# Global Index for Al 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.

Framework for Al Safety Readiness

# Global Gallinder Coll

Governance Environment National Institutions Governance Instruments Cybersecurity Status Research Status Existential Safety Preemption International participation Al Safety Incidents Mational Al Safety Institutes Networks Labs Consortiums Wational Laws & Regulations related to Al Safety Technical & Policy Frameworks for Al Safety Al Safety Publications Al Safety Patents Government fingagement Industry Engagement

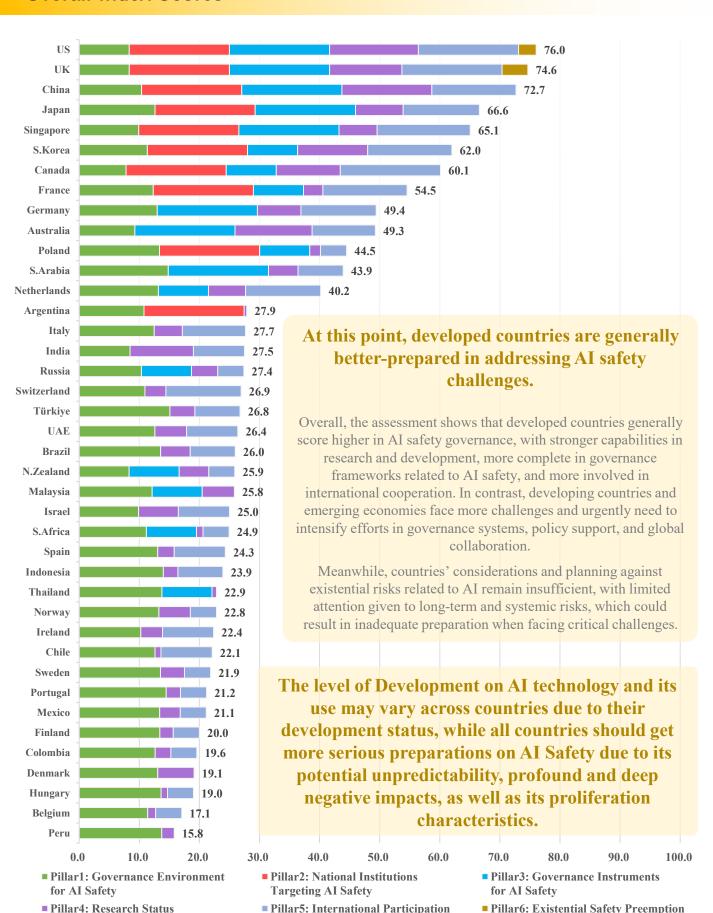
Academia & Civil Society Engagement Government Engagement Industry Engagement

# **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					
Gover	AI Safety Incidents	indicates more exposed safety issues and more governance pressure.					
National Institutions	National AI Safety Institutes/Networks/Labs/Con sortiums	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					
Gover	Technical & Policy Frameworks for AI Safety	that the instruments are more comprehensive and complete in assessing and addressing challenges.					
Research Status	AI Safety Publications	This pillar assesses the research capabilities of each country in discussing, researching, and					
Rese	AI Safety Patents	addressing the risks of AI safety. The higher the score, the greater the research attention and capabilities regarding AI safety risks.					
onal	Government Engagement	This pillar reflects a country's activity in global Al					
nati cips	Industry Engagement	<b>safety governance mechanisms.</b> A higher score indicates a larger role in promoting international					
Internationa Participatio	Academia & Civil Society Engagement	collaboration, setting safety standards, and strengthening global awareness.					
Existential Safety reemption	Government Engagement	This pillar assesses a country's strategic planning in preventing existential risks posed by AI. A higher seems reflects a more sufficient considerations.					
Existential Safety Preemption	Industry Engagement	higher score reflects a more sufficient considerations, actions and strategic planning of governments and industries in proactively addressing existential risks.					

#### **Overall Index Scores**

on AI Safety



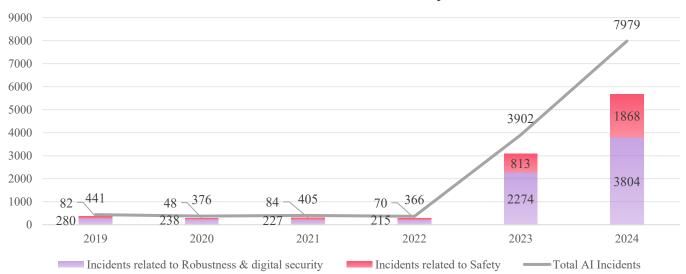
on AI Safety

# Pillar1: Governance Environment for Al Safety

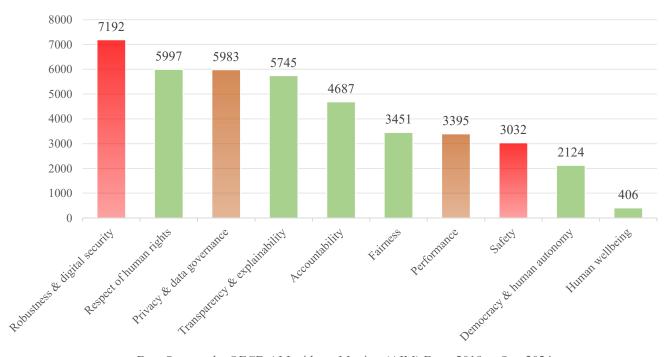
#### The global AI safety environment is becoming increasingly severe in recent years.

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

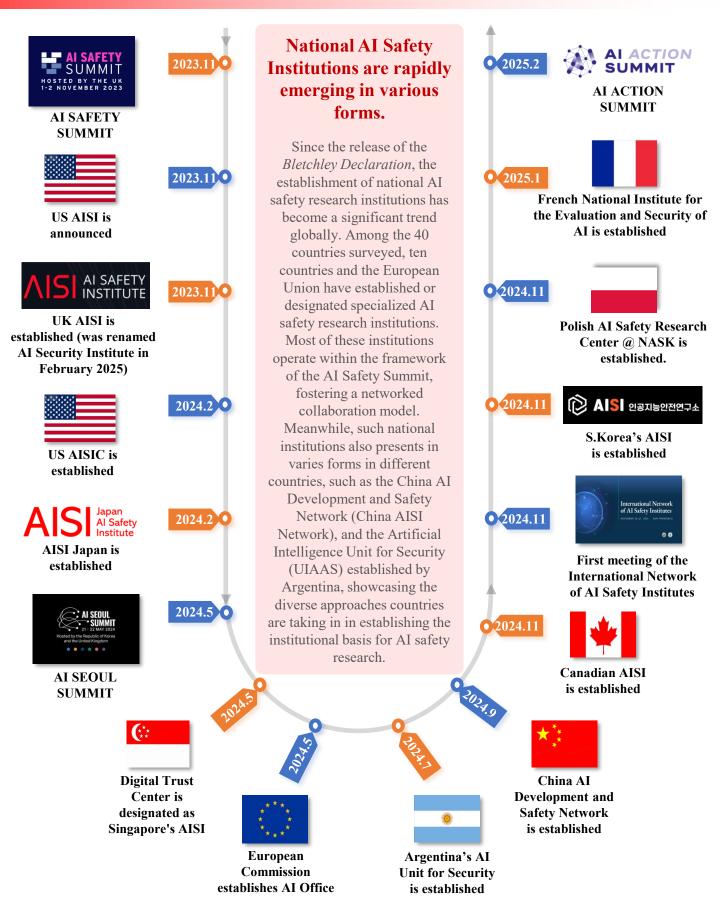


#### **AI Incidents Distribution**



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

# **Pillar2: National Institutions Targeting AI Safety**

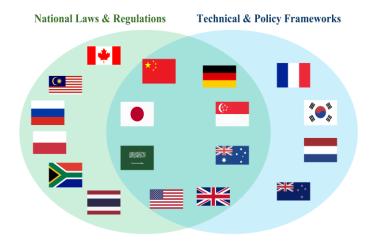


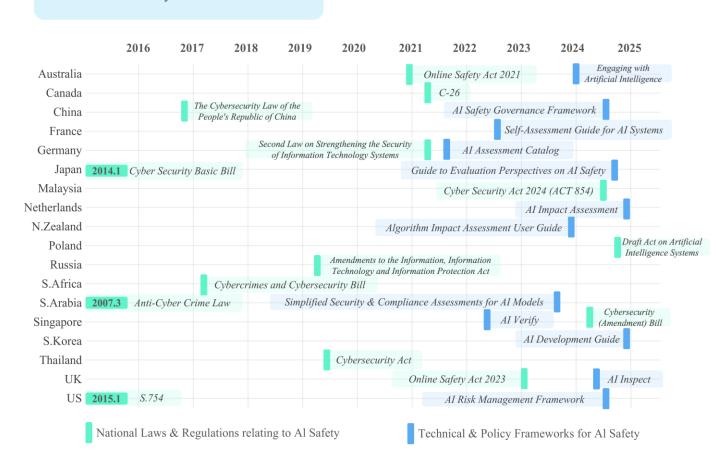
### Pillar3: Governance Instruments for Al Safety

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

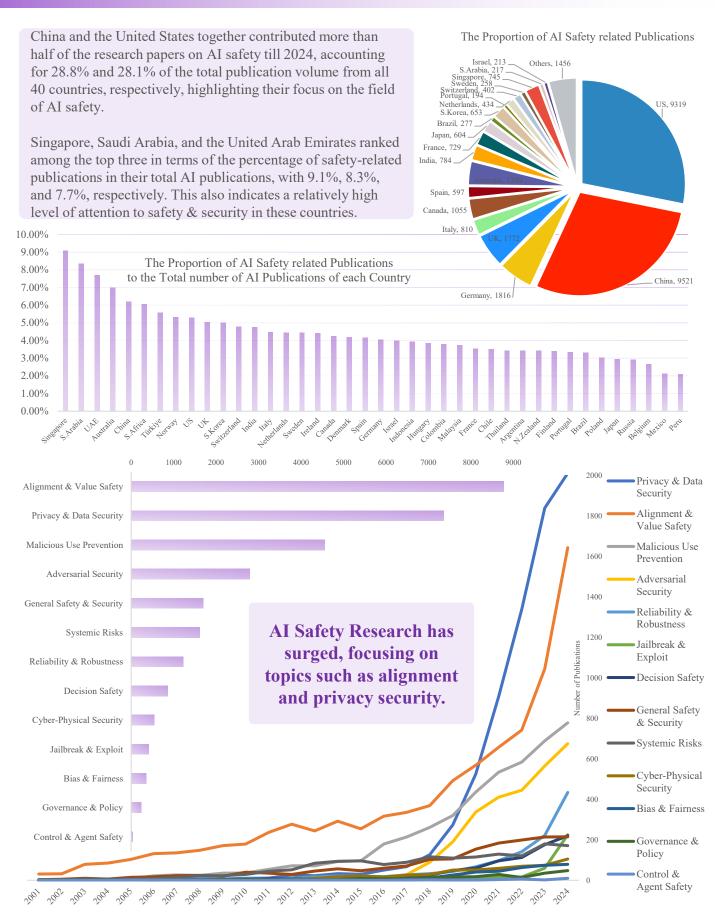
In a survey of 40 countries, 18 have governance instruments related to AI safety. Among these, 8 countries—Australia, China, Germany, Japan, Singapore, Saudi Arabia, the United Kingdom, and the United States—have both national AI-related safety laws and technical/policy frameworks in place. This highlights 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 Al Safety by Countries**





# Pillar4: Research Status on Al Safety



# Pillar5: International Participation on Al Safety

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

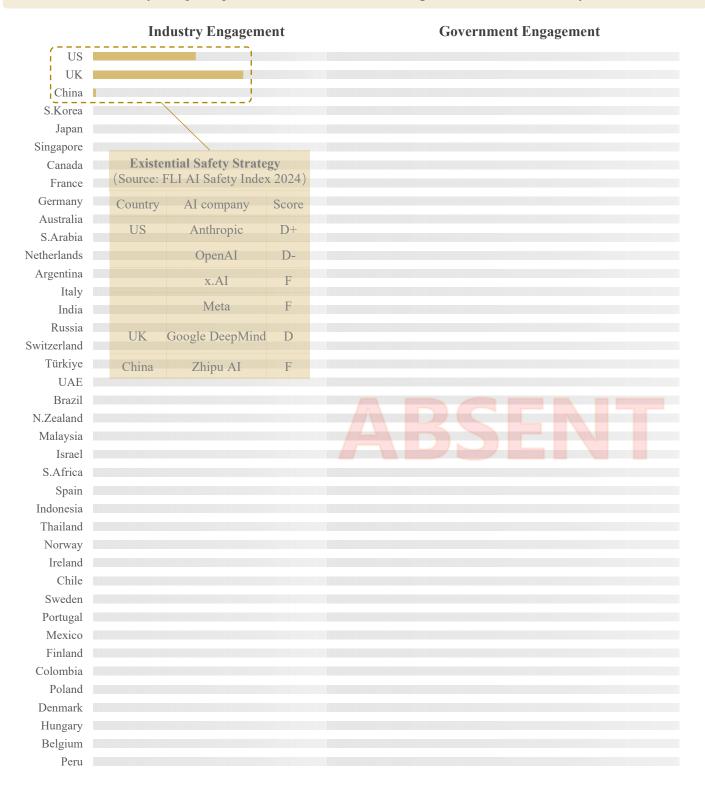
Since 2023, AI safety has emerged as a core issue in global governance frameworks. The 10 international cooperation projects surveyed reflect a multi-stakeholder approach, spanning government, industry, academia, and civil society. These efforts reflect global consensus on AI safety but also reveal the 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	Frontier AI Safety Commitments	International Dialogues on AI Safety (IDAIS)	IDAIS Beijing 2024	
US	V	$\checkmark$	$\sqrt{}$	V	$\sqrt{}$	$\sqrt{}$	V	$\sqrt{}$	$\sqrt{}$	$\checkmark$
Canada	$\sqrt{}$	$\checkmark$	$\checkmark$	×	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
UK	$\sqrt{}$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	×	$\checkmark$	$\checkmark$	$\checkmark$
Singapore	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	$\checkmark$	$\checkmark$	×	×	$\sqrt{}$	$\checkmark$
China	$\sqrt{}$	×	$\checkmark$	×	×	$\checkmark$	$\sqrt{}$	$\checkmark$	$\checkmark$	$\checkmark$
France	$\sqrt{}$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\sqrt{}$	×	×	×
S.Korea	$\sqrt{}$	$\sqrt{}$	$\checkmark$	$\checkmark$	$\sqrt{}$	$\checkmark$	$\sqrt{}$	×	×	×
Japan	$\sqrt{}$	$\checkmark$	$\checkmark$	$\checkmark$	$\sqrt{}$	$\checkmark$	×	×	×	×
Germany	$\sqrt{}$	$\checkmark$	$\checkmark$	$\sqrt{}$	×	$\checkmark$	×	×	×	×
Netherlands	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	×	$\checkmark$	×	×	×	×
Switzerland	$\sqrt{}$	$\sqrt{}$	×	×	×	$\checkmark$	×	$\checkmark$	×	$\sqrt{}$
Australia	$\sqrt{}$	$\sqrt{}$	×	×	$\sqrt{}$	$\checkmark$	×	×	×	×
Italy	$\sqrt{}$	$\checkmark$	$\checkmark$	×	×	$\checkmark$	×	×	×	×
India	$\sqrt{}$	$\checkmark$	×	×	×	$\checkmark$	×	×	×	×
Chile	$\sqrt{}$	$\checkmark$	×	×	×	$\checkmark$	×	×	×	×
Ireland	$\sqrt{}$	×	×	×	×	$\checkmark$	×	×	×	$\sqrt{}$
Israel	$\sqrt{}$	$\checkmark$	×	×	×	$\checkmark$	×	×	×	×
Spain	$\sqrt{}$	$\checkmark$	×	×	×	$\checkmark$	×	×	×	×
UAE	$\sqrt{}$	$\checkmark$	×	×	×	×	$\sqrt{}$	×	×	×
Brazil	$\sqrt{}$	×	×	×	×	$\checkmark$	×	×	×	×
Indonesia	$\sqrt{}$	$\checkmark$	×	×	×	×	×	×	×	×
S.Arabia	$\sqrt{}$	$\checkmark$	×	×	×	×	×	×	×	×
Türkiye	$\sqrt{}$	$\checkmark$	×	×	×	×	×	×	×	×
Belgium	×	×	×	×	×	$\checkmark$	×	×	×	×
Colombia	×	×	×	×	×	$\checkmark$	×	×	×	×
Finland	×	×	×	×	×	$\checkmark$	×	×	×	×
Hungary	×	×	×	×	×	$\checkmark$	×	×	×	×
Mexico	×	$\checkmark$	×	×	×	×	×	×	×	×
Norway	×	×	×	×	×	$\checkmark$	×	×	×	×
N.Zealand	×	$\checkmark$	×	×	×	×	×	×	×	×
Poland	×	×	×	×	×	$\checkmark$	×	×	×	×
Portugal	×	×	×	×	×	$\checkmark$	×	×	×	×
Russia	×	×	×	×	×	$\checkmark$	×	×	×	×
S.Africa	×	×	×	×	×	$\checkmark$	×	×	×	×
Sweden	×	×	×	×	×	$\sqrt{}$	×	×	×	×
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.



# **Appendix A: Index Indicators, Methodology & Results**

PILLAR	DIMENSION	METHODOLOGY & SOURCE				
	D1. Cybersecurity Status	ITU Global Cybersecurity Index <sup>1</sup>				
P1. Governance Environment for AI Safety	D2. AI Safety Incidents	Number of AI safety related risk cases/incidents & the GDP ratio. The data is sourced from OECD AI Incidents Monitor <sup>2</sup> , and the raw data were scored using percentile-f normalization. For specific methodologies, please refer to the appendix of the AGILE Index <sup>3</sup>				
P2. National Institutions Targeting to AI Safety	D3. National AI Safety Institutes/Networks/Labs/Consortiums	Public Information Survey				
P3. Governance Instruments for AI	D4. National Laws & Regulations related to AI Safety	Public Information Survey				
Safety	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 <sup>4</sup> , along with DeepSeek-R1 <sup>5</sup> .				
	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)				
	D8. Government Engagement	The data is sourced from the cumulative number of participations(as an official member state or with the endorsement of national experts) in relevant internationa activities as follows, and then percentile-fit normalization is conducted.				
P5. International Participation on AI Safety	D9. Industry Engagement	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				
	D10. Academia & Civil Society Engagement	·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				
P6. Existential Safety	D11. Government Engagement	Industry engagement is assessed based on the FLI AI Safety Index 2024's Existential Safety Strategy score <sup>6</sup> , _contributing 40% to the dimension's overall weight.				
Preemption	D12. Industry Engagement	Government engagement accounts for 60%, with scores determined by the government's strategic approach to AI's existential risks (Currently None).				

<sup>&</sup>lt;sup>1</sup> https://www.itu.int/epublications/publication/global-cybersecurity-index-2024 <sup>2</sup> https://occd.ai/en/dashboards/overview

<sup>3</sup> https://agile-index.ai/

https://arxiv.org/abs/2501.17805

https://chat.deepseek.com/
https://futureoflife.org/document/fli-ai-safety-index-2024/

# Appendix A: Index Indicators, Methodology & Results

			arce finding the finding to the find				Jesternents  Art. Research States  Tribats: Interne			pational Participation  Agery  Prilings, Existentian Safety  Prilings, Existentian Overall Rea			nking
		. 6	uce I	3	Istitet?	ance		Stat	N	Julal	tials	, d	SRA
		QVerne	B	ational	154	Vernety	osearch	d	A PIRAL	8 .3	stent	- eadin	Seral Index
MIT	1	G Sal		: Areting	3.00	Sal	A. R. Sale		In Sale	16. E.	IIIII	Call Re	Tall III
Country	Pilla	for	Pilla	Care	Pilla for	Pilla	OHA	Pilla	OHA	Pilla Prec	040		Jue.
	D1	D2	D3	D4	D5	D6	D7	D8-D10	D11	D12			
US		0.0	100.0	100.0	100.0	77.5	100.0	100.0	0.0	43.8	1	76	
UK		0.0	100.0	100.0	100.0	57.0	87.5	100.0	0.0	64.0	2		.6
China		33.0	100.0	100.0	100.0	87.0	92.5	84.0	0.0	0.0	3	72	
Japan		54.0	100.0	100.0	100.0	33.0	62.5	76.0	0.0	0.0	4	66	
Singapore	99.9	19.0	100.0	100.0	100.0	76.0	0.0	93.0	0.0	0.0	5	65	
S.Korea		36.0	100.0	0.0	100.0	64.0	76.0	84.0	0.0	0.0	6	62	
Canada		17.0	100.0	100.0	0.0	59.5	68.5	100.0	0.0	0.0	7	60	
France	99.0	49.0	100.0	0.0	100.0	38.5	0.0	84.0	0.0	0.0	8		.5
Germany	93.8	62.0	0.0	100.0	100.0	57.5	29.5	75.0	0.0	0.0	9	49	-
Australia		15.0	0.0	100.0	100.0	81.5	72.5	63.0	0.0	0.0	10	49	
Poland		72.0	100.0	100.0	0.0	21.5	0.0	26.0	0.0	0.0	11	44	
S.Arabia		78.0	0.0	100.0	100.0	59.0	37.0	45.0	0.0	0.0	12	43	1
Netherlands	99.2	59.0 78.0	100.0	0.0	0.0	37.0 5.0	0.0	75.0	0.0	0.0	14	27	
Argentina Italy	and the same of th	50.0	0.0	0.0	0.0	56.0	0.0	0.0 63.0	0.0	0.0	15	27	
India		3.0	0.0	0.0	0.0	61.5	65.0	51.0	0.0	0.0	16	27	
Russia	92.1	32.0	0.0	100.0	0.0	27.0	25.5	26.0	0.0	0.0	17	27	
Switzerland	91.3	40.0	0.0	0.0	0.0	42.0	0.0	75.0	0.0	0.0	18	26	
Türkiye		81.0	0.0	0.0	0.0	50.0	0.0	45.0	0.0	0.0	19	26	
	100.0	51.0	0.0	0.0	0.0	63.5	0.0	51.0	0.0	0.0	20	26	
Brazil	93.7	69.0	0.0	0.0	0.0	22.0	37.0	45.0	0.0	0.0	21	26	
N.Zealand	82.6	17.0	0.0	0.0	100.0	15.0	44.0	26.0	0.0	0.0	22	25	
Malaysia		47.0	0.0	100.0	0.0	26.0	38.0	0.0	0.0	0.0	23	25	
Israel		25.0	0.0	0.0	0.0	27.0	52.5	51.0	0.0	0.0	24	25	
S.Africa		48.0	0.0	100.0	0.0	13.0	0.0	26.0	0.0	0.0	25	24	
Spain		57.0	0.0	0.0	0.0	33.0	0.0	51.0	0.0	0.0	26	24	
Indonesia		68.0	0.0	0.0	0.0	3.0	26.0	45.0	0.0	0.0	27	23	
Thailand		66.0	0.0	100.0	0.0	9.0	0.0	0.0	0.0	0.0	28	22	
Norway		62.0	0.0	0.0	0.0	63.0	0.0	26.0	0.0	0.0	29	22	
Ireland		32.0	0.0	0.0	0.0	43.5	0.0	51.0	0.0	0.0	30	22	
Chile		82.0	0.0	0.0	0.0	12.0	0.0	51.0	0.0	0.0	31	22	
Sweden		63.0	0.0	0.0	0.0	48.0	0.0	26.0	0.0	0.0	32	21	
Portugal		74.0	0.0	0.0	0.0	28.5	0.0	26.0	0.0	0.0	33	21	.2
Mexico		75.0	0.0	0.0	0.0	11.5	29.5	26.0	0.0	0.0	34	21	
Finland		61.0	0.0	0.0	0.0	27.0	0.0	26.0	0.0	0.0	35	20	
Colombia		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	
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 Al Safety Institutions (including Institutes/Networks/Labs/Consortiums)	Date (of establishment or public disclosure)	
Argentina	The Artificial Intelligence Unit for Security (UIAAS) <sup>1</sup>	2024.7	
Canada	Canadian AI Safety Institute (CAISI) <sup>2</sup>	2024.11.12	
China	China AI Development and Safety Network <sup>3</sup>	2024.9	
France	National Institute for the Evaluation and Security of AI (INESIA) <sup>4</sup>	2025.1.31	
Japan	Japan AI Safety Institute (AISI Japan)⁵	2024.2.14	
Poland	Artificial Intelligence Safety Research Centre <sup>6</sup>	2024.11	
Singapore	Singapore's AI Safety Institute (Digital Trust Center) <sup>7</sup>	2024.5	
S. Korea	Korea's AI Safety Institute (AISI) <sup>8</sup>	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) <sup>10</sup>	2024.11.25	
US	U.S. AI Safety Institute (US AISI) <sup>11</sup>	2023.11.1	
US	Artificial Intelligence Safety Institute Consortium <sup>12</sup>	2024.2.8	

 $<sup>^1\,</sup>https://www.argentina.gob.ar/noticias/nuevas-herramientas-para-combatir-el-ciberdelito$ 

<sup>&</sup>lt;sup>2</sup> https://ised-isde.canada.ca/site/ised/en/canadian-artificial-intelligence-safety-institute

<sup>&</sup>lt;sup>3</sup> https://ai-development-and-safety-network.cn/

<sup>4</sup> https://presse.economie.gouv.fr/le-gouvernement-annonce-la-creation-de-linstitut-national-pour-levaluation-et-la-securite-de-lintelligence-artificielle-inesia/

<sup>5</sup> https://aisi.go.jp/

<sup>&</sup>lt;sup>6</sup> https://nask.pl/aktualnosci/powstanie-osrodka-badan-nad-bezpieczenstwem-sztucznej-inteligencji-w-nask/

<sup>7</sup> https://www.ntu.edu.sg/dtc

<sup>8</sup> https://www.aisi.re.kr/kor

<sup>9</sup> https://www.aisi.gov.uk/

<sup>10</sup> https://babl.ai/uk-unveils-ai-security-laboratory-at-nato-cyber-defense-conference/

<sup>11</sup> https://www.nist.gov/aisi

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

 $<sup>^1\,</sup>https://www.legislation.gov.au/C2021A00076/latest/text$ 

<sup>&</sup>lt;sup>2</sup> https://www.cyber.gov.au/resources-business-and-government/governance-and-user-education/artificial-intelligence/engaging-with-artificial-intelligence

<sup>3</sup> https://www.parl.ca/legisinfo/en/bill/44-1/c-26

<sup>4</sup> http://www.npc.gov.cn/zgrdw/npc/xinwen/2016-11/07/content\_2001605.htm

<sup>&</sup>lt;sup>5</sup> https://www.cac.gov.cn/2024-09/09/c\_1727567886199789.htm

 $<sup>^6\</sup> https://www.cnil.fr/en/self-assessment-guide-artificial-intelligence-ai-systems$ 

 $<sup>^7 \</sup> https://www.bmi.bund.de/SharedDocs/downloads/DE/gesetzestexte/it-sicherheitsgesetz-2.pdf?\_blob=publicationFile\&v=1.85 \ publicationFile\&v=1.85 \ publicationFile\&v=1$ 

<sup>8</sup> https://www.iais.fraunhofer.de/en/research/artificial-intelligence/ai-assessment-catalog.html#Receive-our-AI-assessment-catalog-free-of-charge

 $<sup>^9\,</sup>https://www.shugiin.go.jp/internet/itdb\_gian.nsf/html/gian/honbun/houan/g18601035.htm$ 

 $<sup>^{10}\</sup> https://aisi.go.jp/effort/effort\_framework/guide\_to\_evaluation\_perspective\_on\_ai\_safety/$ 

<sup>11</sup> https://www.nacsa.gov.my/act854.php

<sup>12</sup> https://www.government.nl/documents/publications/2023/03/02/ai-impact-assessment

https://data.govt.nz/docs/algorithm-impact-assessment-user-guide

<sup>14</sup> https://www.cliffordchance.com/content/dam/cliffordchance/briefings/2024/11/cb%20-%20projekt%20AI%20eng.pdf

<sup>&</sup>lt;sup>15</sup> https://web.archive.org/web/20190319122732/https://kremlin.ru/acts/news/60083

<sup>16</sup> https://www.gov.za/documents/cybercrimes-and-cybersecurity-bill-b6-2017-21-feb-2017-0000

<sup>17</sup> https://www.mcit.gov.sa/sites/default/files/2021-06/la\_004\_e\_anti-cyber\_crime\_law%20%281%29.pdf

<sup>18</sup> https://www.saifcheck.ai/

<sup>&</sup>lt;sup>19</sup> https://sso.agc.gov.sg/Bills-Supp/15-2024/Published/20240403?spm=5176.28103460.0.0.40f75d27ogoSIr

<sup>20</sup> https://aiverifyfoundation.sg/

<sup>21</sup> https://www.tta.or.kr/tta/selectBbsNttView.do?key=76&bbsNo=107&nttNo=13872&searchCtgry=&searchCnd=all&searchKrwd=&integrDeptCode=&pageIndex=1

<sup>&</sup>lt;sup>22</sup> https://thainetizen.org/wp-content/uploads/2019/11/thailand-cybersecrutiy-act-2019-en.pdf

<sup>&</sup>lt;sup>23</sup> https://www.legislation.gov.uk/ukpga/2023/50/enacted

<sup>&</sup>lt;sup>24</sup> https://inspect.ai-safety-institute.org.uk/

<sup>&</sup>lt;sup>25</sup> https://www.congress.gov/bill/114th-congress/senate-bill/754

<sup>&</sup>lt;sup>26</sup> 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. <a href="https://agile-index.ai/global-index-for-ai-safety">https://agile-index.ai/global-index-for-ai-safety</a>

**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) <a href="https://long-term-ai.center/">https://long-term-ai.center/</a>



Beijing Institute of AI Safety and Governance (Beijing-AISI) <a href="https://beijing-aisi.ac.cn/">https://beijing-aisi.ac.cn/</a>



Beijing Key Laboratory of Safe AI and Superalignment <a href="https://beijing.safe-ai-and-superalignment.cn/">https://beijing.safe-ai-and-superalignment.cn/</a>



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 <u>contact@long-term-ai.cn</u> for more information or with any comments and feedbacks.

#### **Report Version:**

The original version of report was released on February 7<sup>th</sup>, 2025. This updated version (v1.2.0) is released on July 17<sup>th</sup>, 2025.

#### **Copyright information:**

Copyright © 2025 Center for Long-term Artificial Intelligence - All Rights Reserved.