How Does FAIIA Compare to AIARA?

Comparing key provisions in the Future of Artificial Intelligence Innovation Act of 2024 with those of the AI Advancement and Reliability Act of 2024. Both bills aim to codify an entity similar to the current AI Safety Institute.

View online: futureoflife.org/f-vs-a

14th November 2024

US Policy Team
policy@futureoflife.org

Organization

Future of Life Institute

Lead Authors

Alexandra Tsalidis - alex@futureoflife.org

Landon Klein - landon@futureoflife.org

About the Organization

The Future of Life Institute (FLI) is an independent nonprofit organization with the goal of reducing large-scale risks and steering transformative technologies to benefit humanity, with a particular focus on artificial intelligence (AI). Since its founding, FLI has taken a leading role in advancing key disciplines such as AI governance, AI safety, and trustworthy and responsible AI, and is widely considered to be among the first civil society actors focused on these issues. FLI was responsible for convening the first major conference on AI safety in Puerto Rico in 2015, and for publishing the Asilomar AI principles, one of the earliest and most influential frameworks for the governance of artificial intelligence, in 2017. FLI is the UN Secretary General's designated civil society organization for recommendations on the governance of AI and has played a central role in deliberations regarding the EU AI Act's treatment of risks from AI. FLI has also worked actively within the United States on legislation and executive directives concerning AI. Members of our team have contributed extensive feedback to the development of the NIST AI Risk Management Framework, testified at Senate AI Insight Forums, briefed the House AI Task-force, participated in the UK AI Safety Summit, and connected leading experts in the policy and technical domains to policymakers across the US government.

FAIIA/AIARA Comparison Table

Key: Most effective | Somewhat effective | Least effective

Legislative citations are clickable online: futureoflife.org/f-vs-a

		H. R. 9497 | AI Advancement and Reliability Act of 2024 (Congress.gov)	**S.4178** | Future of Artificial Intelligence Innovation Act of 2024 (Cantwell-Young Substitute)	**Analysis**
Who is responsible for the establishment of the entity?		The Under Secretary of Commerce for Standards and Technology [Proposed § 5304(a)(1)]	The Director [Proposed § 22B(b)]	FAIIA states, "the Director shall establish an institute on artificial intelligence within the Institute," but it does not specify the meaning of "Director". Presumably, this would be the Director of NIST, but this should be clarified given that under Proposed § 102, which concerns the testbed program, "Director" is defined as the Director of the National Science Foundation.
Name of entity is to be established		Center for AI Advancement and Reliability [Proposed § 5304(a)(1)]	Artificial Intelligence Safety Institute [Proposed § 22B(b)(2)]	Calling the new entity the "Artificial Intelligence Safety Institute" is somewhat preferable to the name given by AIARA since it aligns with what similar entities are called in other countries. Aligning with this international convention would signal solidarity within the international research community and will communicate the collaborative intent of the entity.
Date by which the entity is to be established		N/A	90 days after enactment [Proposed § 22B(b)(1)]	Unlike AIARA's open-ended timeline, FAIIA provides a more concrete timeline, ensuring the timely establishment of the institution to address urgent issues.
Mission of the entity		To advance the measurement science for AI reliability, robustness, resilience, security, and safety. [Proposed § 5304(a)(2)(A)] To develop voluntary best practices and technical standards for evaluating the reliability, robustness, resilience, security, and safety of AI. [Proposed § 5304(a)(2)(C)] To increase the understanding of government, institutions of higher education, private sector entities, and the public of artificial intelligence reliability, robustness, resilience, security, and safety-related challenges and remediations. [Proposed § 5304(a)(2)(D)]	To assist the private sector and agencies in developing voluntary best practices for the robust assessment of AI, which may be contributed to or inform the work on such practices in standards development organizations. [Proposed § 22B(b)(3)]	AIARA provides a more expansive and safety-focused mission for its Center, explicitly emphasizing reliability, robustness, resilience, security, and safety across multiple objectives. In contrast, FAIIA's mission statement is narrower, focusing primarily on developing voluntary best practices for AI assessment. That said, both FAIIA and AIARA rely on voluntary cooperation for standards development and testing, which is likely to constrain both entities' ability to succeed in their missions.
Who will lead the entity?		The Under Secretary of Commerce for Standards and Technology or an appropriate designee. [Proposed § 5304(a)(3)]	Not specified.	The Cantwell Substitute does not specify who will lead AISI. This is an issue as it could hinder effective decision-making and coordination within AISI, especially in its first year of operation.
Activities of the entity	Evaluations and red-teaming	Conducting evaluations and benchmarking capabilities, including through red teaming. [Proposed § 5304(a)(5)(A)] Conducting measurement research to inform methodologies and best practices related to aspects like safety definitions, system robustness, testing, and red teaming. [Proposed § 5304(a)(5)(B)]	Assessing AI systems and developing best practices for reliable and secure development, deployment, and use of AI. [Proposed § 22B(c)(1)(A)] Supporting red-teaming, sharing best practices, and coordinating on building testbeds and test environments with allies and international partners of the United States. [Proposed § 22B(c)(1)(B)]	FAIIA takes a more comprehensive approach to AI evaluation by explicitly including red-teaming and international collaboration on testing environments. While AIARA mentions evaluations and benchmarking generally, FAIIA provides more specific mechanisms for assessing AI systems and developing practical guidelines for their deployment. However, the language in both FAIIA and AIARA around testing is phrased indirectly. Under FAIIA, the AISI will be "assessing" AI, but it will also be "supporting" red-teaming, "sharing" best practices, and "coordinating" on building testbeds. Similarly, AIARA's Center will be "conducting" evaluations, but it will also be "conducting measurement research" to inform methodologies and best practices related to aspects like red teaming and testing. It is important that AISI/the Center has the capacity to directly conduct testing including red-teaming, instead of simply developing standards around this. This is because AISI/the Center can provide invaluable testing resources and expertise that companies themselves are unlikely to be able to conduct in-house.
	Testing standards	Conducting measurement research for the evaluation and assurance of AI, including developing common definitions for safety across use cases, the ability of AI systems to withstand adversarial attacks, testing/evaluation methods throughout the lifecycle, and reference use cases (and appropriate criteria to evaluate AI in these). [Proposed § 5304(a)(5)(B)]	Coordinating or developing metrics and methodologies for testing artificial intelligence systems, including cataloging/testing existing metrics. [Proposed § 22B(c)(1)(F)] Support and contribute to voluntary, consensus-based technical testing standards which may address: ▪ Physical infrastructure for training and operating AI. [Proposed § 22B(e)(1)] ▪ Training data. [Proposed § 22B(e)(3)] ▪ Data for evaluations. [Proposed § 22B(e)(4)] ▪ Trained or partially trained models and any resulting software systems or products. [Proposed § 22B(e)(5)] ▪ Human-in-the-loop testing. [Proposed § 22B(e)(6)]	AIARA emphasizes measurement research and safety evaluation across the AI lifecycle with a focus on adversarial testing and use cases, but FAIIA's approach to testing standards is preferable because it is more granular. FAIIA specifies testing components like physical infrastructure, training data, and human-in-the-loop testing. That said, AIARA's phrasing is more active and direct than FAIIA's, enabling it to lead on the development of standards as opposed to simply "coordinating" or "supporting". Given the number of stakeholders who may contribute to standards development, including industry and international standard-setting bodies, the U.S. can lead on establishing standards for AI systems by putting the Center/AISI at the helm of these efforts.

Continued on next page →

FAIIA/AIARA Comparison Table *(continued)*

Key: Most effective | Somewhat effective | Least effective

Legislative citations are clickable online: futureoflife.org/f-vs-a

		H. R. 9497 | AI Advancement and Reliability Act of 2024 (Congress.gov)	**S.4178** | Future of Artificial Intelligence Innovation Act of 2024 (Cantwell-Young Substitute)	**Analysis**
Activities of the entity	International collaboration	Supporting the US government in collaborating with international standards organizations, multilateral organizations, and topically relevant bodies and organizations among allies and partners. [Proposed § 5304(a)(5)(D)]	Coordinating with counterpart international institutions, partners, and allies, to support global interoperability in the development of research and testing of standards. [Proposed § 22B(c)(1)(G)] The Under Secretary shall seek to form alliances or coalitions with like-minded governments to, among other objectives, develop the government-to-government infrastructure. [Proposed § 111(b)] The criteria to form an alliance or coalition with another country will include: ▪ Having a high level of scientific and technological advancement. [Proposed § 111(c)(1)] ▪ Supporting the principles for international standards development set out in Committee Decision on the WTO Agreement on Technical Barriers to Trade. [Proposed § 111(c)(2)] ▪ Having in place sufficient IP protections, safety standards, and risk management approaches. [Proposed § 111(e)(1)] ▪ Developing and coordinating research security measures, export controls, and IP protections relevant to innovation, development, and standard-setting. [Proposed § 111(e)(2)] The People's Republic of China is not permitted to participate in an alliance or coalition until the U.S. Trade Representative reports that they have come into compliance with WTO commitments. There must then be a report on the national security, human rights, and commitment monitoring implications of such collaboration. [Budd 2 Amendment]	While the meaning of "allies and partners" is ambiguous, AIARA's more inclusive approach to international collaboration, which leaves room for cautious engagement with significant global AI actors, better serves collective AI safety objectives than FAIIA's explicitly exclusionary policy. Ideally, AIARA's language would be expanded to also include collaboration with strategic competitors, as this type of exchange could be most valuable to the Center/AISI. While FAIIA's restrictions on Chinese participation reflect legitimate security concerns, they could impede crucial technical cooperation on AI safety standards at a time when coordination with all major AI-developing states is essential for establishing effective global safety measures. Since U.S. companies are currently the primary publishers of open-source AI models, balanced international collaboration could transform what is now a one-sided sharing of information into a mutually beneficial exchange. A better approach would balance precaution with collaboration in a flexible manner which enables the Director (or equivalent) of the Center/AISI to ascertain when and to what extent collaboration would be both safe and mutually advantageous. As such, the Director could be given discretion as to technical dialogues with China on critical safety issues while still protecting U.S. national security interests and maintaining appropriate safeguards.
	Helping the government reduce AI risks	Providing NIST and Federal agencies with support on AI risk mitigation. [Proposed § 5304(a)(5)(C)] As appropriate, coordinating Federal research related to AI safety. [Proposed § 5304(a)(5)(E)] The Under Secretary must also assess how AI could create economic or national security risks. [Proposed § 5304(a)(6)(B)]	N/A	AIARA's explicit provisions for supporting government agencies with AI risk mitigation fill a crucial gap that FAIIA does not address. These provisions ensure that NIST can directly assist federal agencies in addressing AI risks, while the requirement for security assessments adds an important layer of proactive risk monitoring that is absent from FAIIA's approach.
	Methods or tools to defend against attacks	N/A	Developing and publishing physical and cybersecurity tools, methodologies, best practices, voluntary guidelines, etc. to assist persons who maintain systems used to create or train AI models with discovering and mitigating vulnerabilities and attacks, including manipulation through data poisoning, including those that may be exploited by foreign adversaries. [Proposed § 22B(c)(1)(C)] Establishing blue-teaming capabilities to support mitigation approaches and partnering with industry to address AI reliability. [Proposed § 22B(c)(1)(D)]	The development of specific tools and methodologies to protect AI training systems from manipulation is especially crucial as data poisoning becomes a more significant threat vector, while the establishment of blue-team testing creates practical mechanisms for identifying and fixing security weaknesses before they can be exploited. Blue-teaming is different from red-teaming in that it focuses on assessing and strengthening the security of an AI system by identifying vulnerabilities and providing mitigation techniques, while red-teaming simulates adversarial attacks to test the system's defenses and expose weaknesses. These concrete security provisions are vital for ensuring AI model integrity throughout the development pipeline. Furthermore, publishing these tools, methodologies, and best practices allows smaller businesses to implement robust AI security measures cost-effectively, enabling them to safeguard their systems against threats without needing extensive resources.
	Synthetic content	N/A	Developing tools, methodologies, best practices, and voluntary guidelines for detecting synthetic content, content authentication, provenance tracking, and labeling. [Proposed § 22B(c)(1)(E)]	The risks posed by deepfakes range from undermining individual autonomy and facilitating fraud to manipulating democratic processes. This is a pressing issue for which further research is needed in the areas of detection, watermarking, and labeling. These are areas where the Center/AISI could contribute valuable research.
Reporting requirements	Who shall make the report?	The Under Secretary of Commerce for Standards and Technology [Proposed § 5304(a)(7)]	The Under Secretary of Commerce for Standards and Technology [Proposed § 105(a)]	
	To whom shall the report be made?	The House Committee on Science, Space, and Technology and the Senate Committee on Commerce, Science, and Transportation [Proposed § 5304(a)(7)]	Congress [Proposed § 105(a)]	Requiring that reports are made to specific committees of jurisdiction make these provisions more actionable as it ensures that progress reports are seen by the most relevant parties.
	When shall the report be made?	Beginning in 2026, 90 days after the President's annual budget request. [Proposed § 5304(a)(7)]	No later than 1 year after enactment. [Proposed § 105(a)]	
	What shall be reported?	A summarized budget for the Center (this fiscal year and the previous). [Proposed § 5304(a)(7)(A)] Goals, priorities, and metrics for guiding and evaluating the Center's activities. [Proposed § 5304(a)(7)(B)]	Progress on the implementation of the testbed program. [Proposed § 105(a)] Reimbursable expenses, project schedules, and deliverables for the testbed program. [Proposed § 105(b)]	The AIARA reporting provisions are stronger than those in FAIIA because AIARA extends reporting requirements to the activities of the entire Center, while FAIIA only requires reporting on the testbed program. Furthermore, AIARA requires a comprehensive budget summary for the entire Center, covering both the current and previous fiscal years. This holistic approach ensures transparent financial oversight of the Center's operations, whereas FAIIA's reporting is limited to reimbursable expenses for the testbed program, lacking an equivalent provision for the broader Institute's financial needs. At the same time, it is valuable for Congress to be consistently updated on the progress of FAIIA's testbed program, as this incentivizes prompt implementation.

Continued on next page →

FAIIA/AIARA Comparison Table *(continued)*

Key: Most effective | Somewhat effective | Least effective

Legislative citations are clickable online: futureoflife.org/f-vs-a

		H. R. 9497 | AI Advancement and Reliability Act of 2024 (Congress.gov)	**S.4178** | Future of Artificial Intelligence Innovation Act of 2024 (Cantwell-Young Substitute)	**Analysis**
Establishment of a consortium	Who shall be part of the consortium?	Stakeholders from academic or research communities, Federal laboratories, private industry, and civil society. [Proposed § 5304(b)(1)]	Stakeholders from academic or research communities, Federal laboratories, private industry, including developers, deployers, and users, and civil society. [Proposed § 22B(d)(1)(A)]	
	Reporting requirements	The Under Secretary shall submit a report on the contributions of Consortium members (no later than two years after enactment). [Proposed § 5304(b)(3)]	The Director shall submit an annual report summarizing the contributions of the members of the consortium. [Proposed § 22B(d)(3)]	FAIIA's Consortium reporting requirements are preferable as they will provide regular updates as opposed to a single report within two years of enactment.
	Purpose of the consortium	Support the Center in its activities. [Proposed § 5304(b)(1)(A)] Evaluate the needs of stakeholders. [Proposed § 5304(b)(2)(A)] To identify and make recommendations on gaps remaining the Center's activities. [Proposed § 5304(b)(2)(B)]	Supporting the Institute in carrying out its functions. [Proposed § 22B(d)(1)(A)] Consulting with the Director not less frequently than quarterly. [Proposed § 22B(d)(2)]	The functions of the consortium envisioned in AIARA are preferred because they focus on evaluating stakeholder needs and recommending ways to address gaps in the Center's activities.
Establishment of a testbed program	Who shall be in charge of establishing the program?	N/A	The Under Secretary of Commerce for Standards and Technology [Proposed § 102(b)] The Secretary of Energy [Proposed § 102(b)] The Director of the National Science Foundation [Proposed § 102(b)]	Given that three entities are listed in FAIIA's provisions, it would be preferable for there to be a single point of contact or responsible entity, such as the Under Secretary. The Under Secretary can then coordinate with the Director and the Secretary in order to ensure effective implementation of the testbed program.
	Purpose	N/A	To encourage collaboration and support partnerships between the National Laboratories, Federal laboratories, the NIST, NAIRR (or any successor program), and public and private sector entities. [Proposed § 102(b)] To conduct tests, evaluations, and security or vulnerability risk assessments, and to support research and development, of AI systems, including measurement methodologies developed by the Institute, in order to develop standards and encourage development of a third-party ecosystem. [Proposed § 102(b)]	The FAIIA testbed program is an important provision which would be key for fostering collaboration between federal institutions, private entities, and public sector bodies. The program would help identify AI capabilities, limitations, and security vulnerabilities, which is crucial for developing safer, more reliable AI technologies. Its focus on building a third-party ecosystem further encourages industry-wide trust and innovation. However, the voluntary nature of the test program for vendors may limit participation.
	Functions	N/A	Run tests and evaluations on the capabilities and limitations of AI. [Proposed § 102(c)(1)] Develop automated and reproducible tests and evaluations for AI. [Proposed § 102(c)(3)] Assess the computational resources necessary to run tests and evaluations and research how these can be minimized. [Proposed § 102(c)(4)] Develop tests and evaluations that are high-, medium-, and low-computational intensity.[Proposed § 102(c)(6)] Identify security vulnerabilities such as: ▪ autonomous offensive cyber capabilities [Proposed § 102(c)(7)(A)] ▪ cybersecurity vulnerabilities in the software ecosystem and beyond. [Proposed § 102(c)(7)(B)] ▪ chemical, biological, radiological, nuclear, critical infrastructure, and energy-security threats or hazards. [Proposed § 102(c)(7)(C)] Take into consideration the applicability of any tests, evaluations, and risk assessments that could enhance a system's ability to contribute to the creation of a pandemic or biological weapon. [Proposed § 102(c)(7)(d)] Carry out a voluntary test program for vendors of foundation models. [Proposed § 102(i)]	It is valuable for the testbed program to have the capacity to run its own tests and evaluations on powerful AI systems. As mentioned above, this program can contribute significant resources and expertise to conduct certain tests and risk assessments that companies are unlikely to conduct themselves. As an impartial entity, FAIIAs testbed program would also provide independent, objective, and scientifically grounded testing, without the conflicts of interests that could be created by conducting these internally within companies. FAIIA also identifies a non-exhaustive list of pressing security vulnerabilities for the testbed program to focus on. Giving the Under Secretary and Secretary the capacity to expand this list as new security risks emerge also makes these provisions future-proof.
Staffing		Can appoint up to 15 staff members until 2029. [AIARA § 2(c)]	Can appoint 30 staff members until 2035. [Proposed § 302(a)]	The goals and missions of both entities are ambitious. They therefore require sufficient resources to realize these ambitions. FAIIA allows for the appointment of up to double the number allowed by AIARA, and extends this provision until 2035, offering greater capacity and a longer timeframe to build expertise. This broader staffing allocation enables the Institute to handle a wider range of projects and responsibilities, ensuring it can support its long-term initiatives more effectively than AIARA's more limited, shorter-term staffing provision. In addition, giving the Director of the Center/AISI the discretion to appoint additional staff members as needed (and when properly justified by increased resource needs) would give these bodies the necessary flexibility to function effectively, especially in the first years of operation. The market for AI is projected to rise to $826 billion by 2030 which is nearly eight times what it is today. Compared to the size and extent of the research that will be happening by 2035, even FAIIA's provision of 30 staff members seems extremely restrictive for effective safeguarding of this technology.
Sunset clauses		The section terminates 6 years after it is enacted. [AIARA § 2(g)]	The hiring authority expires in 2035. [Proposed § 302(b)] The testbed program sections shall end 7 years after enactment. [Proposed § 102(l)]	AIARA's sunset clause is more restrictive than FAIIA's which only apply to certain parts of the bill's programs. Furthermore, it should be noted that the staffing authority under AIARA expired before the sunset clause, which could lead to unexpected staffing shortages if the Center's programs are to continue past the expiry of the staffing provisions.

Continued on next page →

FAIIA/AIARA Comparison Table *(continued)*

Key: Most effective | Somewhat effective | Least effective

Legislative citations are clickable online: futureoflife.org/f-vs-a	**H. R. 9497** | AI Advancement and Reliability Act of 2024 (Congress.gov)	**S.4178** | Future of Artificial Intelligence Innovation Act of 2024 (Cantwell-Young Substitute)	**Analysis**
Appropriations funding	$10,000,000 for fiscal year 2025. [AIARA § 2(h)]	N/A	The budget earmarked for AIARA in its first year is limited relative to the mission and function envisioned. Designating this precise amount provides some security that the Center will be funded in its first year of operation, but even this budget seems insufficient for efficient and effective implementation given the critical role the resulting entity will play.
Confidentiality	N/A	Any confidential content provided by a private sector person shall be exempt from public disclosure rules. [Proposed § 22B(f)(1)] Access to a private sector person's confidential content shall be limited to the private sector person and the Institute, but deidentified data may be made available. [Proposed § 22B(f)(2)]	Given the voluntary nature of model access and information disclosure, the FAIIA public disclosure exemption seems appropriate. However, it is important for information that is in the interest of national security to be disclosed by AISI to relevant agencies, such as the Departments of Defense and Homeland Security. For this reason, there should be a provision within this section of FAIIA to create an exception to the confidentiality clause when disclosing content that furthers the national security interest. Furthermore, the provision's reference to 'deidentified data' should be defined explicitly to clarify that it refers to the anonymization of user information only. This would ensure the provision is not misinterpreted as preventing the disclosure of model names, version numbers, and other relevant technical identifiers that are important for transparency, accountability, and further research.
Limitations on regulation and enforcement	Nothing in this bill shall be interpreted to give the Director any enforcement authority [AIARA § 2(f)(2)] The information provided to NIST under this section cannot be used by the government (both Federal and State) to regulate the activities of the entity that provided the information. [AIARA § 2(e)]	Nothing in this bill shall be interpreted to give the Director any enforcement authority. [Proposed § 22B(g)]	The prohibition against using information to regulate the activities of the entity that provided the information is highly risky. Such a rule could undermine the government's ability to enact necessary regulations based on insights gathered from the Center's assessments and research, even if the information reveals risks or issues that warrant urgent intervention. By shielding companies from potential regulatory consequences, this provision may prioritize industry interests over public safety and accountability. It is also exceedingly far-reaching, preventing regulatory action at any level of government.
Miscellaneous obligations	The Under Secretary must also: Support research assessing and mitigating AI safety across timescales [Proposed § 5304(a)(6)(A)]	FAIIA requires the President to issue a technology directive with respect to AI or other automated systems which prohibits any action by a Federal agency that promotes certain concepts like that AI should be designed in an equitable way. [Cruz Amendment 4] Temporary fellows, including consultants and contractors, who are not Federal employees working for any agency under FAIIA will not be able to perform any inherently governmental function and will have to be audited annually. [Cruz Amendment 5] FAIIA also includes provisions to identify regulatory barriers to innovation [Proposed § 102] and to support capacity building through initiatives like Federal Grand Challenges in AI [Proposed § 202]. FAIIA would also establish a tax-exempt foundation which can receive donations to support NIST in the advancement of measurement science. [Young Amendment 1]	The provisions resulting from Cruz Amendment 4 in FAIIA are problematic because they impose a broad restriction on federal agencies that could unnecessarily limit the capacities of these agencies, including AISI, to conduct valuable research. Other provisions in FAIIA that aim to promote research, development, and innovation, along with the establishment of a foundation represent positive steps forward. The restrictions applying to temporary fellows may have negative externalities in that the programs under FAIIA will likely require, and benefit from, the specialized expertise of independent contractors. It would be more effective for these restrictions and auditing requirements to apply to certain types of fellows, as they may inadvertently dissuade qualified researchers and experts from assisting the government in carrying complex, technical programs like the testbed program. Furthermore, while FAIIA's program to identify regulatory barriers to innovation is valuable, the program should also include an identification of insufficiencies in regulation to ensure that further innovation is safe, especially as it accelerates and risks increase. Finally, without very explicit limitations on the structure of the foundation and how it uses donations, private funding of AISI may impose undue influence on its functions. It is important that conflicts of interest are not introduced via this funding source, so that AISI prioritizes research which is in the national interest, as opposed to what corporate funders may want prioritized. Any projects funded by the foundation should be purely research-based, and should be in addition to, rather than in place of, in-house research ambitions. Moreover, these externally-funded projects should have significant oversight to minimize conflicts of interest. FAIIA should also clearly state that private donations to the foundation are not to be taken into account when determining appropriations for AISI's annual budget, given that this may incentivize corporate capture. Funding sourced from donations should be in addition to, rather than in lieu of, public funding streams.

End of table

ANALYSIS

The central goal of both the AI Advancement and Reliability Act (AIARA) and Future of Artificial Intelligence Innovation Act (FAIIA) is the codification of an entity similar to the U.S. AI Safety Institute (AISI) which is currently authorized by Executive Order #14110. Both bills have some beneficial provisions, including the statutory authorization of the US AISI, but they also contain a range of limitations that make them difficult to adopt as wholesale packages.

FAIIA contains several particularly strong provisions, including detailed provisions on the red-teaming, evaluations, and standard-setting functions of its proposed AISI. This AISI would also develop and publish physical and cybersecurity tools, methodologies, and best practices to help those managing AI systems identify and mitigate vulnerabilities, including threats posed by foreign adversaries - something which AIARA lacks. Importantly, FAIIA establishes an AI testbed program to conduct tests, evaluations, and security or vulnerability risk assessments. Providing this structured testing environment would enable stakeholders to systematically identify and address potential weaknesses in powerful AI models, creating an assurance of safety and reliability prior to deployment. Another crucial distinction between the two bills is that AIARA contains a prohibition against the government using any information provided to NIST as part of the program in order to regulate the entity which provided the information. This rule engenders significant risks of preventing the government (both at the Federal and state level) from intervening when it becomes clear that such intervention is urgently necessary, including if a system directly compromises national security. Part of the Center's mission, as envisioned by AIARA, is to increase the government's understanding of AI security and resilience. If dangerous capabilities are identified, the priority should be for the government to have flexibility in considering whether intervention is necessary, taking into account national security interests.

On the other hand, FAIIA includes several aspects that make it comparatively weaker. First, the approach to international collaboration it envisions is severely limited. While FAIIA's exclusion of Chinese engagement addresses mounting geopolitical concerns, it risks hindering essential technical cooperation on AI safety standards at a time when global coordination is critical. The U.S. stands to benefit significantly from establishing a line of communication between the U.S. AISI and its Chinese counterpart, given that U.S. companies frequently open-source their AI models, resulting in what is currently a largely one-sided flow of information. In contrast, AIARA's more balanced and inclusive approach, which allows for cautious engagement with key international AI actors, better supports collective safety efforts while maintaining the flexibility to protect U.S. security interests. Finally, an amendment adopted in the Senate Commerce Markup has resulted in a set of provisions being embedded in FAIIA which would require the President to issue a technology directive with respect to AI or other automated systems prohibiting any action by a Federal agency that would promote certain concepts such as the equitable development of AI. This restriction on federal agencies could undermine AISI's ability to conduct crucial research, including with respect to political biases previously identified in industry-leading AI systems and for improving AI system compliance with existing law.

Critically, both of these bills place AISI (or AIARA's Center) on statutory footing, allowing it the security to engage in longer-term research projects without fear of losing executive authorization prior to their completion. That said, both bills also contain significant shortcomings, many of which are more adequately addressed in their respective

legislative counterparts. A formal or informal conference process adopting the strongest provisions from both bills would therefore strengthen the overall framework guiding AISI or the Center. Without a conference process or other means of integrating preferred provisions from each bill, FAIIA remains the marginally better bill despite its imperfections, as it provides more comprehensive direction for implementing the key objectives of the AISI. Should Congress instead elect to adopt the AIARA in full, we strongly advise, at a minimum, removing the provision prohibiting the use of information collected by the Center for the purpose of regulation, as this prohibition could threaten both the efficacy of the Center and the security of the American public.

PRIORITY PROVISIONS

Future of Artificial Intelligence Innovation Act (FAIIA)

Restrictions on International Collaboration The People's Republic of China is not permitted to participate in an alliance or coalition until the U.S. Trade Representative reports that they have come into compliance with WTO commitments. There must then be a report on the national security, human rights, and commitment monitoring implications of such collaboration. [Budd 2 Amendment]	While there are legitimate security concerns around collaboration with China, this explicit exclusion risks hindering essential AI safety cooperation at a time when global coordination is crucial. Given that U.S. companies tend to open-source their AI models, balanced international collaboration could shift information sharing from one-sided to mutually beneficial. A more flexible approach would empower the Center's Director to determine safe, advantageous levels of collaboration with China on critical safety issues.
Staffing and Funding Under FAIIA, 30 staff members can be appointed until 2035 [Proposed § 302(a)]. FAIIA would also establish a tax-exempt foundation which can receive donations to support NIST in the advancement of measurement science. [Young Amendment 1]	While FAIIA's staffing provisions offer more capacity than AIARA, the cap of 30 appointees through 2035 remains insufficient for achieving its ambitious goals, especially given the rapid expansion of the AI market. Similar bodies typically require closer to 50–75 staff members to function effectively. The U.K. AISI has started with over 30 technical staff members and is rapidly expanding beyond that number. Moreover, FAIIA does not include any provisions of appropriations. It is important that sufficient funding is provided for AISI, such that it does not have to rely on donations received via the foundation. For reference, the U.K. AISI received £100m in initial funding. If the U.S. is to lead on AI, FAIIA must at least match this level of funding.
Technology Directive Restricting Federal Agencies FAIIA requires the President to issue a technology directive on AI which prohibits any action by Federal agencies that promotes certain concepts like that AI should be designed in an equitable way. [Cruz Amendment 4]	These provisions impose an overly broad restriction that could unnecessarily constrain federal agencies, such as the AISI, in their capacity to conduct essential research and development. This restriction could also limit investigations into critical issues, including political biases, or capacity to comply with existing US anti-discrimination laws.

Artificial Intelligence Advancement and Reliability Act (AIARA)

Limitations on regulation and enforcement The information provided to NIST under AIARA cannot be used by the government (both Federal and State) to regulate the activities of the entity that provided the information. [AIARA § 2(e)]	The prohibition in AIARA against using information provided to NIST for regulatory purposes is problematic, as it limits the government's ability to act on critical insights that could indicate significant risks. If information gathered through the research of the Center reveals issues needing urgent intervention, for example on grounds of national security, the government should not be restricted in its ability to regulate. Removing this restriction would ensure that regulatory bodies can respond appropriately to identified risks. AIARA makes it clear that information is voluntarily provided to the Center, and the bill takes adequate measures to protect proprietary rights and strategic interests. These concerns are therefore adequately addressed in other parts of the bill and do not warrant this restriction on the government's ability to regulate in the future.
Lack of a Testbed Program AIARA lacks the testbed program set out in FAIIA which would conduct tests, evaluations, and security or vulnerability risk assessments.	Including a testbed program in AIARA is essential to foster collaboration among government agencies, laboratories, and the private sector, fast-tracking the U.S.'s ability to lead on safe AI. By adopting a testbed program similar to FAIIA, AIARA would encourage and support the development of leading AI systems that meet security and robustness standards, enhancing public trust and supporting companies of all sizes.
Staffing and Funding Under AIARA, up to 15 staff members can be appointed until 2029 [AIARA § 2(c)]. $10,000,000 is allocated for fiscal year 2025. [AIARA § 2(h)]	The staffing provision in AIARA is inadequate for meeting the substantial demands of overseeing and guiding AI development. For the reasons mentioned above, a team of only 15 staff members would struggle to address a wide range of responsibilities AIARA envisions, from developing standards to conducting assessments and fostering public-private collaborations. While some funding is set aside for the Center, $10 million represents a sliver of what has been allocated to similar entities, like the U.K.'s AI Safety Institute.