



Models for International Private Sector Cooperation on Frontier AI Safety

Conor McGlynn^{1*}, Saad Siddiqui^{2*}, Bruce Tsai³, B.A. Bacigal²,
Jeffrey Ding⁴, Robert Trager⁵, Zaheed Kara⁶, Koen Holtman⁷,
Huw Roberts⁵, Nora von Ingersleben-Seip⁸, Matt White^{9,10}

¹ Harvard University

² Safe AI Forum

³ Constellation Institute (work completed on Future Impact Group fellowship)

⁴ George Washington University

⁵ University of Oxford

⁶ Centre for the Governance of AI

⁷ AI Standards Lab

⁸ Leiden University

⁹ UC Berkeley

¹⁰ Columbia University

*denotes lead authors

The views expressed in this report are not representative of the institutions the authors are affiliated with or employed by, and should be considered purely personal opinion.

Executive Summary

Private companies can and should play a greater role in international cooperation on frontier AI safety. We specifically examine options to facilitate private sector cooperation on frontier AI safety based on expert interviews and case studies of international private sector cooperation in other sectors. While we think states have an important role to play in coordination on frontier AI safety, significant expertise on and leverage over this fast-developing technology is housed at a small number of private firms. The capabilities, safety research, and operational control over frontier systems are concentrated in a handful of developers, and any workable cooperation will have to run partly through them.

Firms in biotechnology, telecommunications, semiconductors, and aerospace already cooperate on sensitive technologies. Institutionalized private sector cooperation is well established in fields with national security implications, including between American and Chinese firms. Cooperation on frontier AI safety would not be a deviation from this norm.

Cooperation takes a variety of different forms across sensitive technology sectors, from light-touch coordination to high-touch collaboration. We identify ten different forms of private sector cooperation. Cooperation in these sensitive technology areas ranges from light-touch coordination activities like information sharing to high-touch collaboration activities such as joint research and development projects. There are a variety of precedents for international private sector cooperation on frontier AI safety.

Shared reputational exposure and pooled expertise can drive private sector cooperation on AI, though securitization and race dynamics pose particular challenges. Many of the drivers that motivate cooperation between companies in other domains are present for frontier AI safety.

However, the increasingly securitized and overtly politicized nature of frontier AI development also creates unique challenges, which any institutionalized form of cooperation will have to address.

A range of institutional models for private sector cooperation are feasible given these challenges and constraints. Even within existing institutional structures, there are a range of pathways towards formal cooperation. The models we consider are a new working group in an existing consortium, an alliance of alliances, an open source foundation, and a coordination forum.

The right model depends on the form of cooperation. Light-touch coordination, such as information and best-practice sharing, is suitable for a coordination forum or alliance of alliances. Technical collaboration on evaluations, tooling, and protocols is suitable for an existing consortium or an open source foundation.

Contents

Introduction	5
Background and Scope	5
International Cooperation for Sensitive Technologies	9
Motivations for Cooperation on Frontier AI Safety	17
Challenges for Cooperation on Frontier AI Safety	20
Institutional Models for Cooperation	22
Conclusion	29
Acknowledgments	29
Appendix: Case Studies of Private Standards Consortia	30

Introduction

Most international cooperation on frontier AI safety to date has been led by states. While private sector actors have been involved in state-led international coordination, such as the Hiroshima Code of Conduct, direct international private sector cooperation on frontier AI safety and governance has remained limited. This state-centered approach has faced headwinds from broader geopolitical tensions, particularly in the US-China relationship, as well as limited existing institutional capacity at the international level.

In this paper, we explore institutional models for private-sector-led cooperation on AI safety, which we view as complementary to state-led approaches. Our focus is on frontier AI safety, by which we mean efforts that are concerned with potential catastrophic risks from leading AI systems that could inflict severe, large-scale harm to society.

We show that private sector cooperation on sensitive technologies is not unusual and can take a variety of forms. This ranges from low-touch coordination activities like information sharing to high-contact collaboration activities, such as joint research and development projects. While there are unique drivers and challenges in the case of frontier AI safety, we suggest that a range of institutional models for private sector cooperation are feasible given these challenges and constraints. We go in depth on four possible models.

This report is based on expert interviews and research into case studies of international private sector cooperation in other sectors. We drafted an initial memo that focused on case studies of private consortia and incentives of actors to coordinate within these bodies. The authors then hosted an expert convening to discuss this, bringing together scholars in international relations and organizational theory with practitioners working in international bodies and standard setting fora. Based on the feedback received at the workshop, we adjusted the scope of the paper and conducted additional research and interviews to focus more squarely on assessing different models of cooperation specific to frontier AI safety.

Background and Scope

This section will motivate and situate the paper's focus on *institutionalized* private sector coordination amidst broader international cooperation efforts before outlining the types of cooperative activity under consideration.

International cooperation on frontier AI safety has been largely driven by states or involved states as primary participants through initiatives led by multilateral organizations. Governments have spearheaded many initiatives. These include multilateral initiatives, such as the AI Summit Series, the G7 Hiroshima Process, and the International Network of AI Safety Institutes, as well as bilateral initiatives, such as the US-China Track 1 Dialogue on AI risks, the UK-US joint AI testing partnership, and the Singapore-US Risk Cross-walk (OECD, 2026). States are a central actor in the activities of international organizations (e.g., in the UN's Global Dialogues on AI Governance; CFR, 2026).

States have not done this on their own; civil society and academia have played a crucial role.

Scientists, such as Yoshua Bengio, Andrew Yao, and many others, have supported state-sponsored cooperative efforts like the International AI Safety Report (supported by the UK AISI) and the Singapore Conference on AI (Bengio et al., 2026; Singapore Conference on AI, 2025). Experts have also convened and participated in Track 2 dialogues, including the International Dialogue on AI Safety and the AIxBio Global Forum (IDAIS, 2026; NTI, 2025).

The private sector has also played a role in state-led initiatives. Most prominently, several leading AI companies signed a set of AI Impact Commitments in Delhi in 2026. They also signed Frontier AI Safety Commitments at the AI Seoul Summit in May 2024, attended a UK AISI/GovAI conference on frontier AI safety frameworks in November of that year, and released Frontier Safety Policies around the French AI Action Summit in February 2025 (IndiaAI, 2026; DSIT, 2025; AISI, 2024).

These state-driven efforts have already had significant tangible impacts. They have elevated conversations about AI safety to the highest levels of international diplomacy, created new cooperative mechanisms like the International Network of AI Safety Institutes, and secured voluntary safety commitments from leading AI developers (METR, 2025).

However, state-led initiatives face several limitations. Constraints on state initiatives include shifts in the US government's attitude towards international cooperation on frontier AI safety and are subject to the health and functioning of interstate relations broadly, specifically US-China tensions. The Trump administration has emphasized commercial advantage, as well as bilateral trade and technology export promotion agreements, rather than bilateral or multilateral safety and risk mitigation. Vice President J.D. Vance underscored this at the 2025 French AI Action Summit, stating that he was in Paris to talk about “AI opportunity,” not “AI safety” (UCSB, 2025). The American AI Action Plan outlines several recommended measures for the US government to lead in international AI diplomacy and security, most of which center on exporting American AI to allies, countering Chinese influence in international governance bodies, and coordinating on global export control enforcement (The White House, 2025). As of mid-2026, prospects for the US–China official AI dialogue are positive but tentative: US and Chinese government sources reported that Trump and Xi had agreed to hold an intergovernmental dialogue on AI, but public details on its timing, scope, and focus on frontier AI safety risks remained limited (State Council, 2026; CNBC, 2026). China also remains excluded from the International Network for Advanced AI Measurement, Evaluation, and Science, formerly known as the International Network of AI Safety Institutes (AISII, 2026).

Existing state-led multilateral efforts, such as the AI Summit Series and the UN Global Dialogue on AI Governance, also face specific challenges. This partially stems from general limits of multilateral processes themselves, which tend to be consensus-based and result in only broad, normative statements. More specifically, the AI Summit Series has had a broadening and shifting focus—from a clear focus on frontier AI safety at the first summit in the UK to an expanded focus on AI opportunities at the French Summit, and a focus on impact at the India Summit (IndiaAI, 2026). The annual UN Global Dialogue on AI Governance, which will take place on the sidelines of existing UN meetings, will largely aim to produce co-chair summaries from each meeting, which are linked to negotiations or policy commitments from participants rather than consensus-based (UNGA, 2025; Kohler et al., 2025).

Given that development of frontier AI takes place largely in the private sector, state-led initiatives also struggle to draw in appropriate expertise on cutting-edge frontier AI R&D issues. While states have attempted to address this through the establishment of governmental centres of excellence in AI R&D, such as the AI Safety/Security Institutes (AISIs), most such centers are nascent efforts with varying levels of institutionalization, political support, and involvement in international coordination (IAPS, 2024; AI First Policy Institute, 2026).

Given the limitations of state-led efforts, private-sector-led cooperation could play an important complementary role in international cooperation on frontier AI safety. Here, “private sector” is defined broadly to include not only frontier model developers themselves, but also third-party evaluation organizations, industry associations, and other non-state actors that can play convening, norm development, or standard-setting roles. Our concern is with a variety of non-state, industry-centered institutional cooperation models. While we believe that private-public cooperation involving both governments and the private sector is feasible and desirable for effective global AI safety outcomes, we focus on private sector coordination activities given the challenges identified above.

Private-sector cooperation is itself not insulated from geopolitics but operates at different and useful levels of sensitivity and technical specificity. Frontier AI has reached such a level of political salience that it is difficult to imagine formalized cooperation taking place without attracting state attention. In addition, technical standard-setting and evaluation are not apolitical activities, but rather are loaded with normative choice about how technologies affect society (Busch, 2011). Nonetheless, private sector cooperation may be useful because it can operate at different levels of sensitivity, formality, and technical specificity than state-led diplomacy (von Ingersleben-Seip, 2023).

Coordination between these actors can take a variety of forms with different degrees of institutionalization. The less institutionalized end of the spectrum includes Track 2 dialogues, informal working groups, and confidence-building mechanisms that help firms align expectations, exchange best practices, and reduce the risks of misinterpretation or escalation. More formalized initiatives include established standards-development organizations, private consortia, and open source foundations.

This paper will focus on institutionalized models of private-sector-led cooperation. While informal dialogues and ad hoc Track 2 processes have and will continue to generate useful exchanges, they are somewhat more documented in the literature, and by their nature, they are more ad hoc (Shoker et al., 2023). By contrast, institutionalized forms of cooperation, such as standards bodies and industry consortia, are more neglected in both frontier AI safety-focused scholarship and in practice.

Moreover, there is already some precedent for such private-sector-led coordination on AI safety and security broadly as well as occasionally on frontier AI safety. In AI, several industry-backed initiatives and multi-stakeholder organizations involving academia and civil society, including ML Commons, the Frontier Model Forum, the Cloud Security Alliance, and the Partnership on AI, provide spaces for domestic (and occasionally international) cooperation on AI safety and security. This mirrors a broader pattern in other sensitive technology industries where private sector actors have often played a significant role in establishing safety standards and coordination regimes, such as in gene synthesis screening. This paper will draw significantly on case studies from different sensitive technology areas.

Finally, private-sector-led coordination could also offer different advantages to state-led approaches. Private-sector coordination will be less directly affected by (though not entirely insulated from) geopolitical and diplomatic shifts. Such coordination will also more directly involve technical experts with expertise and potentially allow for a clearer focus on technical solutions and best practice sharing. Private initiatives may also be more nimble than state-led ones as they can be established more quickly, involve fewer stakeholders, can adapt their scope over time, and update technical practices at the pace of research and development, rather than waiting for lengthy diplomatic negotiations. For a technology moving as quickly as frontier AI, speed and agility are critical in shaping norms before unsafe practices become entrenched.

There are also clear limitations and drawbacks to private-sector-led models. Unlike state-led processes, private-sector-led cooperation mechanisms lack formal authority, making participation voluntary and often uneven. Initiatives may prove fragile if commercial incentives shift, and they can be criticized for privileging the interests of powerful firms, raising legitimacy concerns. However, they are still worth exploring as a complement to state-led efforts, not as a substitute for regulation and interstate coordination.

International Cooperation for Sensitive Technologies

To inform our discussion of the AI case, we consider a range of other sensitive emerging technologies where international cooperation does take place. In this section, we outline types of cooperative activity drawing on examples relevant for frontier AI safety and other emerging technology industries where appropriate. We find that international private-sector cooperation on sensitive technologies, far from being exceptional, is in fact widespread. While the forms of cooperation and institutional models we consider emerged under different geopolitical moments and conditions (e.g., during the Cold War or in the 1990s), in many cases analogous drivers for cooperation exist in the AI case.¹

We define “sensitive technologies” as technologies that tend to be classified by state actors as relevant to national security or military use while also having widely adopted civilian uses. For our list, we have drawn on lists such as the US Critical and Emerging Technology List ([White House, 2024](#)) and the Canadian Sensitive Technology List ([Government of Canada, 2025](#)).

We distinguish two categories of cooperation: coordination and collaboration ([Bucknall et al., 2025](#)). Coordination is defined as multiple parties agreeing to pursue common goals *independently*, while collaboration refers to multiple parties *jointly* pursuing shared goals. Coordination produces a common reference point that each party then applies independently. Collaboration produces something held or operated in common, such as a shared codebase, a jointly maintained database, or a pooled research programme, which exists only through continued joint engagement.

We identify five practices within each category. These are heuristic definitions and our categorization below is not meant to be exhaustive or mutually exclusive. For example, many standardization activities involve parties drafting a standard together, but as implementation is separate and voluntary and requires no further interaction between them, it constitutes coordination by our definition. Other kinds of standards setting may involve continued engagement in implementation, in which case they would constitute collaboration. Most real world cooperation activities will cut across several practices. Our aim in distinguishing them is simply to demonstrate a range of possibilities for international cooperation on sensitive technologies.

This discussion is based on an in-depth case study analysis of cooperation across different critical and emerging technology areas. More detail on each case can be found in Appendix 1. For each practice, we provide brief examples from other industries and note any ongoing cooperation on AI.

¹ For more on the limitations and benefits of historical analogies for AI safety, see [Maas \(2023\)](#).

Table 1 – Overview of Types of Cooperative Activity

Category	Type	Current Frontier AI Safety Cooperation
COORDINATION (Parties pursue shared goals <i>independently</i>)	Information Sharing: Exchange of public or private research findings, risk assessments, or other general information.	Limited, mostly domestic
	Best Practice Sharing: Facilitating the sharing of effective methods through private exchange and/or joint publication.	Limited, mostly domestic
	Incident Sharing: Distributing information about incident reports across private sector organizations.	Limited to broader AI incident reporting, not frontier safety specific
	Pre-Standardization: Preliminary work producing provisional or informational instruments where a subject is still under development or full consensus is not yet available.	Ongoing through EU General Purpose AI Code of Practice
	Standardization: Establishing formal management processes and technical specifications.	Some activities ongoing in ISO/IEC, ITU-T, and IEEE
COLLABORATION (Parties pursue shared goals <i>jointly</i>)	Shared Tooling & Infrastructure: Contributing to common tools, benchmarks, evaluations, or open-source frameworks.	Open standards produced by companies and AISIs
	Technology Registries: Establishing a shared list or database of high-risk systems or components.	Precedents led by public rather than private actors
	Joint R&D: Co-developing pre-competitive research that is broadly beneficial for all industry actors.	Limited, mostly domestic
	Mutual Verification/Assurance: Audits, peer reviews, and certifications that verify claims against shared criteria.	Some domestic agreements among US companies
	Capacity Building: Providing support to other firms to get them up to speed on relevant safety skills and knowledge.	Limited, mostly domestic

Coordination

Information Sharing

Companies sometimes exchange research findings, risk assessments, and operational insights—sometimes confidentially under NDAs or within closed groups—to raise the floor on safety and awareness without coordinating product roadmaps or pricing. Typically, anti-trust considerations and concerns about the leaking of commercially valuable information make information sharing challenging.

Information Sharing and Analysis Centers (ISACs) are a common mechanism through which companies can share critical information with each other. In the US, ISACs emerged in the late 1990s after a Presidential Directive through which the federal government asked critical infrastructure sectors to establish sector-specific information about threats and vulnerabilities ([National Council of ISACs, n.d.](#)). The Information Technology ISAC (IT-ISAC), which aims to enable collaboration and sharing of relevant cyber threat information, does not include any AI developers and is limited to American companies ([IT-ISAC, n.d.](#)).

With respect to frontier AI safety, there is only a domestic information sharing mechanism in place. The US-based Frontier Model Forum facilitated the development and signing of an information-sharing agreement between its member firms that covers information sharing about vulnerabilities, weaknesses, and exploitable flaws (e.g., jailbreaks), threats (e.g., potential threat actors), and concerning capabilities (e.g., model autonomy) ([FMF, 2025](#)).

Best Practice Sharing

Best practice sharing can be considered as a subset of information sharing focused specifically on sharing methods and strategies that companies believe others in the industry should also consider adopting in order to protect a shared reputation. Depending on the industry, companies may also have an incentive to stay involved in best practice sharing to stay in line with best practices developed by others and avoid liability.² Best practice sharing may take the form of closed door sharing of information between members of an association or codification of “what good looks like” in guides, playbooks, and case collections.

In other critical and emerging technology areas, there are established institutions for best practice sharing. In the nuclear industry, the World Association of Nuclear Power Operators provides nuclear power operators around the world with a platform to coordinate, share, and publish best practices like pre-operational safety review guidelines ([WANO, 2025](#)).

With respect to frontier AI safety, there are some domestic best practice sharing mechanisms, limited international consortia that focus on AI safety best practices, and ad hoc international conferences. The Frontier Model Forum’s issue briefs on evaluations and technical reports on frontier safety frameworks are examples of best practice sharing amongst American firms ([FMF](#),

² Courts often assess negligence against what is “reasonable,” and industry custom or “best practice” is treated as admissible evidence of due care; courts have previously ruled that failure to adopt a readily available safety measure can be negligent even if the broader industry has not yet made it standard ([Abraham, 2009; Epstein, 1992](#)).

2024; FMF, 2025). In China, firms have also committed to a set of high-level principles through the AI Industry Alliance's (AIIA) Safety Commitments (CAICT, 2024), which are more abstract than best practices per se. There are some international consortia, such as the US-based AI Alliance, which brings together university researchers, government labs, and companies from the US, UK, Middle East, Singapore, and India, amongst others (AI Alliance, 2025). The AI Alliance runs a Trust and Safety working group, but it does not have much of a frontier AI safety focus and does not include many of the leading frontier AI developers from the US and China. Finally, multi-stakeholders conferences such as the Frontier AI Safety Conference (FAISC) held in November 2024 in San Francisco have also provided an avenue for international sharing of best practices but remain ad hoc and not institutionalized. Thus, there are still no established institutionalized mechanisms for international best practice sharing on frontier AI safety between firms.

Incident Sharing

Incident sharing can also be considered a subset of information sharing and is a relatively common practice across many industries, whereby firms report adverse events (e.g., safety failures) or “near-misses” to a common channel so others can detect, triage, and prevent similar failures (Wei & Heim, 2025). For example, in the cybersecurity industry, the Forum of Incident Response and Security Teams (FIRST) has existed since the 1990s to bring together incident response and security teams from across the world, including those from US and Chinese companies.³

There are no frontier AI safety specific incident sharing mechanisms, although there are broader AI-related incident databases. The AI Incident Database aims to document the collective history of near-misses and harms from the deployment of AI systems. It invites community submissions and is maintained by the Responsible AI Collaborative, an organization dedicated to maintaining the database, and supported by the Partnership on AI, Mozilla, and the Waking Up Foundation (AI Incident Database, n.d.). The OECD also maintains a similar database called the AI Incidents and Hazards Monitor (AIM), which was initiated by the OECD expert group on AI incidents and supported by the Patrick J. McGovern Foundation (OECD, n.d.). Companies may require separate, more secure, and private mechanisms to disclose non-public incidents to each other and to their respective home governments.

Pre-standardization

Pre-standardization refers to work that produces provisional or informational instruments rather than approved standards. Standards bodies use these deliverables where a subject is still under technical development or where consensus on a full standard is a future, rather than an immediate, possibility. ISO and IEC publish Technical Specifications and Technical Reports alongside International Standards for exactly this purpose, and 3GPP separates exploratory Study Items, which produce Technical Reports, from Work Items, which produce binding Technical Specifications.

In other fields, pre-standardization is institutionalized within standards bodies themselves, through deliverables that sit below the threshold of a full standard. In telecommunications, for example, 3rd Generation Partnership Project's (3GPP) pre-standardization entails exploratory Study Items to produce Technical Reports, which are informational without containing normative force (Baron et al.,

³ Alibaba, Tencent Cloud, ZTE, China Mobile, and National Computer Network Emergency Response Technical Team / Coordination Center of China are FIRST members.

2018). Ahead of 6G mobile systems being developed and deployed, 3GPP has used the Study Item route to define technical performance requirements and performance evaluation methodologies before committing them to specification (Larsson et al., 2024).

With respect to frontier AI safety, there are several ongoing pre-standards efforts. The EU General Purpose AI Codes of Practice (COP), which provide general purpose AI model providers a method for complying with parts of the EU AI Act, are a type of pre-standard. The COP precedes an official standard from CEN-CENELEC, the European standards body which will eventually produce a standard for complying with the EU AI Act. The COP was developed by a group of civil society experts in consultation with industry and the EU AI Office. In addition to this, technical reports about frontier safety frameworks issued by the Frontier Model Forum may also constitute a type of pre-standard that informs future formal standards-setting. Finally, while not specific to frontier AI, the Partnership on AI (PAI) has also convened companies and civil society to issue frameworks on managing risks from synthetic media (PAI, 2023).

Standardization

There are a variety of international bodies in which members come together to agree on shared processes and technical standards. These include formal standards development organizations (SDOs) like ISO or IEC, where country-level standards organizations are the voting members that approve the standards produced. They also include standards consortia, which are often created to target a single issue or technology. Some consortia seek to have a broad and open membership with voting rights including private companies, academics, and civil society representatives, while other consortia may seek to admit only selected private companies. Work in standards bodies and consortia is characterized by the members working together to write and approve standards documents or other technical specifications, where approval often happens via a combination of consensus decision making and voting. The kinds of standards that these bodies develop vary widely.

Most standards are designed to be voluntary to implement, even for the members who wrote them. Even so, sometimes regulators will declare the use of certain existing standards offers a route to partial or full compliance with the regulations they oversee. This can create a situation where voluntary standards end up being obligatory in practice, at least for parties who want to participate in a certain market. This approach is widely used in the EU, where it is commonly requested that EU standards organizations write safety standards to support the implementation of new regulations, like the EU AI Act (JTC21).

International standardization has a long history in sensitive technologies. In biotechnology, for example, ISO 35001 was the first international standard for a biorisk management system, setting out requirements for laboratories and other organizations handling biological agents to control and reduce the risks associated with their use. Published in 2019, it addresses biosecurity as well as biosafety, covering misuse, theft, and intentional release. On the consortia side, the Clinical Data Interchange Standards Consortium (see Appendix 1) began as an informal academic working group in the late 1990s and now sets clinical data formats that regulators in the United States, Japan, and China require or prefer for trial submissions.

Standardization relevant to frontier AI safety is underway. Within SDOs, the joint ISO/IEC subcommittee on artificial intelligence (JTC 1/SC 42) has published standards covering AI terminology, risk management, and management systems, while CEN-CENELEC's JTC21 is developing the harmonized European standards that will support implementation of the EU AI Act. Among consortia, the Coalition for Content Provenance and Authenticity develops technical standards certifying the source and history of digital media content, and the World Digital Technology Academy issued three AI standards in 2024, with reviewers drawn from leading developers including Anthropic, Google, Meta, and OpenAI as well as experts affiliated with both US and Chinese companies. What is largely absent is standardization aimed at frontier risks specifically: existing work concentrates on management processes, terminology, and content provenance rather than on evaluation methodologies, capability thresholds, or the safety frameworks that govern the most capable models.

Collaboration

Shared Tooling and Infrastructure

In some cases, even competitors may choose to independently or jointly develop common tooling, datasets, tests, or hosting infrastructure that everyone can use. A classic example outside of AI is the Internet Engineering Task Force, a decentralized organization where contributors collectively define and maintain protocols that underpin the Internet.

Within AI, several initiatives fit this pattern. MLCommons is a nonprofit consortium of over 100 members from industry and academia that promotes open collaboration on benchmarks, datasets, and model standardization ([MLCommons, n.d.](#)). While it includes both Chinese and Western participants, no major frontier AI developer is currently listed as a member ([MLCommons, n.d.](#)). Similarly, the Open Neural Network Exchange (ONNX), hosted by the Linux Foundation, seeks to make AI tools more interoperable. It is supported by leading companies including Baidu, Alibaba, Tencent, Amazon, and Microsoft, though its focus has been on interoperability rather than safety ([ONNX, n.d.](#)).

In addition to these group initiatives, several product initiatives directly linked to frontier AI safety and security have recently emerged. Anthropic released the Model Context Protocol (MCP), an open standard that lets AI systems connect to external tools, datasets, and APIs via a common client-server protocol (i.e., ‘USB-C for AI’; [Anthropic, 2024](#)). OpenAI and Google, as well as Chinese AI developers such as Bytedance, Alibaba, Tencent, and Baidu, have all also announced support for MCP ([OpenAI, 2025](#); [Google, 2025](#); [KrAsia, 2025](#)).⁴ The UK AISI has also developed Inspect, an open-source framework for reproducible LLM evaluations spanning reasoning, coding, multi-turn/agent tasks, behavior, and multimodal understanding ([UK AISI, 2024](#)). In Singapore, the AI Verify Foundation, a nonprofit established by the government’s Infocomm Media Development Authority (IMDA), has launched Project Moonshot, an open-source evaluation that combines benchmarking with red-teaming to test AI models against a suite of safety benchmarks ([AI Verify, 2024](#)).

⁴ MCP was donated by Anthropic to the new Agentic AI Foundation (AAIF), which is housed within the Linux Foundation.

Technology Registries

Another form of collaboration involves the creation of registries or curated lists of high-risk systems to support traceability and oversight. A leading example comes from biotechnology: the International Gene Synthesis Consortium (IGSC), a California-based 501(c)(3) formed by nucleic acid synthesis companies from around the world. IGSC members jointly maintain a Restricted Pathogen Database, which maps government export-control lists to specific DNA sequences. Members use this database to screen customer orders and verify bona fide research use, establishing a de facto industry standard for safety and compliance. In 2010, this consortium represented approximately 80% of the nucleic acid synthesis industry by volume and still claims to represent a majority of the nucleic acid synthesis market today, having steadily incorporated new entrants ([IGSC, 2024](#)).

Within AI, analogous efforts are mostly run by governments instead of the private sector. The EU AI Act requires providers of high-risk systems to log them in a Commission-run EU database before deployment, creating a structured, partially public registry across the single market ([EU, 2024](#)). In China, the CAC's algorithm filing system, strengthened under the 2023 Generative AI Measures, obliges providers of generative models to file details of their services before public release, effectively functioning as a registry of high-risk AI systems ([CAC, 2023](#)).

Joint R&D

Joint R&D involves competitors pooling resources to co-develop pre-competitive research that benefits all industry actors. This form of collaboration typically emerges when firms face common technical challenges that exceed any single company's capacity to address, or when governments provide incentives for cooperation.

In sensitive technology areas, there are established precedents for joint R&D. SEMATECH, formed in 1987 by 14 US semiconductor manufacturers in response to the US losing market share to Japan, received approximately \$100 million annually from DARPA to improve domestic manufacturing processes through collective R&D, ([Grindley et al., 1994](#)). The consortium later internationalized in 1998 ([SHMJ, 2022](#)). In satellite communications, Intelsat was founded in 1964 as an intergovernmental consortium to jointly develop and operate a global communications satellite system ([IEEE, n.d.](#)).

With respect to frontier AI safety, there are no clear existing examples of joint international R&D.

Mutual Verification/Assurance

Mutual verification involves audits, peer reviews, or certifications in which organizations assess each other's practices against shared criteria. This form of collaboration is particularly valuable in industries where a single firm's failure could harm the entire sector's reputation, creating incentives for collective quality assurance.

The nuclear power industry provides the most developed model through the World Association of Nuclear Operators (WANO). Established in 1989 following Chernobyl, WANO operates on the premise that a serious nuclear incident anywhere affects operators everywhere (WANO, n.d.). Peer reviews became a member obligation in 2006, and every commercial nuclear power plant worldwide had received at least one review by 2010.

With respect to frontier AI safety, the most significant example is the joint evaluation exercise conducted by OpenAI and Anthropic in August 2025. Each company evaluated the other's public models using their internal misalignment-related tests and released their results in parallel posts (OpenAI, 2025; Anthropic, 2025). Evaluations focused on propensities related to sycophancy, self-preservation, and supporting human misuse.

Capacity Building

Another form of collaboration involves safety capacity building, whereby established industry actors provide support to new entrants or less experienced firms to help them meet safety standards. For technologies with significant safety risks, ensuring all participants can adhere to best practices is in the collective interest of the sector.

WANO again provides the most developed model through its New Unit Assistance (NUA) program. The program offers 18 modules covering topics from Nuclear Safety Culture to Operational Decision-Making, which can be tailored to members' needs along the project timeline (WANO, n.d.). WANO recommends that new nuclear operators join years before initial criticality to gain maximum value.

Most of the industry-led capacity building mechanisms for AI safety are domestic or regional (e.g., the FMF's work to support companies developing risk management practices).

Motivations for Cooperation on Frontier AI Safety

Based on our literature review, we identify three key drivers for private sector cooperation and consider how they may apply in the case of frontier AI safety.

Interoperability and Market Access

Cooperating on standards helps ensure compatibility between companies' products or processes, expanding the overall market and enabling firms to buy/sell from each other more easily. By aligning on technical specifications or quality benchmarks, companies can tap into larger networks of suppliers and users, entering new markets with fewer barriers ([British Standards Institute, 2015](#)). One downside of interoperability is that it can help create strong network effects which potentially lock-in incumbents ([Swann, 2000](#)). On the other hand, compatibility focused interoperability, such as open standards that allow different parts of a technology stack to be interoperable with each other, can improve competition between firms as it pushes against monopolistic lock-in effects ([Farrell & Klemperer, 2007](#)). The effectiveness of interoperability as a driver may be modulated by market dynamics and investor expectations. If investors expect significant competitive moats, companies may be less interested in contributing to standards that could reduce this moat.

Interoperability emerges as the most prevalent driver of cooperation, though it manifests differently across technology domains. In telecommunications, interoperability refers to enabling global communication protocols that allow different network equipment to function together across the world. Organizations like the 3rd Generation Partnership Project (3GPP) were formed in the late 1990s as telecommunications providers worldwide sought to standardize protocols to enable global communication ([Baron et al., 2018](#)).⁵ For internet governance bodies, such as the Internet Engineering Task Force (IETF), World Wide Web Consortium (W3C), and Internet Corporation for Assigned Names and Numbers (ICANN), interoperability centers on ensuring consistency of core internet protocols and web standards that form the foundation of the global internet ([Teleanu, 2021](#); [Gordon, 2023](#); [Mueller, 2023](#); [IETF, n.d.](#)).⁶ On the other hand, the semiconductor industry prioritizes interoperability in the form of compatibility between components in an extremely complex global supply chain. Organizations like JEDEC, which began as the Joint Electron Tube Engineering Council in 1944 and became crucial for memory standards, develop standards addressing both physical specifications (such as wafer dimensions and memory interfaces) and testing methodologies that ensure components from different manufacturers can work together seamlessly ([JEDEC, n.d.](#)).

Interoperability also takes different forms depending on technology maturity. In established fields like telecommunications and semiconductors, standards primarily ensure compatibility between existing systems. For emerging technologies like quantum computing, early interoperability efforts

⁵ The 3GPP brings together domestic standards setting organizations from Japan, North America, China, Europe, and Korea to ensure mobile devices can communicate regardless of manufacturer or location.

⁶ The IETF evolved from the early ARPANET technical community in the 1980s and maintains an open structure that emphasizes technical merit rather than formal voting procedures. The W3C, established by Sir Tim Berners-Lee in 1994, focuses on standards for the World Wide Web that ensure long-term growth and compatibility.

focus on pre-standardization activities, developing shared terminology that enables stakeholders to "speak the same language." These early efforts occur at bodies like ISO/IEC JTC 1, which set up a working group on quantum computing in 2020 (Cory, 2024).

In the case of frontier AI safety, interoperability is likely to be a weaker driver of cooperation than it is for other technologies. There are examples of interoperability-drive standards (e.g., the Model Context Protocol), but the commercial incentive to ensure different systems developed by frontier AI companies from different jurisdictions work together is less pronounced than in the case of telecommunications or the internet where interoperability was core to the functioning of the technology. There is also not a strong incentive to manage a complex international supply chain, as in the semiconductor case. Rather than simply ensuring systems can work together, as in the case of common agent infrastructure, interoperability in frontier AI safety standards could instead focus on establishing interoperable evaluations using shared frameworks or common risk taxonomies (Chan et al., 2025; UK AISI, 2024; Slattery et al., 2024; IMDA, 2023). However, lacking the traditional commercial incentive of interoperability, this kind of interoperability would be better understood as a potential consequence rather than a driver of cooperation.

Shared Identity or Reputation Amongst Companies

Companies in the same industry may share an "intangible commons" where errors by one firm can damage all firms through spillover harm. Such industries represent a "reputation collective" (Ding & Li, 2026), as one firm's accident damages the reputation of all others. Following major crises, this effect may intensify as stakeholders struggle to differentiate between individual firms, leading them to apply industry-wide sanctions based on collective perceptions instead (Barnett & King, 2008). Industry self-governance through standards is a way to mitigate this risk by managing the reputation of the industry as a whole (Yue & Ingram, 2015). In other sectors, such as nuclear power, this shared reputation has contributed to the formation of international industry associations, which have contributed to improved safety standards (Ding & Li, 2026).

In our research, we found that this driver manifests most powerfully when industries face identifiable external threats. In biotechnology, the International Gene Synthesis Consortium (IGSC) was established by US-based nucleic acid synthesis companies and has grown to approximately 40 members globally, including Chinese companies (Tucker, 2010). The catalyst for this cooperation was fear of reputational damage from potential misuse of DNA synthesis technology and the threat of government regulation, leading competitors to collaborate on nucleic acid sequence screening protocols.⁷ Similarly for space technologies, safety concerns drive cooperation, as mission failures could impact public trust in the entire sector. The Consultative Committee for Space Data Systems (CCSDS) started in 1982 with a focus on communications protocols that has expanded to include 11 member agencies, 32 observer agencies, and over 140 industrial associates (CCSDS, 2010). Finally, in the aerospace sector, the International Aerospace Quality Group (IAQG) was established in 1998 to bring together companies from the aviation, space, and defense industries to develop quality and safety standards to ensure that the reputation of the aerospace industry was maintained globally (IAQG, n.d.).

⁷ The IGSC was not the only effort that came about due to these pressures. It was preceded by the International Association of Synthetic Biology (IASB), which was focused on a greater role for government and civil society, alongside a fully automated screening process. Two leading companies were not in favor of the IASB approach and started IGSC alongside several other players, keeping meetings closed to just members and only admitting companies that crossed a particular revenue threshold (Heron, 2023).

Shared reputation is likely to be a significant driver for cooperation on frontier AI safety. The International Gene Synthesis Consortium provides an interesting comparison, where DNA synthesis companies collaborated to prevent misuse of their technology because of collective reputational risks. Frontier AI developers could face similar industry-wide reputational considerations. Most frontier AI models are trained using very similar architectures and largely similar data, and innovations are often freely published in preprints that proliferate across different companies. Therefore, they also share many of the same weaknesses, such as susceptibility to emergent misalignment and universal jailbreaks (Betley et al., 2025; Anthropic, 2025; AISI, 2026). AI developers also utilize many of the same open-source frameworks, which leaves them collectively vulnerable. An example is the Ray AI framework, which was hacked in 2024, and affected OpenAI amongst others (Arghire, 2024). A countervailing factor here is that frontier AI companies are often described in national terms (e.g., as “Chinese” or “American” companies), which might limit the degree to which they perceive themselves as sharing a reputation.

Opportunity to Pool Expertise

No single firm may have all the know-how or resources to develop complex new standards or technologies alone. By working together, firms can selectively share R&D results, learn from each other, and co-create standards that push the state of the art. Studies suggest that firms participating in consortia are more likely to cite each other's patents and may see improved innovation performance (Delcamp & Leiponen, 2011; Jiang et al., 2019).

In highly specialized fields with limited experts, such as confidential computing, knowledge pooling tends to be comprehensive and deep, with competitors openly sharing technical insights to advance collective understanding. The Confidential Computing Consortium (CCC), established in 2019 under the Linux Foundation, brings together experts from companies like Intel, ByteDance, and Huawei to discuss complex, niche issues with counterparts in other companies (CCC, n.d.).

Pooling expertise is likely to be a significant driver for cooperation on frontier AI safety. The specialized nature of frontier AI research means that technical expertise is concentrated in a relatively small number of companies and research institutions. While some of this talent lies outside frontier labs, for example in evaluation providers like METR, the level of knowledge within labs is substantial. With regards to frontier safety, the total number of leading researchers working on the problem is just a small fraction of the overall research pool (ETO, n.d.).⁸ There may therefore be parallels to how the Confidential Computing Consortium enables competitors to share technical insights in a field with limited experts. As in other industries where pooling knowledge is common, frontier AI companies will have to manage the risk of technology transfer.

⁸ According to CSET's ETO tool, from 2018 to 2023, only 2% of all AI articles were focused on safety, growing from 3,500 articles to 14,500 articles per year during this time.

Challenges for Cooperation on Frontier AI Safety

Any form of private sector cooperation faces difficulties inherent to the conflicting interests of competing firms. However, there are also a number of challenges specific to frontier AI development that make cooperation more difficult in this domain.

First, broader US-China strategic competition has created an increasingly hostile environment for cross-border AI safety cooperation. US-China relations have generally worsened throughout the 2010s and early 2020s, leading to a reduction in some types of private sector exchanges. For example, it appears that the US-China Internet Industry Forum, which was previously co-organized by Microsoft, has been put on hold after 2015. The forum previously convened senior political and business leaders to discuss issues related to internet governance from 2007 to 2015. The last forum was held in Seattle and attended by President Xi Jinping ([Hormats, 2011](#); [Hormats, 2013](#); [ISC, 2015](#)).⁹

The linking of frontier AI development to national security and strategy has made it challenging for private sector actors to cooperate. Leaders from both countries have emphasized the strategic nature of AI, alongside the need for their country to lead in an AI race.¹⁰ Chinese President Xi Jinping has emphasized that AI is a “strategic technology leading a new round of scientific and technological revolution and industrial transformation” ([Xinhua, 2025](#)). Academic experts such as Gao Wen (leader of a Pengcheng Lab, a national-level AI lab) have written that “whoever masters advanced AI technology will hold the initiative in future competition” ([Gao Wen, 2025](#)). In the US, the former Biden Administration issued the National Security Memorandum on AI, highlighting the close coupling of frontier AI with national security uses ([White House, 2024](#)). Under the Trump administration, full-stack competition on AI has been a central theme, with David Sacks, White House AI and Crypto Czar, stating that “the American AI stack needs to be unleashed to compete” ([Sacks, 2025](#)).

Regulatory changes responding to the two trends described above have also constrained cooperation between American and Chinese frontier AI companies. For example, Huawei’s listing on the Commerce Department’s entity list resulted in the withdrawal of US companies from several standards bodies where Huawei was also a member. While regulations were updated throughout the Biden administration to allow standards-setting activity to continue, the 2–3 year period of limbo may have led to permanent fractured standards-setting blocs in some areas ([Cory, 2023](#)). Separately, the listing of frontier AI companies from China and AI safety organizations, such as the Beijing Academy of AI, have also impacted the ability of Western companies to engage with these companies, as companies are unsure whether engaging with entity-listed entities in talks or discussions about AI safety goes against export controls ([Chen, 2025](#)).

⁹ There may be ongoing private versions of the event but there are no public records after 2015.

¹⁰ It is unclear if both countries are in fact running the same AI race. For more on this see [Chinatalk, 2025](#).

There is also the inherent challenge of teasing apart AI safety-related issues from general AI capability or development-related issues, some of which may be proprietary or sensitive. For example, specific AI safety techniques, such as reinforcement learning from human feedback (RLHF), were first conceived as a safety technique but played a significant role in making pre-trained models function usefully as chatbots like ChatGPT ([Bucknall et al., 2025](#)).

Institutional Models for Cooperation

Cooperation between private sector actors can take a wide range of shapes, from light touch coordination to higher contact collaboration. In this section, we consider four institutional models that could accommodate these forms of cooperation, in light of the particular motivations and challenges that we identified for frontier AI safety. The choice of institutional model should be driven by the kind of cooperation desired, and we assess some of the advantages and disadvantages of each of these models.

We specifically consider a new working group in an existing standards consortium; an alliance of existing regional frontier bodies (an alliance of alliances); a new group in an open source foundation; and a high-level body to coordinate existing cooperation workstreams (a coordination forum). This list is not exhaustive of possible models, and the list could be expanded to include a variety of possible new institutions. These four models were chosen on the basis of interviews with stakeholders working on frontier AI safety. While these institutional models can each be adapted to different purposes, they broadly differ in terms of the level of cooperation they enable and the diversity of participants they include.



The strengths and weaknesses of each of the four models can be assessed along different axes. First is the depth and breadth of frontier AI safety knowledge within an institutional model. Effective safety cooperation requires technical understanding of frontier systems, their capabilities, failure modes, and mitigation strategies. This knowledge is primarily found among researchers at leading AI

labs, specialist academics, and third-party evaluators. Second is speed and flexibility: how quickly can a mechanism adapt to emerging risks and evolving capabilities? Frontier AI development proceeds rapidly with safety challenges that arise unpredictably, requiring approaches that can pivot faster than traditional governance processes.

Political and institutional feasibility is another factor that affects whether a cooperative mechanism can realistically operate within existing constraints. Private sector-led safety efforts must navigate complex international regulatory environments, intellectual property concerns, and competitive dynamics across jurisdictions in order to be effective. An institution also needs buy in from relevant stakeholders and alignment with key participants' motivations, particularly frontier AI developers. Successful cooperation depends on voluntary engagement from companies balancing development timelines and commercial interests. Outputs must be viewed as authoritative and trustworthy by diverse stakeholders. For safety standards and protocols to influence industry practices, inform regulation, and build public confidence, they need to be broadly recognized as credible and representative.

Finally, there is the question of how robust an institutional model is to the geopolitical tensions that mark contemporary AI development, and whether it can sustain international cooperation during periods of geopolitical tension. As US and Chinese entities push the frontier, continuing safety coordination between these entities will become ever more important, even as technological competition and national security concerns may raise further barriers to formal cooperation.

There is no single perfect institution, and different institutional models lend themselves to different issues. There are also trade-offs between different kinds of institutional design. For example, expanding the diversity of participants involved may increase the perceived legitimacy of a body but reduce its speed and flexibility. Selectivity in membership and confidentiality may be necessary for high-touch collaboration on sensitive topics, but this may increase the political cost for firms to participate compared to more open and transparent institutions. Our analysis here is not intended to promote any one particular model, but rather to provide a conceptual framework for thinking about these questions of institutional choice and design.

	Existing consortium	Alliance of alliances	Open source foundation	Coordination forum
Institutional base	Existing body, with a new working group formed inside it	New body federating existing national and regional industry associations	New project hosted by an existing open source foundation	New lightweight body convening existing institutions
Range of participants	Selected member firms, admitted under the host body's existing rules	Industry associations, participating as representatives of national or regional blocs	Firms, individual contributors, and researchers participating on technical merit	Institutions of varied type, including SDOs, AI safety institutes, consortia, and frontier developers
Typical cooperation practices	Pre-standardization; standardization; mutual verification and assurance	Best-practice sharing; information sharing	Shared tooling and infrastructure; joint research and development	Information sharing
Precedents	World Digital Technology Academy; Cloud Security Alliance; C2PA; OWASP	3GPP; International Council of Quantum Industry Associations	Linux Foundation; Agentic AI Foundation; AI Verify Foundation	World Standards Cooperation; Metaverse Standards Forum

Working Group in Existing Private Consortium

One way of deepening private sector cooperation on frontier AI safety is through the institutional infrastructure provided by existing bodies. Some scholars have argued that traditional SDOs may struggle with the speed of frontier AI development, dominance of larger technology companies in standards-setting processes, and the value-laden nature of some frontier AI standards-setting on issues such as fairness (Roberts & Ziosi, 2025). They have argued that division of labor and coordination between three types of standards bodies may better address the unique challenges of setting frontier AI standards—with traditional standards organizations focusing on high-level process and management standards for well-established "trailing-edge" issues, business/private consortia (such as the Frontier Model Forum) developing granular technical specifications for "leading-edge" issues using their agility, and technical expertise and national AI institutes (such as the UK AISI)

bridging these two levels with technical capacity and regulatory connections (Roberts & Ziosi, 2025). Across these processes, coordination with actors from outside the West remains a challenge as key actors like China are not involved in mechanisms such as the network of AISIs.

There are several existing private consortia that play a role in AI standards. These include global cybersecurity-focused consortia that have also moved into AI standards, such as the Cloud Security Alliance (CSA), Open Worldwide Application Security Project (OWASP), and content provenance consortia such as the US-based Coalition for Content Provenance and Authenticity (C2PA), which develop technical standards that certify the source and history of digital media content (CSA, 2024; OWASP, 2024.; C2PA, n.d.).¹¹ Dedicated AI-focused consortia have also been set up in recent years including the World Digital Technology Academy (WDTA). WDTA was founded in 2023 in Geneva and issued three AI standards in 2024. Reviewers on its standards include representatives from leading companies, consortia members such as Google, Meta, and OpenAI, and expert committee members like Geoffrey Hinton and Fei Fei Li (WDTA, n.d.). WDTA has experts affiliated with both US and Chinese companies.

The strengths of a working group in an existing consortium include credibility, an existing membership base from industry, and experience facilitating industry exchanges that would provide a high base of expertise. A new working group would likely benefit from institutionalized relationships and reputation. Consortia that include US and Chinese members could also provide cover to companies that are already cooperating in other areas to begin coordinating on frontier safety.

The main weakness would be potential limitations placed on a new working group by existing institutional features. Companies may perceive there to be reputational or political risks in a consortium deviating from existing missions to focus on frontier safety. Members may resist the formation of a frontier AI safety working group if it is perceived as deviating from the consortium's core mission. For example, a consortium focused on developing benchmarks and evaluations for reliability may be resistant to expanding its mission to include catastrophic risks.

Frontier AI Alliance of Alliances

An alliance of alliances refers to a model where industry bodies from different geographies form an international association that coordinates across their various memberships.¹² Such an alliance can take several forms. Several industry associations can jointly set up a formal standards-setting organization, as was the case for the 3rd Generation Partnership Project (3GPP), an international standards body for telecommunications formed by domestic standards setting organizations from Japan, North America, China, Europe, and Korea in the late 1990s (Baron et al., 2018).¹³ 3GPP consists of a structured system of three Technical Specification Groups (TSG), with several Working Groups within each TSG addressing specific technical challenges under each TSG (3GPP, n.d.).

¹¹ OWASP and CSA are focused on AI agents, LLM vulnerabilities, and privacy; OWASP's initiatives include the Agentic Security Initiative and AI Security Guide, while CSA's AI Safety Initiative and Multi-Agent Systems Framework address AI risks and capabilities.

¹² An SDO like ISO could be seen as an alliance of alliances, as it brings together national standards bodies. Our concern here is with more focused alliances bringing together consortia and industry bodies from particular sectors.

¹³ In 2014, a newly formed body called the Telecommunications Standards Development Society of India became the seventh member.

Alliances can also have a less formal structure and fewer defined goals. For example, the International Council of Quantum Industry Associations was set up in 2023 through the signing of a Memorandum of Understanding between the domestic associations of the US, Canada, Europe, and Japan. These associations had been meeting informally for the past year and decided to come together to strengthen collaboration and best practice sharing between the different domestic associations (QED-C, 2023).

The alliance-of-alliances approach has the advantage of being suitable for a more lightweight model, focused on low-touch forms of coordination such as information or best practice sharing. This increases the geopolitical feasibility of this approach, potentially making participation more palatable to frontier AI developers in the US and China. The lack of formalization allows for quick set up and flexible adaptation to changes in the external environment.

However, the informality of the model also has its drawbacks. This model may not enjoy the entity-list exemptions granted to SDOs, making it significantly harder for US companies to engage with entity-listed Chinese companies through this forum (Cory, 2023). Moreover, examples from other alliances of alliances also point to the challenges of navigating geopolitical tensions. For example, the 3GPP has faced significant internal challenges with controversies surrounding Chinese government pressure on Lenovo and other Chinese companies to back Huawei's standards proposals. While not linked to the model itself, it seems plausible that a model where individuals participate as representatives of a regional or national bloc are more vulnerable to geopolitical tension (Coopersmith et al., 2021).

Open Source Foundation

An open source foundation refers to a nonprofit body that hosts and stewards collaboratively developed software, providing the legal, governance, and intellectual property infrastructure within which contributors from competing firms can maintain a shared codebase. Unlike standards bodies, foundations do not typically issue written specifications for members to implement separately. Their characteristic output is executable code that is held in common and persists only through continued contribution. This makes them a vehicle for the higher-contact forms of collaboration identified above, particularly for shared tooling and infrastructure. Foundations vary in scope, from single-project bodies to multi-project hosts, such as the Linux Foundation, which offers a neutral home for projects that no single company wants to own.

Foundation-hosted projects have proved able to sustain technical cooperation between US and Chinese firms in other sensitive technology domains. The Confidential Computing Consortium, established under the Linux Foundation in 2019, brings together experts from Intel, ByteDance, and Huawei among others. The Open Neural Network Exchange, also hosted by the Linux Foundation, is supported by Baidu, Alibaba, and Tencent as well as Amazon and Microsoft. In both cases, participation is organized around technical contribution rather than national representation.

Within AI, the clearest instance of this institutional model is the Model Context Protocol (MCP). Anthropic developed MCP and donated it to the Agentic AI Foundation—a directed fund established under the Linux Foundation—in December 2025 (Linux Foundation, 2025). Anthropic and OpenAI co-founded the foundation together with Block (Anthropic, 2025). Google, Microsoft, Amazon Web Services, and Cloudflare are among its members, which places several of the largest AI developers

inside a single body. Chinese developers including ByteDance, Alibaba, Tencent, and Baidu have separately adopted the MCP, though they are not members of the foundation. There are some OSF-like groups that focus on AI safety, although none focus on frontier AI safety per se. These include MLCommons (previously called MLPerf), a nonprofit focused on creating ML benchmarks, open datasets, and best practices, as well as AI Verify, a foundation set up by the Singaporean government focused on developing testing tools.

The OSF approach is effective in sectors where stakeholders (particularly companies) have an incentive to work together to build interoperable protocols. [Nanni \(2022\)](#) suggests that major Chinese companies like Huawei recognize that “it would be economically disadvantageous... to create an architectural 'island' separated from the existing IP-based Internet.” While interoperability is unlikely to be as strong a driver of cooperation in frontier AI safety as it is in other domains, some types of frontier AI safety standards, such as those involving multi-agent interaction, may require technical interoperability that is beneficial for all actors.

The OSF approach may be able to facilitate higher touch collaboration among US and Chinese firms through the openness of the institutional structure. Because contributions are made in public and assessed on technical merit, participation does not require a firm to sign a political commitment or to be seen entering a formal safety arrangement with a rival. The logic resembles the transparency approach adopted by the International Gene Synthesis Consortium, whose members treat everything shared as if it were public, on the view that this reduces the perceived sensitivity of taking part.

However, given the premium placed on the talent at frontier AI companies, it is unlikely that the most skilled staff would have significant amounts of time to build open source standards or protocols. Open source projects would also not touch on certain frontier AI safety topics that have national security implications. However, given the involvement of Anthropic and UK AISI staff (organizations with significant frontier AI safety expertise) in building open source agent communication and evaluation frameworks respectively, it seems plausible that at least a moderate amount of frontier safety expertise would be possible in this group.

Coordination Forum

At present, international cooperation on frontier AI between private actors takes place to varying degrees within different institutional settings: in SDOs like ISO, IEC, and ITU; through the International AI Safety/Action Summits; through the nascent AISI network; and in Track 2 dialogues, among others. These forums intersect in membership and often interact among themselves, but there is at present no body in which companies can coordinate their efforts in these different settings.

A coordination forum refers to a body that gathers representatives from companies or organizations operating with distinct yet overlapping mandates. In general, this involves coordination (rather than collaboration) in the form of information sharing about ongoing workstreams. SDOs often take part in coordination forums to prevent conflicting or redundant efforts. For example, World Standards Cooperation ([WSC](#)) is a coordination forum that brings together the heads of ISO, IEC, and ITU for strategic coordination on cross-cutting standards setting issues. The ISO/IEC Joint Technical Committees are also in a sense very integrated coordination forums, where the activities are not

limited to sharing information but to developing jointly issued standards. The 2025 International AI Standards Summit, which brought together ISO, IEC, and ITU, is an example of a coordination forum for AI, though it is limited in scope given that it was a one-off event.

The Metaverse Standards Forum (MSF) is an example of a coordination forum that is made up mostly of companies rather than SDOs. MSF does not issue standards itself but rather provides formal structures within which members can coordinate requirements and resources to support the work of other standards organizations working on relevant issues. Membership of MSF is open to different companies and standards organizations associated with developing metaverse infrastructure and products, rather than being tailored around a particular technical domain. Funding for MSF comes from its member companies.

The main strength of a coordination forum is that it would be a relatively light-touch form of cooperation, making it potentially more robust to geopolitical tensions. A coordination forum for frontier AI could help to streamline and build upon existing cooperation without encountering the institutional barriers of operating within existing organizations and without the political barriers to establishing an entirely new organization.

The weakness of this model is that, by involving a wider range of institutional actors (e.g., SDOs, government bodies like AISIs, frontier AI labs), decision making within a coordination forum is likely to be slower, and consensus will be harder to build than in a smaller closed organization. This means that members will likely find it harder to find common positions on more contentious issues.

Conclusion

Private sector cooperation on sensitive technologies is not abnormal and can take place in a variety of ways. International private sector cooperation takes place across sensitive technology areas, from biotechnology to nuclear to aerospace. Cooperation ranges from low-touch coordination activities like information sharing to high-contact collaboration activities, such as joint research and development projects. International cooperation between leading companies on frontier AI safety would not be a deviation from the norm.

There are drivers but also unique challenges for cooperation on frontier AI safety. Many of the drivers that motivate cooperation between companies in other domains, such as shared reputations and beneficial expertise sharing, are present for frontier safety. However, the increasingly securitized and overtly politicized nature of frontier AI development also creates unique challenges, which any institutionalized form of cooperation will have to address.

A range of institutional models for cooperation are feasible given these challenges and constraints. Even within existing institutional structures, there are a range of pathways towards formal cooperation. Each one of these models we consider—a working group in an existing consortium; an alliance of alliances; an open source foundation; and a coordination forum—has unique strengths and weaknesses as forums for cooperation.

No single institutional model is likely to be appropriate for all forms of private-sector cooperation on frontier AI safety. Lower-touch coordination activities, such as information or best-practice sharing, are more likely to be feasible in lightweight coordination forums or alliance-of-alliances models. More technical activities, such as evaluation tooling, benchmarking infrastructure, and interoperable reporting formats, are more naturally suited to existing consortia or open-source standards organizations. The institutional question is therefore not which model is best in the abstract, but which model is appropriate for the sensitivity, technical specificity, and trust requirements of a given cooperative activity.

Acknowledgments

We would also like to thank Konrad Seifert, Renata Dawn, Stephen Clare and Jacob Pratt for helpful feedback on previous versions of this report.

Appendix: Case Studies of Private Standards Consortia in Critical and Emerging Technologies

Given the range of complex technologies we covered, our summaries are not intended to be comprehensive. We attempted to check our understanding with relevant experts to ensure that our high-level takeaways were accurate and also supplement public sources with interviews with members of the mentioned organizations.

We indicate our levels of confidence in each of the summaries below.

Biotechnology

Confidence level High

Biotechnology is a very broad area encompassing synthetic biology, genomic engineering, bioinformatics, cellular and viral systems engineering, and biomanufacturing technologies. We understand biotechnology standards are related and often overlapping with what other analysts have termed biosafety and biosecurity standards. While biosecurity standards tend to focus on managing unauthorized access or loss of valuable biological materials, biosafety standards tend to focus on practices that can be implemented to prevent the accidental release or exposure to pathogens (Hadshar, 2023).

While our analysis focuses on private consortia that set standards, it is worth noting joint advocacy organizations like the Alliance for Regenerative Medicine (ARM); cooperative research projects, such as the International Cancer Genome Consortium and Global Biofoundries Alliance (GBA); and international multi-stakeholder bodies led by regulatory agencies, such as the International Council of Harmonization, represent significant international collaboration but are outside the scope of our study. There are also some relevant biotechnology standardization efforts going on at traditional SDOs or multilateral organizations such as the OECD, ISO, and Codex Alimentarius Commission.¹⁴

Within biotechnology, we find that different factors are responsible for driving cooperation in different sub-areas, primarily shared reputation and interoperability.

International Gene Synthesis Consortium (IGSC)¹⁵

Shared reputation and safety concerns are the primary drivers behind the International Gene Synthesis Consortium (IGSC). This California-based 501(c)(3) organization was established by US-based nucleic acid (DNA and RNA) synthesis companies and has grown to approximately 40

¹⁴ It is also worth noting that there has been interstate cooperation on agricultural biotechnology, including commitments by China to accept US agricultural biotechnology products as part of the 2020 Phase 1 Trade Deal (Trump White House, 2020).

¹⁵ This section draws upon an interview with a member of IGSC leadership.

members globally, including Chinese companies Tsingke, Synbiotech, and BGI.¹⁶ The consortium emerged when industry players recognized that misuse of synthetic DNA technology posed collective reputational risks that would damage all companies in the sector, regardless of which company's products were implicated (Tucker, 2010). Companies were not primarily concerned with the weaponization of DNA sequences, which was believed to be quite challenging, but rather reputational risk from low impact incidents which could nonetheless cause significant controversy. For example, the Guardian ran a story in 2006 about how easy it was to order the DNA sequences for smallpox (Randerson, 2006). Moreover, the US government made it clear to the leaders of key DNA synthesis companies that something needed to be done, and while the government didn't have a very clear idea of what was needed to manage risks, it would do so if companies themselves did not step up to manage this risk.¹⁷

These concerns, about shared reputation and the threat of being subject to onerous government regulations, motivated competitors to collaborate on a unified nucleic acid sequence screening protocol to prevent misuse of synthetic nucleic acids, including developing an annually updated restricted pathogen database that includes restricted pathogens from national or international export control lists (IGSC, 2009). The database essentially maps between the species listed in government lists and other more detailed databases, which list the exact sequences that correspond to the different listed species. Based on this protocol, IGSC members screen orders from customers against the pathogen database, collect identification and intended use data from customers, only supply restricted nucleic acid sequences to researchers in *bona fide* laboratories, and verify independently that the described use is consistent with the activities of the purchasing organization (Regan et al., 2024). In 2010, this consortium represented approximately 80% of the nucleic acid synthesis industry by volume and still claims to represent a majority of the nucleic acid synthesis market today, having steadily incorporated new entrants (IGSC, 2024).

Several factors are responsible for the success of IGSC. First, there are relatively few companies that need to coordinate to ensure that the same standards are applied across most of the nucleic acid synthesis market. The focus of the IGSC was also quite narrow, mapping government or multilateral lists of controlled species to specific sequences. The companies realized that they each needed this mapping to effectively screen nucleic acid synthesis orders and that pooling resources to create a joint map, which governments around the world would be able to trust as a *de facto* standard, would be in their common interests. Members of the IGSC are relatively transparent; companies involved assume that everything they share at meetings could be made public and there are no non-disclosure agreements signed between companies. This openness is seen as facilitating the inclusion of Chinese companies in the process, as it reduces the perceived sensitivity of taking part in meetings: all members treat their contributions as akin to “shouting in public” in the words of one IGSC member interviewed. The IGSC was also able to obtain an EAR99 designation for its RPD from the Department of Commerce and has generally maintained frequent contact with the relevant parts of the US government to ensure that its activities are not running afoul of regulations.¹⁸

Nonetheless, the success of the IGSC is not absolute and it continues to face challenges.

¹⁶ One of the founding companies (i.e., GenScript) was US-based but had R&D operations in Nanjing at the time and was started by a scientist of Chinese origin, suggesting that Chinese scientists were involved with the IGSC from the beginning (McGregor, 2023).

¹⁷ The IGSC was not the only effort that came about due to these pressures. It was preceded by the International Association of Synthetic Biology (IASB), which was focused on a greater role for government and civil society, alongside a fully automated screening process. Two leading companies were not in favor of the IASB approach and started IGSC alongside several other players, keeping meetings closed to just members and only admitting companies that crossed a particular revenue threshold (Heron, 2023).

First, beyond nucleic acid sequences for pathogens, which are covered by the RPD, there is the open question about how to manage risks related to sequences of concern (SOCs). SOCs are currently defined by the Office of Science and Technology Policy (OSTP) as “a nucleotide sequence or its corresponding amino acid sequence that is a Best Match to a sequence of federally regulated agents (i.e., the Biological Select Agents and Toxins List (BSAT), or the Commerce Control List (CCL)), except when the sequence is also found in an unregulated organism or toxin” (OSTP, 2024). From October 2026, however, this definition will expand to also include sequences known to contribute to pathogenicity and toxicity. This effectively gives the industry until October 2026 to determine a way to assess which sequences contribute to pathogenicity or toxicity, a fundamental question that has no simple existing answer. Currently, there are some volunteer efforts within the IGSC that are attempting to put together a dataset that maps different sequences to pathogenicity and toxicity levels.

Another challenge that the IGSC has faced regards engagement with Asian companies. Given the need to straddle multiple timezones, the IGSC often has meetings that are extremely late in the Asian timezone, which makes it much more challenging for Chinese, Japanese, and Korean companies to attend meetings. These companies still utilize IGSC’s RPD but struggle to be more deeply involved.

The IGSC also faces challenges in moving from a volunteer-driven organization to one run by full-time staff. There is believed to be a relatively low rate of misuse of nucleic acid synthesis processes and the current system seems to have worked to prevent any significant incidents. As such it has been challenging for the IGSC to persuade stakeholders that more funding and resources are needed to move beyond its current volunteer-driven format.

There are also still gaps in the current approach to coordinating between companies about potentially problematic actors. Companies do not share information about suspicious customers with each other, given the small customer base of around 20,000 entities, which makes each customer extremely valuable (IGSC, 2024). IGSC does have an internal process for the reporting of suspicious orders, but it has only been used once in its entire history.

Finally, in recent years, the securitization of some aspects of biotechnology (e.g., genomic data) has introduced challenges to international cooperation writ large. The American intelligence community has become increasingly concerned about Chinese attempts to collect American genomic data, with Congress passing the Biosecure Act in 2024, which restricts federal contracts for companies such as BGI, as well as companies that do business with BGI (National Counterintelligence and Security Center, 2021; Freifeld, 2024). Concerns about genetic data and biotechnologies are also rife in China, where in 2020, Frank Zhang, founder of GenScript, was placed under residential surveillance and later arrested by Chinese authorities for alleged smuggling of prohibited goods, highlighting the intensified scrutiny on genetic research activities (Ng, 2020).¹⁹ From speaking to individuals involved with the IGSC, it appears that securitization has yet to significantly impact the IGSC’s ability to

¹⁸ An EAR99 designation under the U.S. Export Administration Regulations (EAR) indicates that an item is subject to the EAR but is not specifically listed on the Commerce Control List (CCL). Such items are typically low-technology consumer goods and generally do not require an export license for most destinations. While an EAR99 designation for the restricted pathogen database suggests that it is not specifically controlled under the CCL and may not require an export license under typical circumstances, sharing this database with Chinese entities necessitates thorough due diligence. Exporters must ensure that the recipients are not on any restricted lists, that the intended use is permissible, and that no other regulatory prohibitions apply.

¹⁹ From July 2023, the Human Genetic Resources (HGR) Regulation has also come into effect, imposing stringent controls on the collection, preservation, and utilization of human genetic materials, particularly by foreign entities, to safeguard national biosecurity (Yi Wu, 2023).

involve Chinese companies amongst their membership. Nonetheless there have been efforts to house some data within more international bodies. For example, the International Biosecurity and Biosafety Initiative for Science (IBBIS) also houses some of the data maps that the IGSC has put together and released an open-source, globally available tool that sets the baseline for synthesis screening (IBBIS, n.d.).²⁰²¹ Housing this data at IBBIS increases the chances that it remains accessible to the global community, including China, even as tensions between the US and China escalate.

Clinical Data Interchange Standards Consortium (CDISC)

Interoperability needs, alongside a desire to pool niche knowledge in specific industries, serve as a key driver across multiple biotechnology standards organizations, most notably in clinical trial data.

The Clinical Data Interchange Standards Consortium (CDISC) is a public-private consortium that began as an informal academic working group in the late 1990s. Today, it brings together academic research institutions, medical device companies, clinical research organizations, pharmaceutical companies and regulators from the US, EU, Japan, and China to standardize consistent clinical data formats across regulatory jurisdictions (CDISC, 2024; Edwards, 2024). Domestic regulators across China, Japan, and the US require or prefer CDISC standards for clinical data submission (CDISC, n.d.).

Telecommunications and the Internet

Confidence level High

Compared to the other critical and emerging technologies included in our study, standardization efforts for technologies grouped under Data Privacy, Data Security, and Cybersecurity have been much more widely studied and reported on.²² Our summary here aims to mainly distill the key trends analyzed in existing literature and highlight how it relates to our analysis.

The broad trend across different industry consortia appears to be significant cooperation between Chinese and Western companies and academics, driven mostly by a desire for interoperability of systems.

There has been some recent politicization of the standards setting process as US-China tensions have escalated.²³ However, it is worth noting that politicization has not been equal across different technologies in this grouping. While there has been significant controversy around alleged Chinese influence on standards setting processes at the 3GPP (telecommunications) and Huawei-led initiatives proposing a top-down new architecture for the internet, Chinese involvement is not a

20 IBBIS was started by the Nuclear Threat Initiative in 2019, and is developing a tool that all providers can use for screening called the international Common Mechanism for DNA synthesis screening.

21 There are also other efforts that provide freely available DNA synthesis screening tools, notably SecureDNA hosted by the SecureDNA Foundation in Switzerland.

22 It is also worth noting that compared to many of the other technologies, this grouping is far more broad and encompasses many unrelated technologies as a sort of “catch-all” for many digital technologies.

23 One example of these tensions is the pausing of the US-China Internet Industry Forum. The forum, which appears to have been co-organized by Microsoft, convened senior political and business leaders to discuss issues related to internet governance from 2007 to 2015 (Hormats, 2011; Hormats, 2013; ISC, 2015). The last fora was held in Seattle and attended by President Xi Jinping. There may be ongoing private versions of the event, but there are no public records after 2015.

source of controversy in several key consortia that govern internet standards, including the Internet Engineering Task Force, Internet Corporation for Assigned Names and Numbers, and the World Wide Web Consortium (Voo & Creemers, 2021).

Telecommunications (3GPP and O-RAN Alliance)

The 3rd Generation Partnership Project (3GPP) case highlights that interoperability-driven standards setting can also be characterized by intensive competition between and within industry consortia. 3GPP was formed in the late 1990s as telecommunications providers around the world looked to standardize protocols to enable global communication (Baron et al., 2018). Housed at the European Telecommunications Standards Institute, it was formed by domestic standards setting organizations from Japan, North America, China, Europe, and Korea.²⁴ Early on in its history, there were disputes between 3GPP and 3GPP2, a rival group which focused on developing a rival standard that was more widely used in the US. This led to a public 3G standards war between US-based Qualcomm and Europe-based Ericsson. More recently, several different sources that examine 3GPP in great detail point to China's evolution from a standards-taker, with relatively few standards-essential patents filed in the 3G and 4G era, to a standards-maker in the 5G era. These experts argue that Chinese companies, particularly Huawei, have begun to dominate this standards setting body, being granted many more standard-essential patents (SEPs) (Triolo, 2020; Gonzales et al., 2022; Russel et al., 2021).²⁵ However, more recent research, which accounts for how integral the patents are to the technology using a Patent Asset Index, suggests that Qualcomm is still ahead of Huawei (Morris, 2025).

The O-RAN Alliance follows a similar pattern of politicization to 3GPP. It was motivated by a desire to coordinate on technology standards that would allow companies to interoperably use hardware and software from different vendors—both direct rivals and firms across the supply chain came together to set up O-RAN as a private nonprofit in Germany.²⁶ The O-RAN Alliance is responsible for publishing RAN specs and software while supporting member implementation. O-RAN standards complement 3GPP without contradicting them to prevent market fragmentation (Summer, 2020). However, the technology that the O-RAN alliance promotes, the Open Radio Access Network (ORAN), is itself politicized. The US government has actively supported O-RAN development to diversify the telecommunications supply chain and mitigate security concerns associated with Chinese vendors, ZTE, and Huawei. In 2021, Congress passed the Open RAN Outreach Act, which offers outreach and technical support to small providers regarding O-RAN, with further expansions passed in 2023 (Yoon, 2024). These efforts are complicated by the fact that China Mobile holds veto power as a founding member and other Chinese firms are also actively involved in ORAN (Harjani, 2024).

²⁴ In 2014, a newly formed body called the Telecommunications Standards Development Society of India became the seventh member.

²⁵ A Standard Essential Patent (SEP) is a patent covering technology indispensable for complying with a specific technical standard. Companies holding SEPs are obligated to license them on Fair, Reasonable, and Non-Discriminatory (FRAND) terms, which balances their interests with industry-wide accessibility. SEPs ensure a market for your technology but also put a ceiling on the price that a company can charge. Some companies prefer not to have their patents declared during the standards setting process, leading to what is known as a “patent ambush,” where they subsequently assert that a patent is infringed by use of the standard as adopted.

²⁶ Japanese companies have also been deeply involved in the O-RAN Alliance and O-RAN technologies, driven by a similar desire to avoid vendor lock-in through hardware and software interoperability across a range of different suppliers (Nanni, 2023).

Internet governance (IETF, W3C, ICANN)

Compared to telecommunications, there is significantly less politicization of internet technologies, especially with relation to concerns around undue Chinese influence. Some of this may be due to the fact that Chinese actors are not represented in leadership positions for key organizations such as the IETF and ICANN (Voo & Creemers, 2021). However, this does not fully explain the relative lack of politicization, as Beihang University, an entity on the Bureau of Industry and Security's entity list, is one of the hosts of the World Wide Web Consortium, with relatively little public controversy. This section looks at patterns of cooperation, almost entirely driven by a desire for interoperability, across three key consortia.²⁷ Moreover, despite not having senior leadership positions in the IETF, Chinese companies, particularly Huawei, are very active and have working group leadership positions (Arkko, n.d.).

The Internet Engineering Task Force (IETF) is a unique consortium. Without formal membership requirements, IETF allows anyone interested to participate in its work developing internet standards. The IETF evolved from the early ARPANET technical community in the 1980s, formally established in 1986 with an open structure that emphasized technical merit and "rough consensus" rather than formal voting procedures (IETF, n.d.). It was initially supported by the US federal government before being housed in the Internet Society in 1992, a nonprofit with local chapters across the world (ISOC, 1995). While initially dominated by US-based engineers and academics, the IETF has gradually internationalized over decades, with significant growth in Asian participation (particularly from China, Japan, and South Korea) since the early 2000s. Data shows that while Chinese participation in the IETF has grown significantly over time, the United States still maintains a dominant presence. According to the Atlantic Council's GeoTech Center research, "56 percent of the IETF working group chairs are American" compared to Chinese actors who hold only about 3% of these leadership positions (Neaher et al., 2021). However, document authorship patterns reveal interesting trends: "the number of contributions submitted by Chinese authors has increased continuously since 2000, while the number of contributions from US authors has decreased" (Teleanu, 2021). It is clear that interoperability has driven much of this desire for standardization.

Within the context of the IETF, Nanni (2022) suggests that major Chinese companies like Huawei recognize that "it would be economically disadvantageous... to create an architectural 'island' separated from the existing IP-based Internet." Huawei's patterns of engagement across IETF working groups demonstrate connections with competitor affiliates, and IETF standards often feature co-authorship between Chinese and US companies. Commercial competition exists alongside this cooperation, even among Chinese companies, with Huawei and ZTE displaying a "love-and-hate relationship."²⁸ This commercial competition creates incentives for engagement that sometimes diverge from national interests (Nanni, 2022).

The World Wide Web Consortium (W3C) is another example of interoperability-driven standardization that has proceeded with fairly little politicization or controversy. W3C, which focuses on standards for the World Wide Web that ensure long-term growth and interoperability, began at MIT's Laboratory for Computer Science in 1994, established by Sir Tim Berners-Lee.²⁹ Under a

²⁷ Other organizations may also be relevant but we did not have capacity to explore them. These include the Cloud Security Alliance, FIDO International, Internet Governance Forum, and the Wi-Fi Alliance.

²⁸ This reflects existing market realities. Huawei had internationalized its activities earlier, establishing itself in Europe in 2000 and achieving a 39% global market share in LTE contracts by 2013, compared to ZTE's 2% (Pawlicki, 2017, as cited in Nanni, 2022).

"hosted model," W3C was initially led by professors from MIT, expanding to additional hosts, including China's Beihang University in 2013, along with various regional offices worldwide. W3C transitioned to become an independent public-interest nonprofit organization in 2023 (Gordon, 2023). Several Chinese companies are members of W3C, including Baidu, Huawei, and Tencent, and membership is also open to academics and individuals (Ong, 2013).

Finally, the Internet Corporation for Assigned Names and Numbers (ICANN), which oversees the global Domain Name System (DNS) and distribution of IP and AS addresses, is another example of cooperation between government and industry for interoperability, with uncontroversial involvement of US and Chinese representatives. Established in 1998 as a California-based nonprofit, its initial oversight by the US Department of Commerce led to international criticism after revelations of US surveillance by Edward Snowden (Li Yan, 2015).³⁰ This culminated in the 2016 transition of the Internet Assigned Numbers Authority (IANA) functions from US government oversight to the global internet community (ICANN, n.d.). ICANN maintains its authority by controlling DNS root access, granting top-level domain names (e.g., .org, .uk) to registries only through binding contracts that enforce specific policies and require fee payments. Similarly, registrars can only receive ICANN accreditation by signing standardized agreements that mandate certain practices and financial obligations to ICANN. The policy parameters of what the contractual agreements between ICANN and the registries are set by a multistakeholder, consensus-driven model, with governments participating through a Governmental Advisory Committee (Mueller, 2023).

Space Technologies and Systems

Confidence level Medium

Space technologies and systems is a broad domain made up of a variety of subtechnologies that enable activities conducted in outer space. These technologies and systems contribute to spacecraft, satellites, launch vehicles, spacecraft operations, ground-based systems for data reception and processing, and other supporting infrastructure (OECD, 2022).

Standards for space technologies and systems are typically generated voluntarily and via consensus building within SDOs, like the ISO Technical Committee 20 (TC 20), the UN's International Telecommunication Union (ITU), Institute of Electrical and Electronics Engineers (IEEE), and ASTM International. A number of national space organizations conduct standards setting work, including NASA (US), CNSA (China), and Roscosmos (Russia), all of which establish some of their own technical and regulatory standards for space technologies and systems. Additionally there are regional bodies like the European Cooperation for Space Standardization (ECSS), which functions similarly to national organizations but establishes standards for ESA.

Interoperability is a major driver of cooperation in standards setting within the space domain, and the knock-on effects of interoperability are substantial for commercial incentives and, critically, safety.

²⁹ Berners-Lee had previously invented the Web at CERN and created the foundational technologies of HTML, HTTP, and URLs.

³⁰ In October 2013, a coalition of major internet governance organizations, including ICANN, the Internet Engineering Task Force (IETF), the Internet Society (ISO), the World Wide Web Consortium (W3C), and five regional Internet address registration authorities collectively released what became known as the "Montevideo statement." This declaration advocated for the "globalization" of both ICANN and the Internet Assigned Numbers Authority (IANA) functions.

A subgroup of space technologies and systems with notable cooperation for standards setting is communications and data systems, as space missions rely on global infrastructure requiring common data formats and communications protocols (NASA, 2025). The Consultative Committee for Space Data Systems (CCSDS) started in 1982 with a focus on the lower layers of the communications protocol stack. However, over time its scope has expanded to include the entire ISO communications stack and beyond to system architecture, security, and more. CCSDS is composed of 11 member agencies, 32 observer agencies, and 140+ industrial associates,³¹ and now functions as an ISO committee (CCSDS, 2010). Industrial associates participate primarily for the commercial incentives of interoperability, which expands market access and enables seamless global integration.

Even if a mission has no international interoperability goals, or no explicit plans for cooperation, there have been instances in which compliance with established standards from CCSDS has allowed for vital rescue operations. There are two frequently cited examples of this playing out: the 1995 NASA DSN rescue of the United Kingdom's STRV vehicle and the 2008 NASA DSN rescue of ESA's XMM-Newton mission (Saunders & Dumm, 2017).

It should be noted that there are a few factors which contribute to the uniqueness of the space domain and therefore limit its applicability to other industries. The primary consideration is the legal liability regime between countries and private entities in space. Under international space law, countries bear responsibility for the actions of private actors and companies operating in space (UNOOSA, 1967). The Outer Space Treaty (OST) of 1967, specifically Article VI, establishes that national governments must authorize and continuously supervise non-governmental space activities, ensuring compliance with international obligations; the Liability Convention of 1972 expands upon that responsibility. As a result, commercial inter-industry consortia for standards setting are unlikely to emerge without government involvement.

A second unique consideration within the space domain is that since 2011, the US has, codified within law through the “Wolf Amendment,” limitations on direct bilateral cooperation with China and/or China-affiliated organizations by NASA, except in instances with Congressional approval and approval by the Federal Bureau of Investigation (Mervis, 2011). As standards setting bodies are neither direct nor bilateral, participation by both the US and China in organizations like CCSDS is not directly prohibited. However, de facto there exist numerous organizations entrusted with space standards where NASA is a part and CNSA is not, such as the International Planetary Data Alliance (IPDA),³² which conducts projects related to building standards and interoperable systems. The Space Systems MOSA Interface Standards Alliance (MOSA Alliance)³³ similarly has strong participation from NASA, many US based-companies, and allied countries like the UK and Europe, but no participation from China or Chinese-affiliated companies.

Cloud and Other Advanced Computing

Confidence level Low

³¹ Industrial Associates are private sector organizations, such as aerospace companies and technology providers, that collaborate with CCSDS member agencies to contribute expertise, technology, and industry perspectives. They do not have voting rights but help shape space system standards. Both the US and China have sponsored industrial associates, but CNSA currently only sponsors 1 (CCSDS, 2025).

³² List of partners: <https://ipda.jpl.nasa.gov/partners/>

³³ List of members: <https://interoperabilityalliance.org/members/>

In advanced computing, significant cooperation seems to be driven by a desire for companies to cooperate to avoid lock-in to specific vendors. Downstream of this motivation is a desire to ensure that products developed by different parts of the supply chain are interoperable. Finally, for at least one consortium (i.e., the Confidential Computing Consortium (CCC)), the key driver appears to be knowledge pooling.

The Open Compute Project (OCP), is a nonprofit based in the US, which was launched by Facebook (now Meta) in 2011. Meta was frustrated by closed hardware systems where companies had to accept vendors' terms, so Meta engineers redesigned data center components and open-sourced these designs through OCP to gain leverage against traditional vendors while distributing development costs across the industry (Roenigk, 2021). OCP now includes leading companies like ARM, Google, Intel, Microsoft, NVIDIA, Alicloud, Bytedance, Tencent, and Baidu (OCP, n.d.). A survey by research firm Omdia showed that adoption of open standards-based servers grew from just 7% in 2016 to 36% in 2021, with projections to exceed 40% by 2025 (Chang, 2022). The pattern of Chinese company involvement is unsurprising, as IEIT Systems, a provider of general-purpose servers and racks, contributes the most to OCP standards, with Lenovo a close second. There appears to have been a drop off in Chinese company involvement in 2024, with Lenovo and IEIT being the only Chinese companies involved in OCP standards that year. Despite Chinese company involvement, membership of the steering committee and board are predominantly from the US and Europe (OCP, n.d.).³⁴

Similarly, companies such as Alibaba, Meta, Apple, Google, and Intel have come together to form the UALink Consortium, which governs the open UALink standard for communication between AI accelerators (UALink, n.d.). This standard is a direct rival to NVIDIA's proprietary NVLink, a solution for GPU-to-GPU communication. NVIDIA is not part of this consortium.

On the other hand, according to individuals involved in the CCC, the main benefit is pooling knowledge and discussing complex, niche issues with counterparts in other companies (Intel, 2024). The CCC, established in 2019 under the Linux foundation, “brings together hardware vendors, cloud providers, and software developers to accelerate the adoption of Trusted Execution Environment (TEE) technologies and standards.” Its working groups meet every 1-2 weeks, and its founding general members include SUSE, Baidu, ByteDance, Decentriq, Fortanix, Kindite, Oasis Labs, Swisscom, Tencent, and VMware (CCC, n.d.).

Semiconductors and Microelectronics

Confidence level

Medium

Key private consortia within the semiconductor industry are JEDEC, which focuses on memory standards; SEMI, a broader industry consortium; and the World Semiconductor Council (WSC), which brings together key industry associations from around the world to coordinate on key issues without formally issuing any standards itself. American and Chinese entities are active across all three organizations. This cooperation takes place against a backdrop of competition.³⁵

³⁴ Analysis done based on publicly available data on the OCP website.

³⁵ China's TC599 committee was created to help develop a domestic standard in order to “take advantage of China's whole industrial chain, break the blockade [of foreign standards], and have a voice in the international arena” (Global Times, 2023).

Knowledge pooling and interoperability are the primary drivers of standardization in the semiconductor industry, though shared reputation and identity may play a secondary role. Pooled contributions from enterprises and researchers across the world account for the advances in performance, lowering power consumption, and shrinking size while interoperability is necessary for ensuring that components from different manufacturers function together within one of the most complex global value chains (Ezell, 2020; Adams, 2023). On some issues, the semiconductor industry appears to have a globally unified voice, which may stem from a sense of joint reputation or identity, as industry associations from key markets issue joint recommendations to their governments through the WSC (WSC, 2024).

The WSC is an international forum that brings together leading semiconductor industry associations from the US, China, Taiwan, Europe, Japan, and Korea to address global challenges and promote the healthy growth of the semiconductor industry (WSC, n.d.). Established in 1996, the WSC meets annually at a CEO level, issues a Joint Statement that includes a set of recommendations to the governments and authorities, and is followed by a meeting known as the “Government/Authorities Meeting on Semiconductors (GAMS)” where government officials discuss WSC’s recommendations (WSC, n.d.).³⁶ It has four key workstreams, covering environment, safety and health, intellectual property, market openness, and market trends. While it does not issue standards, the WSC does have the power to convey unified industry recommendations to governments.

Established in 1944, the Joint Electron Tube Engineering Council (JETEC) aimed to coordinate vacuum tube type numbering systems.³⁷ As the radio industry expanded, JETEC started to include standards for solid-state and semiconductor devices and was renamed JEDEC in 1958. Initially formed as part of the Electronic Industries Alliance (EIA), it became independent in 2000 when EIA dissolved (JEDEC). Historically focused on standardizing semiconductor packages and testing methods, but evolved to become crucial for memory standards (DRAM, flash memory etc). There is comparatively high involvement from Chinese firms, who make up about 20% of total members, although committee chairs and leadership are mainly US-dominated. In 2016 JEDEC also approved a Huawei proposal for a Quality and Reliability Task Group, created to assist Chinese companies with contributing to international standards development (JEDEC, 2016). JEDEC has also been engaging with TC599, a newly formed technical committee dedicated to integrated circuit standards in China (JEDEC, 2025; CESI, 2024).

SEMI was established in 1973 to help standardize wafer dimensions in silicon manufacturing. Today these include both safety and compatibility standards (SEMI, n.d.). As of 2024 it included 3000+ companies (SEMI, n.d.). Apart from standards, SEMI also helps produce market research reports for its members in the semiconductor/LED industries, conferences and trade shows for semiconductor production equipment. Their Board of Directors is quite diverse, with significant East Asian representation (~70%, though mainly from Japan rather than China), while their Board of Industry Leaders are fairly US heavy (~40% US) (SEMI, n.d.; SEMI, n.d.). SEMI China has also been forming local chapters (e.g. in compound semiconductor materials and information & control technical committees), which have led to developments of new standards.

³⁶ The WSC chairmanship and location rotates annually. The WSC is supported by a Joint Steering Committee (JSTC) and by task force and committee units composed of industry experts appointed by the member associations.

³⁷ Vacuum tubes, also known as thermionic valves, are electronic devices that control electric current flow through a vacuum between electrodes. They were essential components in early electronics, enabling amplification and rectification in radios, televisions, and early computers before being largely replaced by transistors.

Quantum

Confidence level High

There are many different relevant quantum technologies, but with respect to standards, there are three main issues: metrology, computing, and cryptography ([Cory, 2024](#)).

Early quantum standards efforts were driven by a desire for knowledge pooling and interoperability of definitions and standards in this small, niche field, while recent efforts have been characterized by securitization and a focus on aligning quantum standards with national interests, potentially hindering future prospects for cooperation. Across the board, private consortia have played a fairly limited role, with most standardization work happening at the ISO, IEC, ITU, and European Telecommunications Standardization Institute (ETSI).



Diagram of key international quantum standardization efforts
(Source: [SQuaD-Quantenkommunikation](#))

Early efforts included targeted, initial standards discussions (often called pre-standardization) to develop common terminology, definitions, use cases, and measurement and benchmarking tools to ensure stakeholders are talking about the same thing and measuring it the same way (Cory, 2024). Firms, national standards bodies, and other quantum experts discussed early-stage quantum technology standards issues at ISO/IEC JTC 1 on information technology, which set up working group 14 on quantum computing in 2020 (Cory, 2024).³⁸ ETSI, a body set up by the European Commission, has an Industry Specification Group on Quantum Key Distribution (ISG-QKD) that focuses on standardizing quantum cryptographic techniques to enhance secure communications. Since 2008,

³⁸ For example, in 2024, the ISO/IEC released a standard focused on standardizing the vocabulary used for quantum computing (ISO, 2024).

telecom and quantum companies from the US (e.g., HP, Airbus), China (e.g., Huawei, China Telecom), and other countries have collaboratively developed quantum key distribution standards (ETSI, n.d.). In 2023, the group released the world's first Protection Profile for QKD systems (ETSI, 2023).

Amid intensifying quantum technology competition with China, the Biden administration adopted a more interventionist approach by pushing for a new ISO/IEC joint technical committee on quantum standards despite the technology's immaturity, departing from traditional private sector-led US policy by appointing NIST to lead the US technical advisory group (Cory, 2024). Meanwhile, Chinese actors have taken steps in a similar direction; China-based organizations have formed an Intellectual Property Alliance for Quantum Computing Industry, and The Institute of Commercial Cryptography Standards has been set up under the Chinese Cryptography Standardization Technical Committee to solicit proposals for post-quantum cryptographic algorithms in parallel with ongoing efforts by NIST (Digital Watch Observatory, 2023; Quantum Insider, 2025).³⁹

Given that many relevant standardization discussions are taking place in established SDOs, industry consortia have focused on supporting the growth of a quantum industry, not technical standardization. Industry consortia in the quantum space have also adopted an alliance-of-alliances model, with industry associations from the US, Canada, Europe, and Japan coming together to form the International Council of Quantum Industry Associations in 2023. This consortium critically does not include any Chinese actors (QED-C, 2023). Another broader effort, the Open Quantum Institute (OQI), is a multilateral governance initiative hosted at CERN, which brings together industry, policymakers, and academics. As an educational initiative, it also aims to broaden access to quantum technology and bring together experts to explore how quantum technology can be used in the public interest. OQI also does not include any experts from China (OQI, n.d.).

Human-Machine Interfaces

Confidence level

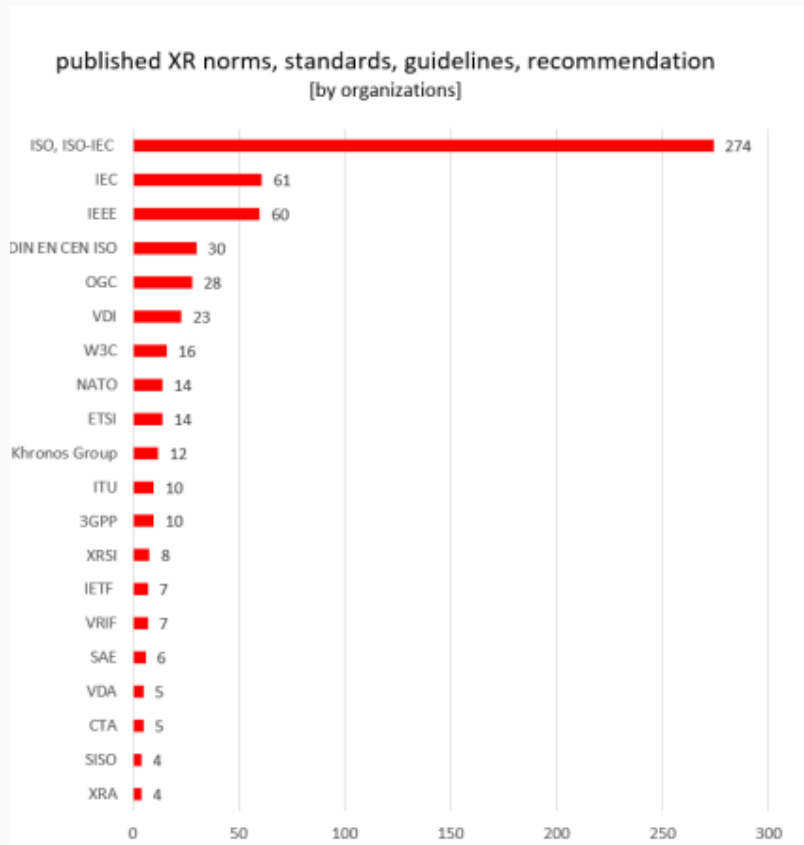
Medium

Human-machine interfaces refer to a broad set of technologies that encompass extended reality and neurotechnologies. Standards setting for both technologies appears to be mostly driven by interoperability and knowledge pooling and take place in traditional SDOs.

Extended reality (XR)

XR standardization involves 40 organizations developing hundreds of standards across seven major categories: fundamentals, management, UX/ergonomics, coding/interoperability, graphics software, hardware, and applications. Most standards are set by traditional SDOs, such as the ISO/IEC, IEEE, and ETSI, which have published over 350 XR-related documents, as well as existing technology standards bodies such as the 3GPP, the IETF, W3C, and the Khronos Group (Runde, 2023). American and Chinese companies are active across all of these key standards setting bodies, either directly or through domestic industry associations.

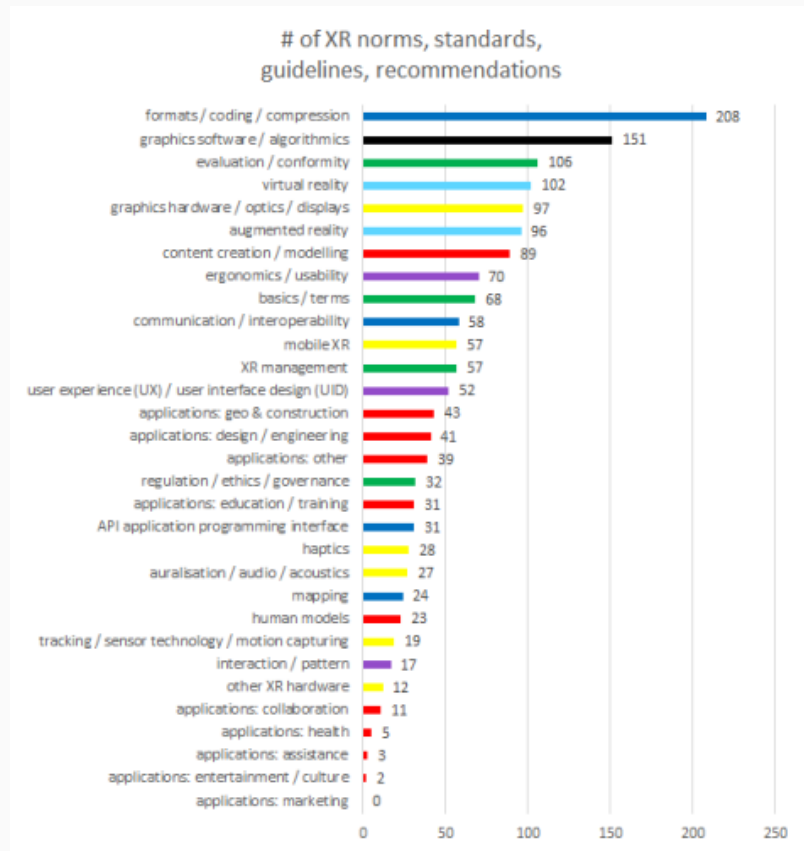
³⁹ Post-quantum cryptography is not a quantum technology but rather a classic technology that is designed with quantum enabled cryptography in mind.



(Source: *Runde, 2023*)

Given that most of the standards set across all the bodies are focused on formats, algorithms, and other technical features, the drivers of cooperation appear to be knowledge pooling and interoperability. Only 5% of all the standards focus on ethics and governance, which suggests that shared safety concerns are less important.⁴⁰ Smaller XR-focused consortia reflect both safety and interoperability concerns. The Metaverse Standards Forum (which includes companies from across the world, including China) states that “interoperability is the bedrock of the metaverse, to enable multiple disruptive technologies to work together,” while the XR Safety Initiative (which does not seem to include any Chinese experts), focuses on safety, privacy, security, human rights, human well-being, responsible innovation, governance, and regulation (*Metaverse Standards Forum, n.d.; XRSI, n.d.*).

⁴⁰ It is also understandable that most of the standards are related to the development of the technology and that safety takes up a much smaller portion of the overall standards set. Therefore, we do not take this as strong evidence that XR standards are not driven by shared safety concerns and reputational considerations.



(Source: [Runde, 2023](#))

Neurotechnologies

Neurotechnology refers to electronic devices or methods used to read or modify neural activity ([Harper, 2023](#)). Key technologies include:

- Brain machine interfaces (BMI) that translate brain activity into technical commands
- External brain interference systems that modulate neural activity
- Non-invasive "brainware" devices for monitoring neural activity

Interoperability appears to be a key driver of cooperation, but overall availability of data and information for neurotechnologies is severely limited.

Standards development for brain-computer interfaces is advancing through several traditional SDOs and domestic standards bodies. ISO/IEC JTC 1/SC 43 has been responsible for creating foundational standards and reference architectures for BCIs ([ISO, 2022](#)). The IEEE also leads multiple initiatives, including Working Groups P2731 and P2794, focused on unified terminology and reporting standards respectively, while the Consumer Technology Association focuses on interoperability standards for consumer EEG data, notably with the ANSI/CTA 2060 specification for EEG data file formats ([Chavarriaga, 2021](#); [IEEE, 2020](#)).

Non-traditional SDOs play a more limited role, though Chinese actors only appear to be involved in academic initiatives. The International Neuroinformatics Coordinating Facility (INCF), an organization formed by 12 member countries to facilitate the development of neuroinformatics by developing web portals for accessing integrated neuroscience databases and tools and creating standards for data sharing, does not include China ([Bjaalie et al., 2007](#); [INCF, n.d.](#)). Chinese academics are involved in academic initiatives, such as the Organization for Human Brain Mapping (OHBM), an international society dedicated to advancing the understanding of the anatomical and functional organization of the human brain through neuroimaging ([OHBM, n.d.](#)).

Lasers

Confidence level High

Directed energy (the term used in the CET list) refers to use of lasers, microwaves, or particle beams, typically in weapons. Directed energy weapons (DEW) use concentrated electromagnetic energy to combat enemy forces and assets. These weapons include high energy lasers and other high power electromagnetics, such as millimeter wave and high power microwave weapons. Unlike weapons that fire bullets or missiles, DEWs can respond to a threat in different ways. For example, they can temporarily degrade electronics on a drone or physically destroy it ([GAO, 2023](#)). Directed energy also has wide ranging commercial uses, including the use of lasers and jammers for infrastructure protection, battery cooling, industrial cutting and welding, and 3D-printing and additive manufacturing.

The military and commercial realms for lasers are largely separable.

There is coordination between companies across the world on commercial laser safety standards but almost entirely through the IEC. The IEC developed the 60825 series of standards to ensure laser product safety across various applications. These standards are the basis of most national standards ([Laser Focus World, 2001](#); [Lumispot, n.d.](#)). Outside of the IEC, there does appear to be coordination; we found some evidence of cooperation between Chinese and European bodies, but none directly between US and Chinese entities. For example, the Guangdong Laser Industry Association and the European Photonics Industry Consortium signed an MOU in 2015 ([Laser Focus World, 2015](#)).

In the military realm, directed energy weapons are considered a relatively early and nascent technology, given limited use in at-scale operations thus far. Most consortia are national or between like-minded allies (e.g., the [UK's Dragonfire](#)), with no cooperation between companies from rival countries.

There is some interstate coordination, as the US and China have both ratified the Protocol on Blinding Laser Weapons (Protocol IV) to the 1980 Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons, which prohibits the use of lasers intended specifically to blind enemy combatants ([ICRC, n.d.](#)).

However, most of the coordination takes place between likeminded allies against countries like China. High power lasers are a dual-use technology; they are controlled under Section 6 of the multilateral Wassenaar Arrangement on Export Controls for Conventional Arms and Dual Use Goods

and Technologies (WA). WA was formed in 1995 and established with the goal of enhancing international security by promoting responsibility and transparency over transfers of dual use goods and technologies (WA, 2023).⁴¹ The WA lays out clear specifications for controlled lasers. For example, carbon monoxide lasers are export controlled if they meet the following requirements: “having any of the following: a. Output energy exceeding 2 J per pulse and “peak power” exceeding 5 kW; or b. Average or CW output power exceeding 5 kW” (WA Control List, 2024).

Hypersonics

Confidence level Low

Referring to systems that reach speeds of Mach 5 (or five times the speed of sound) or faster, hypersonic systems that require incredibly specialized technology and materials to withstand unique aerodynamic, thermal, and propulsion challenges met by traveling at such speeds (Center for Arms Control and Non-Proliferation, 2023).

There are no dedicated international industry bodies for standards setting for hypersonics. Many existing aerospace standards (i.e., ISO, ECSS, CCSDS) influence hypersonic systems but aren’t exclusively focused on them. Of those standards produced explicitly for the governing of hypersonics, they seem to originate predominantly from the US fairly unilaterally through organizations like the American Institute of Aeronautics and Astronautics (AIAA), NASA, and the Department of Defense (NASA, 2025; Vergun, 2021).

While not standards-producing, the AIAA has hosted conferences on hypersonics in which both US and Chinese affiliated organizations and institutions have attended (AIAA, 2017). These conferences are indicative of the need felt by entities across the US and China to pool knowledge on a niche topic, but any such coordination has likely been impeded by the pre-existing arms race dynamics and competition for hypersonic weapons development between the US, China, and Russia (Ioanes, 2023).

Aerospace⁴²

Confidence level Low

The aerospace industry encompasses the design, development, production, and maintenance of aircraft, spacecraft, satellites, and related systems. It is a highly advanced and technology-driven industry that includes both commercial and military applications, ranging from passenger airlines and cargo transport to defense aviation and space exploration.

Despite certain disanalogies with the AI sector (slower pace of innovation; slower update frequency for standards), international standards setting in the aerospace industry may provide lessons for US-China AI safety cooperation. Through the International Aerospace Quality Group (IAQG), US and

⁴¹ China was not engaged in initial negotiations for the WA’s establishment, and is not a member. While interested in membership at various points, China ultimately did not meet the membership requirements due to reservations around the supplying of arms and dual use goods to countries of concern (Pan, 2018; Bates, 2010).

⁴² Information in this section is drawn from the IAQG public website.

Chinese companies cooperate on quality and safety standards for aircraft.⁴³ While the contents of the standards may not be highly sensitive from a national security perspective, participants in the process include major players in the US defense sector. Thus, IAQG may provide a model for ensuring some level of cooperation between companies working in sensitive areas.

Of particular note is the hub and spoke model of the IAQG. The organization is made up of three semi-autonomous and geographically subdivided sectors covering the Americas (AAQG), Asia-Pacific (APAQG), and Europe (EAQG). Each sector is responsible for disseminating IAQG standards within their respective region. It is plausible that this decentralized structure facilitates greater buy-in, with regional bodies able to act with greater autonomy in how they implement standards and police compliance in their own jurisdictions, but with strong shared reputation incentives across regions.

IAQG develops quality and safety standards for aerospace technologies. The organization is comprised of member companies within the aviation, space, and defense industries that design, develop, manufacture, and support original equipment at system or subsystem levels. IAQG standards are documented, published, and applied internationally by original equipment manufacturers (OEMs) and circulated throughout the supply chain. IAQG was established in 1998 under the umbrella of the Society of Automotive Engineers (SAE), headquartered in Brussels, Belgium. In 2013, it became an International Not-For-Profit Association under Belgian Law.

The organization has something of a hub and spoke structure, made up of three regional sectors covering the Americas (AAQG), Asia-Pacific (APAQG), and Europe (EAQG). All 10 of the Full Members in AAQG are from the US, and include the largest defense contractors (i.e., Raytheon, Lockheed Martin, and Northrop Grumman). Two of the seven Full Members of APAQG are from China (i.e., Commercial Aircraft Corporation of China in Shanghai and Aviation Industry Corporation of China in Beijing), as are several Affiliate and Associate members.

⁴³ Other international industry associations, including ICCAIA and IAEG, engage in standards setting and coordination activities. For the purposes of this memo, we focus solely on IAQG, but this could be expanded in a future paper.

The logo for the Singapore Association of International Finance (SAIF) is centered on the page. It features a stylized white asterisk-like symbol to the left of the acronym "SAIF" in a bold, white, sans-serif typeface. The background is a solid teal color with several thick, light-colored diagonal lines crossing the frame.

*** SAIF**