the governance status of AI safety readiness across 40 countries.

Through this assessment, we can found:

- Developed countries are generally better-prepared in addressing AI safety challenges.
- The global AI safety environment is becoming increasingly severe in recent years.
- National AI safety institutions are rapidly emerging in various forms.
- Related laws, policies, and tools are being implemented, but only in some countries.
- AI safety research has surged, focusing on topics such as alignment and privacy security.
- International AI safety cooperation is forming but needs wider participation.
- AI existential safety preemption and planning are lacking in all countries.

The assessment does not seek to categorize countries as either paragons or laggards. AI's safety challenges affect us all, and no country can solve them alone, no matter how well the country itself has done. The level of development on AI technology and its use may vary across countries due to their development status, while all countries should get more serious preparations on AI Safety due to its potential unpredictability, profound and deep negative impacts, as well as its proliferation characteristics. Therefore, the GIAIS acts as an assessment tool, helping countries recognize their circumstances and deficiencies. By sharing experiences and learning from each other, we can enhance global cooperation, coordinate resources, and guide AI towards a safer, more sustainable future. This united effort is not just for the present but a legacy for future generations, ensuring AI becomes both a safe and powerful force propelling humanity forward.

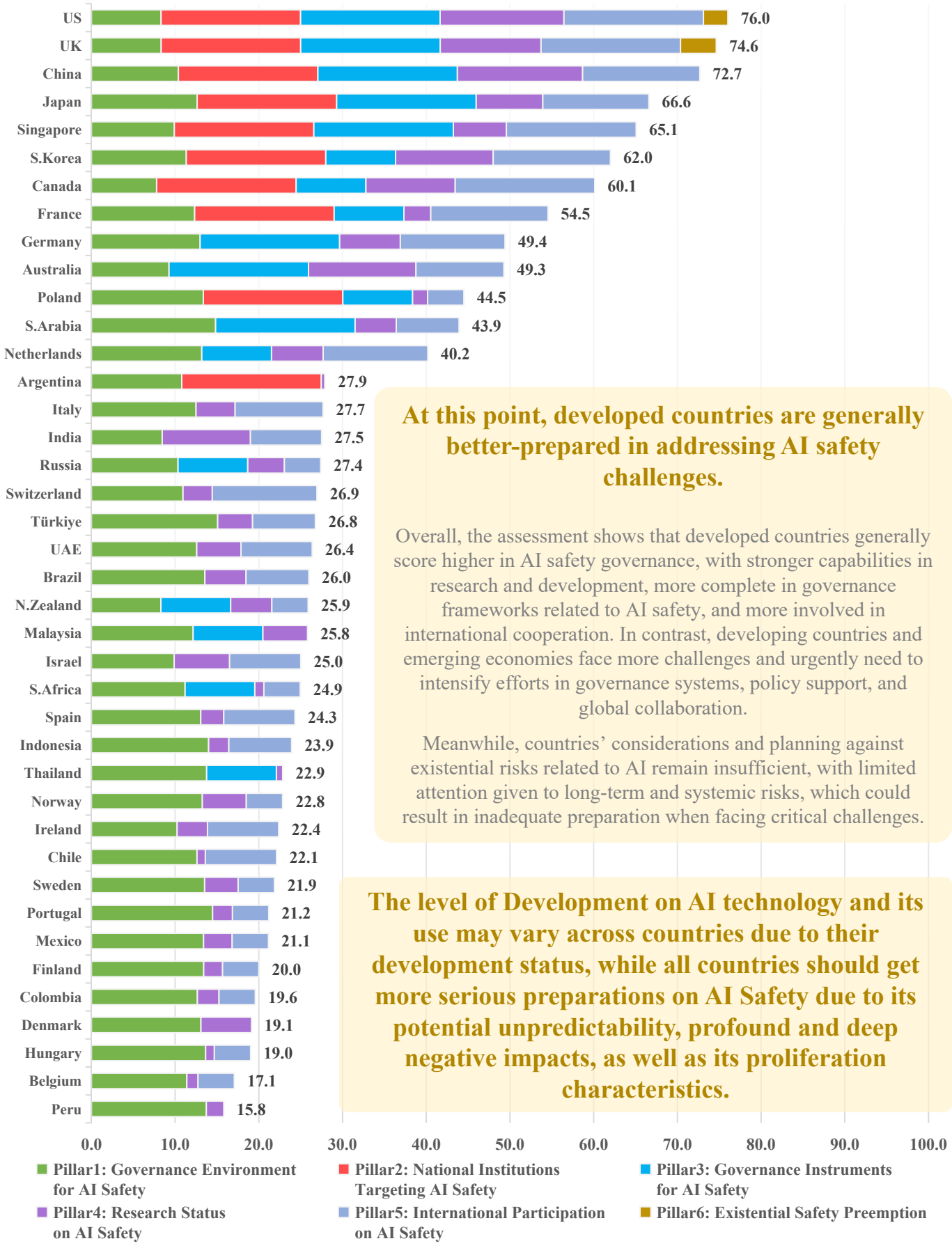
Yi Zeng
Dean and Professor
Beijing Institute of AI Safety and Governance (Beijing-AISI)
Beijing Key Laboratory of Safe AI and Superalignment
Center for Long-term AI
International Research Center for AI Ethics and Governance,
Institute of Automation, Chinese Academy of Sciences

Global Index for AI Safety

GLOBAL INDEX FOR AI SAFETY

PILLARS	DIMENSIONS	EVALUATION CRITERIA
Governance Environment	Cybersecurity Status	**This pillar primarily assesses AI safety challenges faced by the country.** A lower score indicates more exposed safety issues and more governance pressure.
	AI Safety Incidents	
National Institutions	National AI Safety Institutes/Networks/Labs/Consortiums	**This pillar evaluates the national institutional readiness on AI safety.** A relatively high score reflects that the country is ahead in establishing the institutional basis for assessing and addressing AI safety risks.
Governance Instruments	National Laws & Regulations related to AI Safety	**This pillar evaluates the completeness and effectiveness of a country's laws, policies, and tools related to AI safety.** A higher score indicates that the instruments are more comprehensive and complete in assessing and addressing challenges.
	Technical & Policy Frameworks for AI Safety	
Research Status	AI Safety Publications	**This pillar assesses the research capabilities of each country in discussing, researching, and addressing the risks of AI safety.** The higher the score, the greater the research attention and capabilities regarding AI safety risks.
	AI Safety Patents	
International Participation	Government Engagement	**This pillar reflects a country's activity in global AI safety governance mechanisms.** A higher score indicates a larger role in promoting international collaboration, setting safety standards, and strengthening global awareness.
	Industry Engagement	
	Academia & Civil Society Engagement	
Existential Safety Preemption	Government Engagement	**This pillar assesses a country's strategic planning in preventing existential risks posed by AI.** A higher score reflects a more sufficient considerations, actions and strategic planning of governments and industries in proactively addressing existential risks.
	Industry Engagement	

Overall Index Scores



At this point, developed countries are generally better-prepared in addressing AI safety challenges.

Overall, the assessment shows that developed countries generally score higher in AI safety governance, with stronger capabilities in research and development, more complete in governance frameworks related to AI safety, and more involved in international cooperation. In contrast, developing countries and emerging economies face more challenges and urgently need to intensify efforts in governance systems, policy support, and global collaboration.

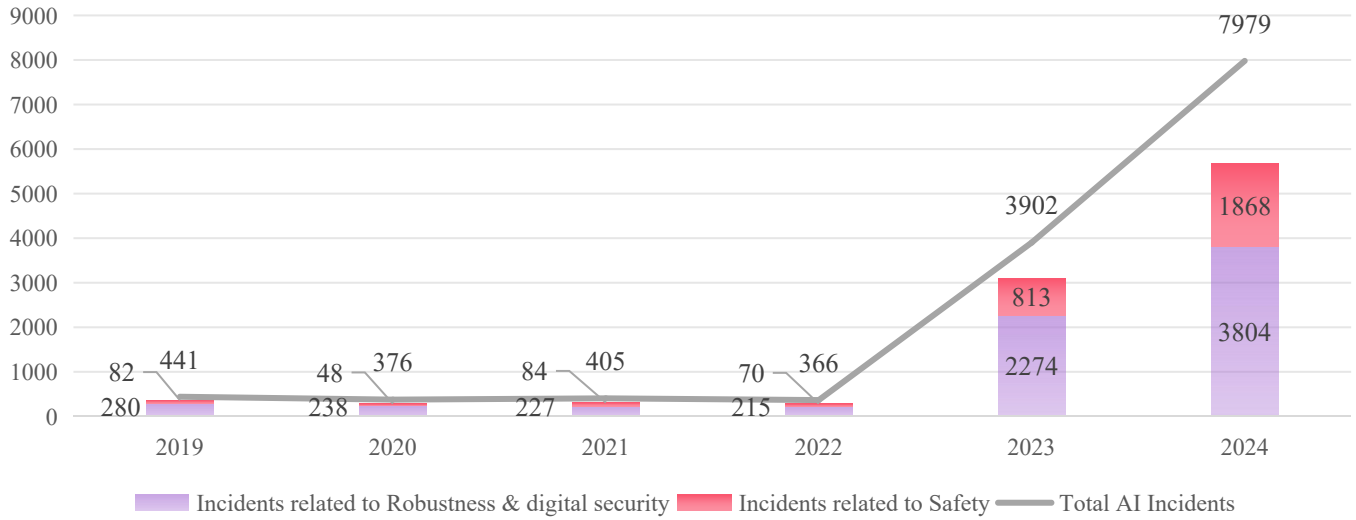
Meanwhile, countries' considerations and planning against existential risks related to AI remain insufficient, with limited attention given to long-term and systemic risks, which could result in inadequate preparation when facing critical challenges.

The level of Development on AI technology and its use may vary across countries due to their development status, while all countries should get more serious preparations on AI Safety due to its potential unpredictability, profound and deep negative impacts, as well as its proliferation characteristics.

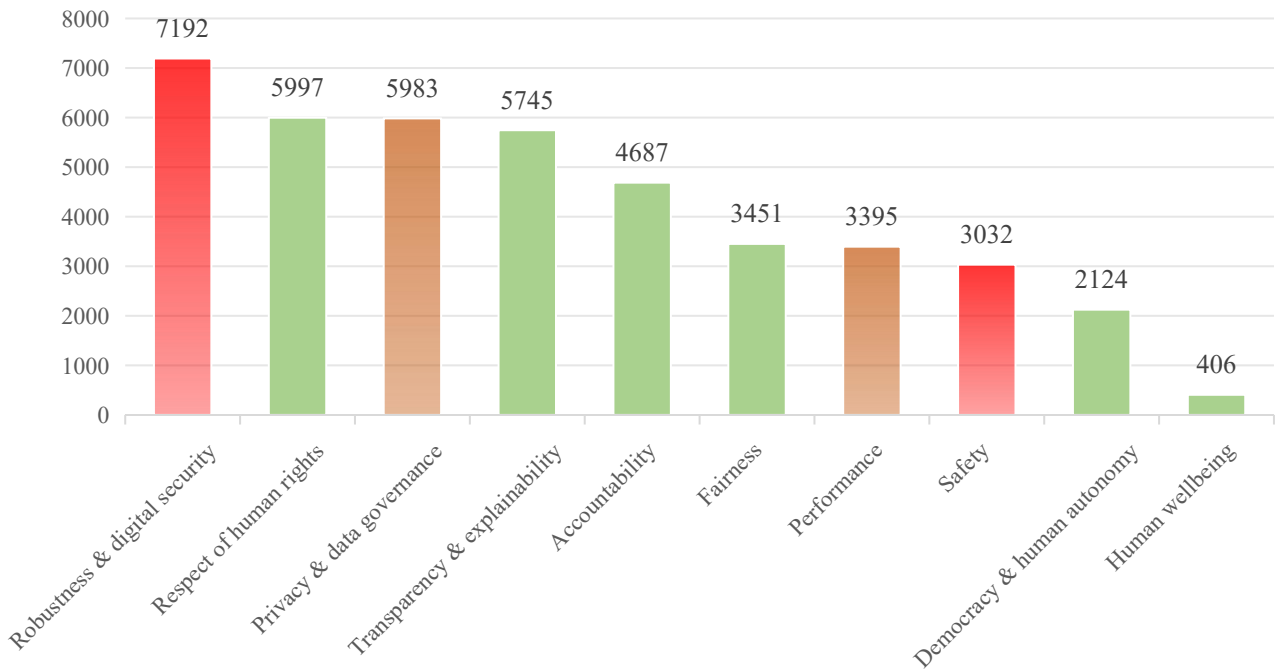
The global AI safety environment is becoming increasingly severe in recent years.

Since 2022, with the breakthroughs in generative AI technology and its deepening application across various fields, the total number of AI risk incidents has surged. According to the OECD AI Incidents Monitor (AIM), the total number of risk incidents in 2024 has increased by approximately 21.8 times compared to 2022, showing a rapid growth trend. Of the AI risk incidents that occurred between 2019 and 2024, about 74% were directly related to AI safety issues. The number of AI incidents that directly related to safety & security in 2024 grew by approximately 83.7% compared to 2023.

Trends in the Increase of AI Safety Incidents



AI Incidents Distribution



Data Source: the OECD AI Incidents Monitor (AIM) From 2019 to Oct. 2024

Pillar2: National Institutions Targeting AI Safety

National AI Safety Institutions are rapidly emerging in various forms.

Since the release of the *Bletchley Declaration*, the establishment of national AI safety research institutions has become a significant trend globally. Among the 40 countries surveyed, ten countries and the European Union have established or designated specialized AI safety research institutions. Most of these institutions operate within the framework of the AI Safety Summit, fostering a networked collaboration model. Meanwhile, such national institutions also presents in varies forms in different countries, such as the China AI Development and Safety Network (China AISI Network), and the Artificial Intelligence Unit for Security (UIAAS) established by Argentina, showcasing the diverse approaches countries are taking in in establishing the institutional basis for AI safety research.

2023.11 — AI SAFETY SUMMIT HOSTED BY THE UK 1-2 NOVEMBER 2023
AI SAFETY SUMMIT

2023.11 — 
US AISI is announced

2023.11 — AISI AI SAFETY INSTITUTE
UK AISI is established (was renamed AI Security Institute in February 2025)

2024.2 — 
US AISIC is established

2024.2 — AISI Japan AI Safety Institute
AISI Japan is established

2024.5 — AI SEOUL SUMMIT 21 - 22 MAY 2024 Hosted by the Republic of Korea and the United Kingdom
AI SEOUL SUMMIT

2024.5 — 
Digital Trust Center is designated as Singapore's AISI

2024.5 — 
European Commission establishes AI Office

2024.7 — 
Argentina's AI Unit for Security is established

2024.9 — 
China AI Development and Safety Network is established

2024.11 — 
Canadian AISI is established

2024.11 — 

First meeting of the International Network of AI Safety Institutes

2024.11 — AISI 인공지능안전연구소
S.Korea's AISI is established

2024.11 — 
Polish AI Safety Research Center @ NASK is established.

2025.1 — 
French National Institute for the Evaluation and Security of AI is established

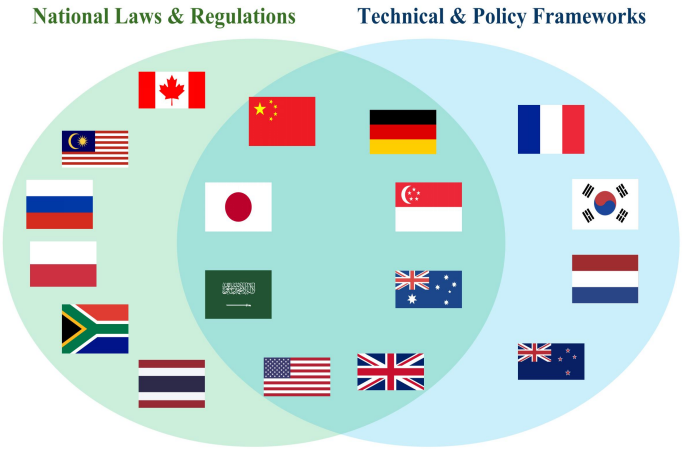
2025.2 — AI ACTION SUMMIT
AI ACTION SUMMIT

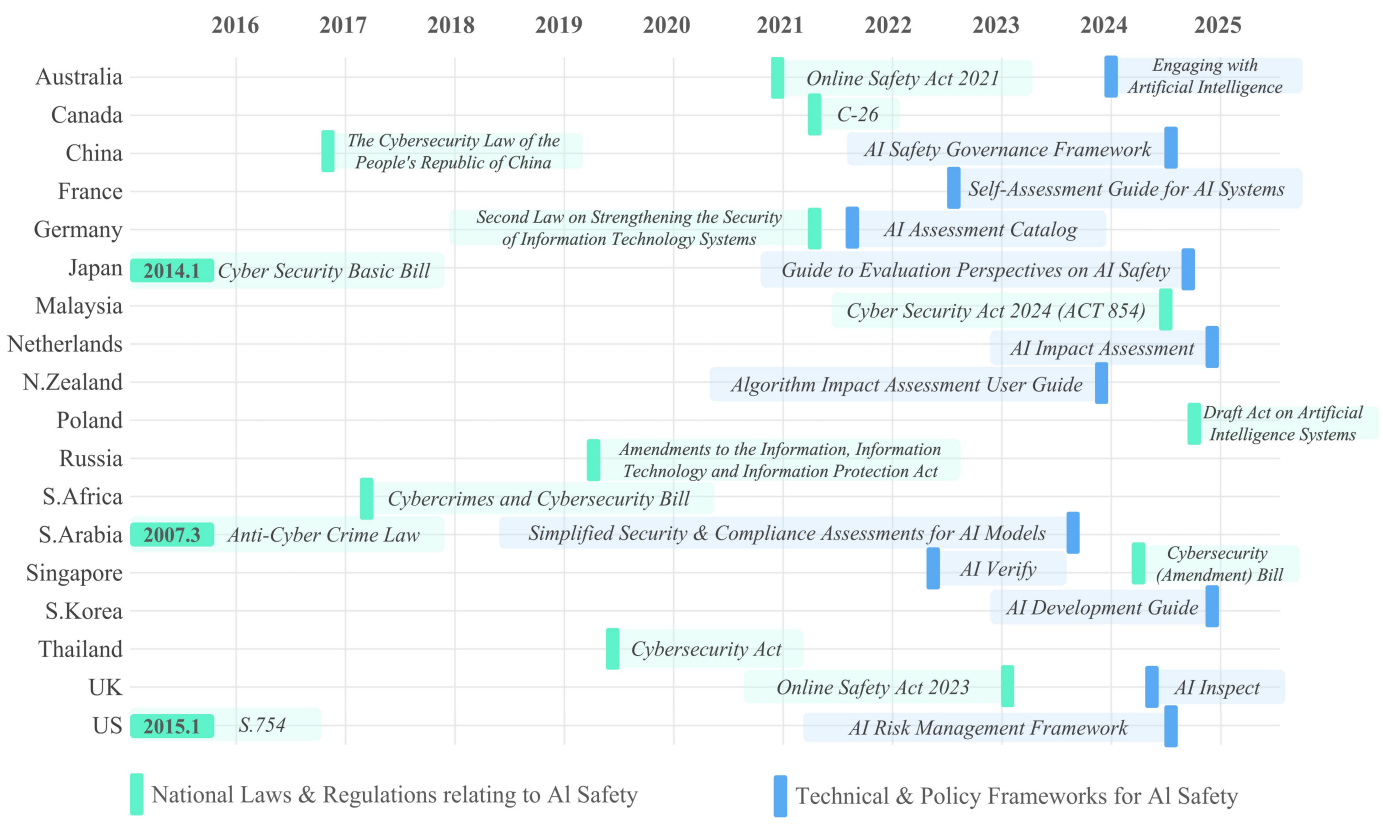
Pillar3: Governance Instruments for AI Safety

Laws, policies, and tools related to AI safety are being implemented, but only in some countries.

In a survey of 40 countries, 18 have governance instruments related to AI safety. Among these, 8 countries—Australia, China, Germany, Japan, Singapore, Saudi Arabia, the United Kingdom, and the United States—have both national AI-related safety laws and technical/policy frameworks in place. This highlights a global trend towards regulating AI safety through such frameworks. However, most AI-related safety laws are still primarily focused on cybersecurity and information security, with laws specifically targeting AI safety remaining relatively scarce. The majority of technical and policy frameworks were released in 2024, reflecting the concerted efforts to tackle AI safety issues in the past year.

Governance Instruments for AI Safety by Countries



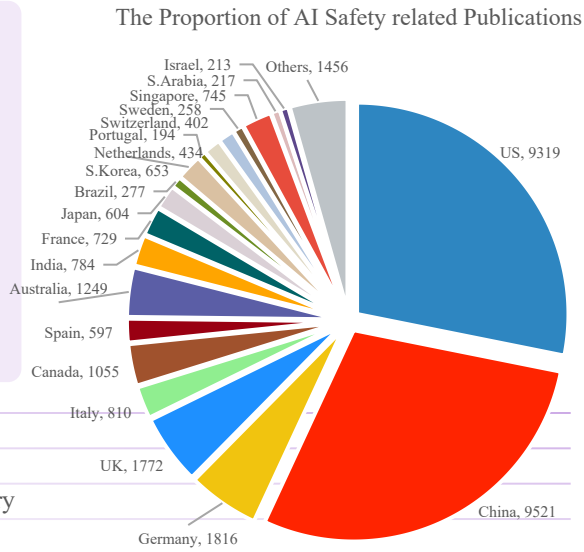


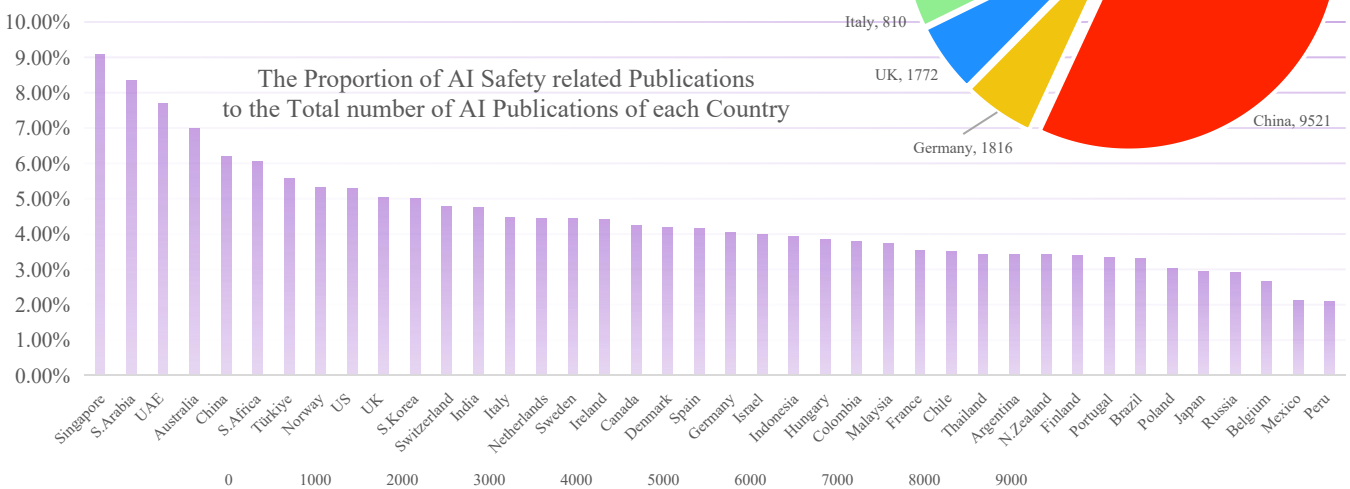
More detailed information can be found in Appendix B.

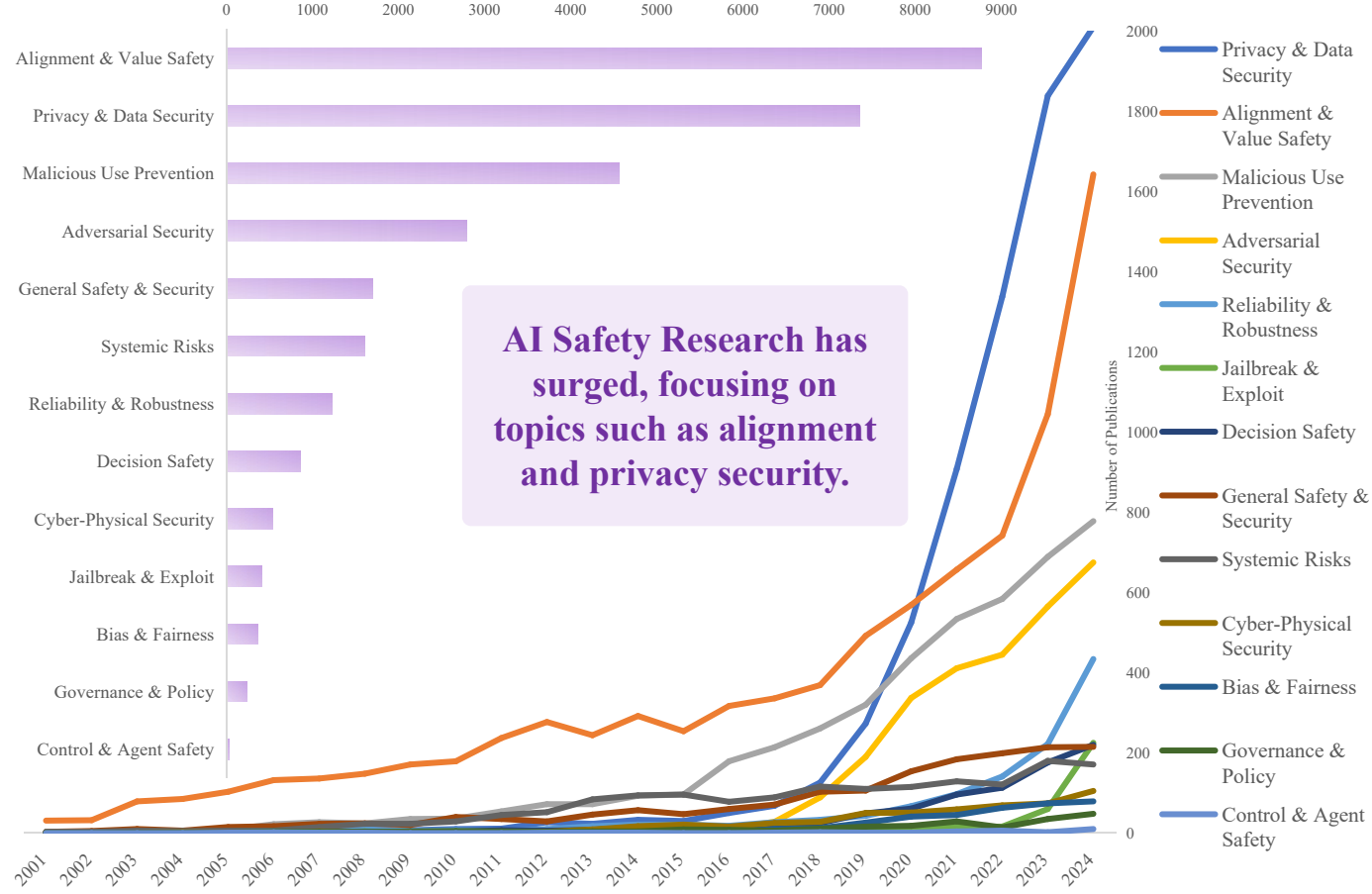
Pillar4: Research Status on AI Safety

China and the United States together contributed more than half of the research papers on AI safety till 2024, accounting for 28.8% and 28.1% of the total publication volume from all 40 countries, respectively, highlighting their focus on the field of AI safety.

Singapore, Saudi Arabia, and the United Arab Emirates ranked among the top three in terms of the percentage of safety-related publications in their total AI publications, with 9.1%, 8.3%, and 7.7%, respectively. This also indicates a relatively high level of attention to safety & security in these countries.







AI Safety Research has surged, focusing on topics such as alignment and privacy security.

Data Source: the DBLP Computer Science Bibliography literature database (Data as of Feb, 2025)

Pillar5: International Participation on AI Safety

International AI safety cooperation is forming, but needs wider participation.

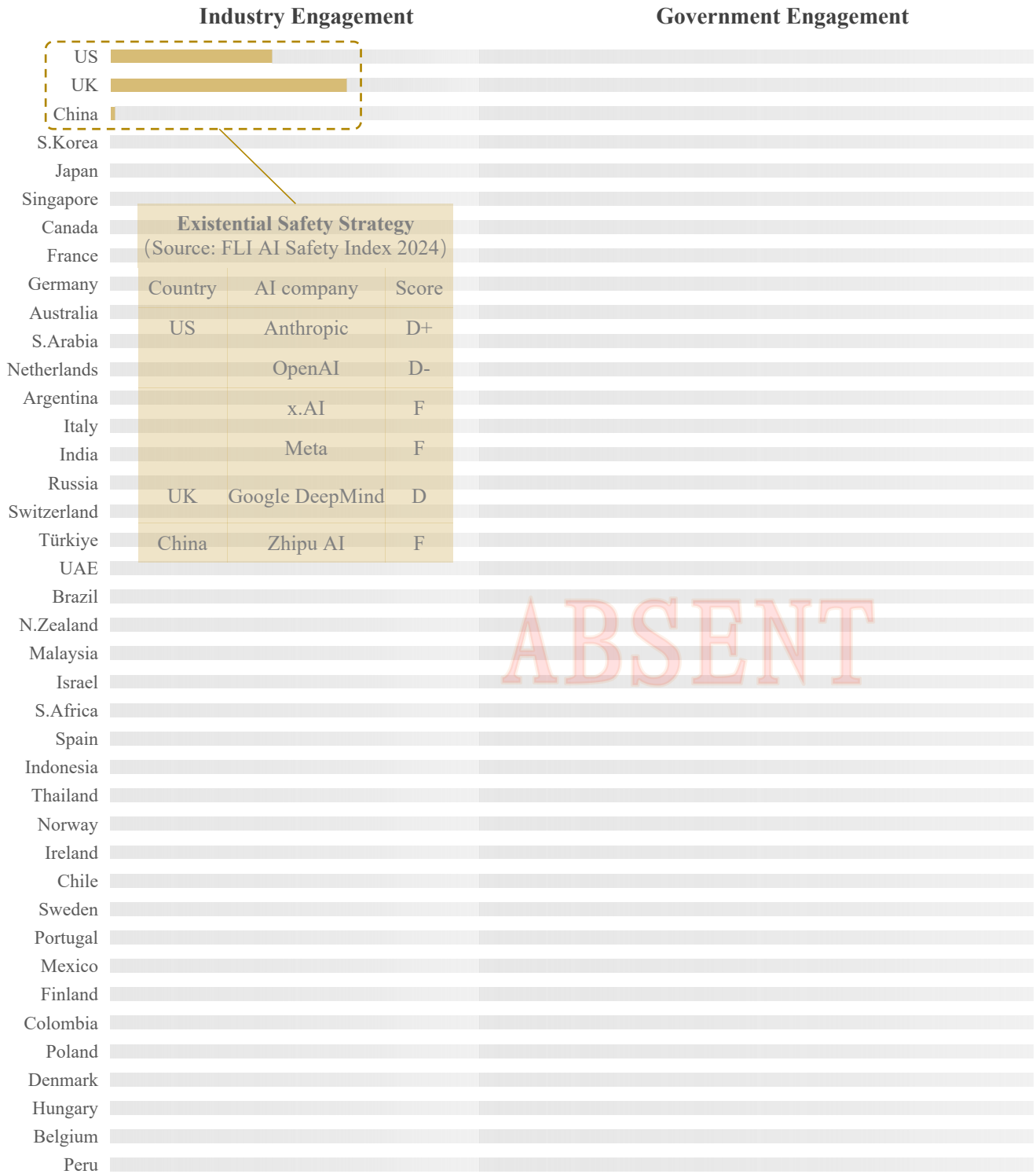
Since 2023, AI safety has emerged as a core issue in global governance frameworks. The 10 international cooperation projects surveyed reflect a multi-stakeholder approach, spanning government, industry, academia, and civil society. These efforts reflect global consensus on AI safety but also reveal the need for broader inclusivity in governance mechanisms.

	Government Engagement					Industry, Academia & Civil Society Engagement				
Country	Bletchley Declaration (AI Safety Summit 2023)	Seoul Ministerial Statement (AI Seoul Summit 2024)	Call to Action (REAIM 2023)	Blueprint for Action (REAIM 2024)	First meeting of the International Network of AI Safety Institutes 2024	Statement on AI Risk 2023 (Center for AI Safety)	Frontier AI Safety Commitments (AI Seoul Summit 2024)	International Dialogues on AI Safety (IDAIS) Oxford 2023	IDAIS Beijing 2024	IDAIS Venice 2024
US	√	√	√	√	√	√	√	√	√	√
Canada	√	√	√	×	√	√	√	√	√	√
UK	√	√	√	√	√	√	×	√	√	√
Singapore	√	√	√	√	√	√	×	×	√	√
China	√	×	√	×	×	√	√	√	√	√
France	√	√	√	√	√	√	√	×	×	×
S.Korea	√	√	√	√	√	√	√	×	×	×
Japan	√	√	√	√	√	√	×	×	×	×
Germany	√	√	√	√	×	√	×	×	×	×
Netherlands	√	√	√	√	×	√	×	×	×	×
Switzerland	√	√	×	×	×	√	×	√	×	√
Australia	√	√	×	×	√	√	×	×	×	×
Italy	√	√	√	×	×	√	×	×	×	×
India	√	√	×	×	×	√	×	×	×	×
Chile	√	√	×	×	×	√	×	×	×	×
Ireland	√	×	×	×	×	√	×	×	×	√
Israel	√	√	×	×	×	√	×	×	×	×
Spain	√	√	×	×	×	√	×	×	×	×
UAE	√	√	×	×	×	×	√	×	×	×
Brazil	√	×	×	×	×	√	×	×	×	×
Indonesia	√	√	×	×	×	×	×	×	×	×
S.Arabia	√	√	×	×	×	×	×	×	×	×
Türkiye	√	√	×	×	×	×	×	×	×	×
Belgium	×	×	×	×	×	√	×	×	×	×
Colombia	×	×	×	×	×	√	×	×	×	×
Finland	×	×	×	×	×	√	×	×	×	×
Hungary	×	×	×	×	×	√	×	×	×	×
Mexico	×	√	×	×	×	×	×	×	×	×
Norway	×	×	×	×	×	√	×	×	×	×
N.Zealand	×	√	×	×	×	×	×	×	×	×
Poland	×	×	×	×	×	√	×	×	×	×
Portugal	×	×	×	×	×	√	×	×	×	×
Russia	×	×	×	×	×	√	×	×	×	×
S.Africa	×	×	×	×	×	√	×	×	×	×
Sweden	×	×	×	×	×	√	×	×	×	×
Argentina	×	×	×	×	×	×	×	×	×	×
Denmark	×	×	×	×	×	×	×	×	×	×
Malaysia	×	×	×	×	×	×	×	×	×	×
Peru	×	×	×	×	×	×	×	×	×	×
Thailand	×	×	×	×	×	×	×	×	×	×

Pillar6: Existential Safety Preemption

AI existential safety preemption and planning are lacking in all countries.

For years, the academic community has been actively engaged in discussions and has shown great concern for the potential existential risks that could be brought about by advanced artificial intelligence. Nevertheless, the industry and governments, being the key players in the development, promotion, and governance of cutting-edge AI systems, still widely lack preemptive considerations and actions against AI's existential safety risks.



Existential Safety Strategy
(Source: FLI AI Safety Index 2024)

Country	AI company	Score
US	Anthropic	D+
	OpenAI	D-
	x.AI	F
	Meta	F
UK	Google DeepMind	D
China	Zhipu AI	F

Appendix A: Index Indicators, Methodology & Results

PILLAR	DIMENSION	METHODOLOGY & SOURCE
P1. Governance Environment for AI Safety	D1. Cybersecurity Status	ITU Global Cybersecurity Index[1]
	D2. AI Safety Incidents	Number of AI safety related risk cases/incidents & the GDP ratio. The data is sourced from OECD AI Incidents Monitor[2], and the raw data were scored using percentile-fit normalization. For specific methodologies, please refer to the appendix of the AGILE Index[3]
P2. National Institutions Targeting to AI Safety	D3. National AI Safety Institutes/Networks/Labs/Consortiums	Public Information Survey
P3. Governance Instruments for AI Safety	D4. National Laws & Regulations related to AI Safety	Public Information Survey
	D5. Technical & Policy Frameworks for AI Safety	Public Information Survey
P4. Research Status on AI Safety	D6. AI Safety Publications	Total number & the proportion of publications on AI safety topics. The data is sourced from the DBLP Computer Science Bibliography literature database. To determine whether the literature is related to AI safety, we combined the keywords for AI safety publication analysis based on the International AI Safety Report[4], along with DeepSeek-R1[5].
	D7. AI Safety Patents	Number of granted AI patents & the per capita ratio. The data is sourced from World Intellectual Property Organization open data, combined with safety-related keywords (AI Safety, AI Security)
P5. International Participation on AI Safety	D8. Government Engagement	The data is sourced from the cumulative number of participations in relevant international activities as follows, and then percentile-fit normalization is conducted. ·*The Bletchley Declaration*, AI Safety Summit 2023 ·Seoul Ministerial Statement for advancing AI safety, innovation and inclusivity, AI Seoul Summit 2024 ·*Call to Action*, Summit on Responsible Artificial Intelligence in the Military Domain (REAIM) 2023 ·*Blueprint for Action*, REAIM 2024 ·First meeting of the International Network of AI Safety Institutes 2024 ·*Statement on AI Risk*, Center for AI Safety 2023 ·*Frontier AI Safety Commitments*, AI Seoul Summit 2024 ·International Dialogues on AI Safety (IDAIS) Oxford 2023 ·IDAIS-Beijing 2024 ·IDAIS-Venice 2024
	D9. Industry Engagement	
	D10. Academia & Civil Society Engagement	
P6. Existential Safety Preemption	D11. Government Engagement	Industry engagement is assessed based on the FLI AI Safety Index 2024's Existential Safety Strategy score[6], contributing 40% to the dimension's overall weight. Government engagement accounts for 60%, with scores determined by the government's strategic approach to AI's existential risks (Currently None).
	D12. Industry Engagement	

[1] https://www.itu.int/epublications/publication/global-cybersecurity-index-2024
[2] https://oecd.ai/en/dashboards/overview
[3] https://agile-index.ai/
[4] https://arxiv.org/abs/2501.17805
[5] https://chat.deepseek.com/
[6] https://futureoflife.org/document/fli-ai-safety-index-2024/

Appendix A: Index Indicators, Methodology & Results

Country	Pillar1: Governance Environment for AI Safety		Pillar2: National Institutions Targeting AI Safety	Pillar3: Governance Instruments for AI Safety		Pillar4: Research Status on AI Safety		Pillar5: International Participation on AI Safety		Pillar6: Existential Safety Preemption	Overall Readiness Ranking	Overall Index Score
	D1	D2	D3	D4	D5	D6	D7	D8-D10	D11	D12		
US	99.9	0.0	100.0	100.0	100.0	77.5	100.0	100.0	0.0	43.8	1	76.0
UK	100.0	0.0	100.0	100.0	100.0	57.0	87.5	100.0	0.0	64.0	2	74.6
China	91.6	33.0	100.0	100.0	100.0	87.0	92.5	84.0	0.0	0.0	3	72.7
Japan	97.6	54.0	100.0	100.0	100.0	33.0	62.5	76.0	0.0	0.0	4	66.6
Singapore	99.9	19.0	100.0	100.0	100.0	76.0	0.0	93.0	0.0	0.0	5	65.1
S.Korea	100.0	36.0	100.0	0.0	100.0	64.0	76.0	84.0	0.0	0.0	6	62.0
Canada	76.5	17.0	100.0	100.0	0.0	59.5	68.5	100.0	0.0	0.0	7	60.1
France	99.0	49.0	100.0	0.0	100.0	38.5	0.0	84.0	0.0	0.0	8	54.5
Germany	93.8	62.0	0.0	100.0	100.0	57.5	29.5	75.0	0.0	0.0	9	49.4
Australia	96.2	15.0	0.0	100.0	100.0	81.5	72.5	63.0	0.0	0.0	10	49.3
Poland	88.4	72.0	100.0	100.0	0.0	21.5	0.0	26.0	0.0	0.0	11	44.5
S.Arabia	100.0	78.0	0.0	100.0	100.0	59.0	0.0	45.0	0.0	0.0	12	43.9
Netherlands	99.2	59.0	0.0	0.0	100.0	37.0	37.0	75.0	0.0	0.0	13	40.2
Argentina	51.5	78.0	100.0	0.0	0.0	5.0	0.0	0.0	0.0	0.0	14	27.9
Italy	100.0	50.0	0.0	0.0	0.0	56.0	0.0	63.0	0.0	0.0	15	27.7
India	98.5	3.0	0.0	0.0	0.0	61.5	65.0	51.0	0.0	0.0	16	27.5
Russia	92.1	32.0	0.0	100.0	0.0	27.0	25.5	26.0	0.0	0.0	17	27.4
Switzerland	91.3	40.0	0.0	0.0	0.0	42.0	0.0	75.0	0.0	0.0	18	26.9
Türkiye	100.0	81.0	0.0	0.0	0.0	50.0	0.0	45.0	0.0	0.0	19	26.8
UAE	100.0	51.0	0.0	0.0	0.0	63.5	0.0	51.0	0.0	0.0	20	26.4
Brazil	93.7	69.0	0.0	0.0	0.0	22.0	37.0	45.0	0.0	0.0	21	26.0
N.Zealand	82.6	17.0	0.0	0.0	100.0	15.0	44.0	26.0	0.0	0.0	22	25.9
Malaysia	98.8	47.0	0.0	100.0	0.0	26.0	38.0	0.0	0.0	0.0	23	25.8
Israel	93.6	25.0	0.0	0.0	0.0	27.0	52.5	51.0	0.0	0.0	24	25.0
S.Africa	86.3	48.0	0.0	100.0	0.0	13.0	0.0	26.0	0.0	0.0	25	24.9
Spain	99.7	57.0	0.0	0.0	0.0	33.0	0.0	51.0	0.0	0.0	26	24.3
Indonesia	100.0	68.0	0.0	0.0	0.0	3.0	26.0	45.0	0.0	0.0	27	23.9
Thailand	99.2	66.0	0.0	100.0	0.0	9.0	0.0	0.0	0.0	0.0	28	22.9
Norway	97.0	62.0	0.0	0.0	0.0	63.0	0.0	26.0	0.0	0.0	29	22.8
Ireland	90.9	32.0	0.0	0.0	0.0	43.5	0.0	51.0	0.0	0.0	30	22.4
Chile	69.4	82.0	0.0	0.0	0.0	12.0	0.0	51.0	0.0	0.0	31	22.1
Sweden	99.3	63.0	0.0	0.0	0.0	48.0	0.0	26.0	0.0	0.0	32	21.9
Portugal	99.7	74.0	0.0	0.0	0.0	28.5	0.0	26.0	0.0	0.0	33	21.2
Mexico	85.8	75.0	0.0	0.0	0.0	11.5	29.5	26.0	0.0	0.0	34	21.1
Finland	100.0	61.0	0.0	0.0	0.0	27.0	0.0	26.0	0.0	0.0	35	20.0
Colombia	68.9	83.0	0.0	0.0	0.0	31.0	0.0	26.0	0.0	0.0	36	19.6
Denmark	100.0	57.0	0.0	0.0	0.0	31.0	41.5	0.0	0.0	0.0	37	19.1
Hungary	88.7	75.0	0.0	0.0	0.0	12.5	0.0	26.0	0.0	0.0	38	19.0
Belgium	96.8	40.0	0.0	0.0	0.0	16.0	0.0	26.0	0.0	0.0	39	17.1
Peru	83.7	81.0	0.0	0.0	0.0	25.0	0.0	0.0	0.0	0.0	40	15.8

Appendix B: List of National AI Safety Institutions

Country	National AI Safety Institutions (including Institutes/Networks/Labs/Consortiums...)	Date (of establishment or public disclosure)
Argentina	The Artificial Intelligence Unit for Security (UIAAS)[1]	2024.7
Canada	Canadian AI Safety Institute (CAISI)[2]	2024.11.12
China	China AI Development and Safety Network[3]	2024.9
France	National Institute for the Evaluation and Security of AI (INESIA)[4]	2025.1.31
Japan	Japan AI Safety Institute (AISI Japan)[5]	2024.2.14
Poland	Artificial Intelligence Safety Research Centre[6]	2024.11
Singapore	Singapore's AI Safety Institute (Digital Trust Center)[7]	2024.5
S. Korea	Korea's AI Safety Institute (AISI)[8]	2024.11.27
UK	AI Safety Institute (UK AISI, name has been changed to AI Security Institute in February 2025) [9]	2023.11.2
UK	Laboratory for AI Security Research (LASR)[10]	2024.11.25
US	U.S. AI Safety Institute (US AISI)[11]	2023.11.1
US	Artificial Intelligence Safety Institute Consortium[12]	2024.2.8

[1] https://www.argentina.gob.ar/noticias/nuevas-herramientas-para-combatir-el-ciberdelito
[2] https://ised-isde.canada.ca/site/ised/en/canadian-artificial-intelligence-safety-institute
[3] https://ai-development-and-safety-network.cn/
[4] https://presse.economie.gouv.fr/le-gouvernement-annonce-la-creation-de-linstitut-national-pour-levaluation-et-la-securite-de-lintelligence-artificielle-inesia/
[5] https://aisi.go.jp/
[6] https://nask.pl/aktualnosci/powstanie-osrodka-badan-nad-bezpieczenstwem-sztucznej-inteligencji-w-nask/
[7] https://www.ntu.edu.sg/dtc
[8] https://www.aisi.re.kr/kor
[9] https://www.aisi.gov.uk/
[10] https://babl.ai/uk-unveils-ai-security-laboratory-at-nato-cyber-defense-conference/
[11] https://www.nist.gov/aisi
[12] https://www.nist.gov/aisi/artificial-intelligence-safety-institute-consortium-aisic

Appendix C: List of National AI Safety Governance Instruments

Country	National Laws & Regulations related to AI Safety	Technical & Policy Frameworks for AI Safety
Australia	Online Safety Act 2021[1]	Engaging with Artificial Intelligence[2]
Canada	C-26[3]	/
China	《中华人民共和国网络安全法》[4](The Cybersecurity Law of the People's Republic of China)	《人工智能安全治理框架》[5](AI Safety Governance Framework)
France	/	Self-assessment guide for AI Systems[6]
Germany	Zweites Gesetz zur Erhöhung der Sicherheit informationstechnischer Systeme[7] (Second law on strengthening the security of information technology systems)	AI Assessment Catalog[8]
Japan	サイバーセキュリティ基本法案[9](Cyber Security Basic Bill)	AIセーフティに関する評価観点ガイド[10] (Guide to Evaluation Perspectives on AI Safety)
Malaysia	Cyber Security Act 2024[11] (ACT 854)	/
Netherlands	/	AI Impact Assessment[12]
N. Zealand	/	Algorithm Impact Assessment User Guide[13]
Poland	Polish Draft Act on Artificial Intelligence Systems[14]	/
Russia	Внесены изменения в закон об информации, информационных технологиях и о защите информации[15](Amendments to the Information, Information Technology and Information Protection Act)	/
S. Africa	Cybercrimes-and-cybersecurity-bill[16]	/
S. Arabia	Anti-Cyber Crime Law[17]	Simplified security & compliance assessments for AI models[18]
Singapore	Cybersecurity (Amendment) Bill[19]	AI Verify[20]
S. Korea	/	AI 개발 안내서[21](AI Development Guide)
Thailand	Cybersecurity Act[22]	/
UK	Online Safety Act 2023[23]	AI INSPECT[24]
US	S.754[25]	AI Risk Management Framework[26]

[1] https://www.legislation.gov.au/C2021A00076/latest/text
[2] https://www.cyber.gov.au/resources-business-and-government/governance-and-user-education/artificial-intelligence/engaging-with-artificial-intelligence
[3] https://www.parl.ca/legisinfo/en/bill/44-1/c-26
[4] http://www.npc.gov.cn/zgrdw/npc/xinwen/2016-11/07/content_2001605.htm
[5] https://www.cac.gov.cn/2024-09/09/c_1727567886199789.htm
[6] https://www.cnil.fr/en/self-assessment-guide-artificial-intelligence-ai-systems
[7] https://www.bmi.bund.de/SharedDocs/downloads/DE/gesetzestexte/it-sicherheitsgesetz-2.pdf?__blob=publicationFile&v=1
[8] https://www.iais.fraunhofer.de/en/research/artificial-intelligence/ai-assessment-catalog.html#Receive-our-AI-assessment-catalog-free-of-charge
[9] https://www.shugiin.go.jp/internet/itdb_gian.nsf/html/gian/honbun/houan/g18601035.htm
[10] https://aisi.go.jp/effort/effort_framework/guide_to_evaluation_perspective_on_ai_safety/
[11] https://www.nacsa.gov.my/act854.php
[12] https://www.government.nl/documents/publications/2023/03/02/ai-impact-assessment
[13] https://data.govt.nz/docs/algorithm-impact-assessment-user-guide
[14] https://www.cliffordchance.com/content/dam/cliffordchance/briefings/2024/11/cb%20-%20projekt%20AI%20eng.pdf
[15] https://web.archive.org/web/20190319122732/https://kremlin.ru/acts/news/60083
[16] https://www.gov.za/documents/cybercrimes-and-cybersecurity-bill-b6-2017-21-feb-2017-0000
[17] https://www.mcit.gov.sa/sites/default/files/2021-06/la_004_e_anti-cyber_crime_law%20%281%29.pdf
[18] https://www.saifcheck.ai/
[19] https://sso.agc.gov.sg/Bills-Supp/15-2024/Published/20240403?spm=5176.28103460.0.0.40f75d27ogoSIr
[20] https://aiverifyfoundation.sg/
[21] https://www.tta.or.kr/tta/selectBbsNttView.do?key=76&bbsNo=107&nttNo=13872&searchCtgry=&searchCnd=all&searchKrwd=&integrDeptCode=&pageIndex=1
[22] https://thainetizen.org/wp-content/uploads/2019/11/thailand-cybersecrutiy-act-2019-en.pdf
[23] https://www.legislation.gov.uk/ukpga/2023/50/enacted
[24] https://inspect.ai-safety-institute.org.uk/
[25] https://www.congress.gov/bill/114th-congress/senate-bill/754
[26] https://www.nist.gov/itl/ai-risk-management-framework

Citation Information:
Yi Zeng, Enmeng Lu, Xiaoyang Guo, Cunqing Huangfu, Jiawei Xie, Jin Wang, Zhengqi Wang, Dongqi Liang, Yu Chen, Gongce Cao & Zizhe Ruan. (2025). *Global Index for AI Safety (AGILE Index on Global AI Safety Readiness).* Center for Long-term Artificial Intelligence (CLAI); Beijing Institute of AI Safety and Governance (Beijing-AISI); International Research Center for AI Ethics and Governance, Institute of Automation, Chinese Academy of Sciences.
https://agile-index.ai/global-index-for-ai-safety

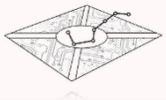
Dynamic Report Website: (With refinement across time)
https://agile-index.ai/global-index-for-ai-safety

Contributing Institutes:

Center for Long-term Artificial Intelligence (CLAI)
https://long-term-ai.center/

Beijing Institute of AI Safety and Governance (Beijing-AISI)
https://beijing-aisi.ac.cn/

Beijing Key Laboratory of Safe AI and Superalignment
https://beijing.safe-ai-and-superalignment.cn/



International Research Center for AI Ethics and Governance,
Institute of Automation, Chinese Academy of Sciences
https://ai-ethics-and-governance.institute/

Funding information:
This research was financially supported by the National Science and Technology Major Project of China (Grant No. 2022ZD0116202) and by the Beijing Municipal Science & Technology Commission.

Contact information:
Please contact contact@long-term-ai.cn for more information or with any comments and feedbacks.

Report Version:
The original version of report was released on February 7th, 2025. This updated version (v1.1.1) is released on May 17th, 2025.