DECEMBER 2024

BIPARTISAN HOUSE TASK FORCE REPORT ON ARTIFICIAL INTELLIGENCE

Guiding principles, forward-looking recommendations, and policy proposals to ensure America continues to lead the world in responsible AI innovation

118TH CONGRESS

U.S. HOUSE OF REPRESENTATIVES

Table of Contents

Bipartisan House Task Force on Artificial Intelligence

The Honorable Mike Johnson
Speaker
United States House of Representatives
Washington, DC 20515

The Honorable Hakeem Jeffries
Democratic Leader
United States House of Representatives
Washington, DC 20515

Dear Speaker Johnson and Leader Jeffries:

We, the Co-Chairs of the Bipartisan Artificial Intelligence Task Force, submit to you our key findings in this report.

Although artificial intelligence (AI) is not a new concept, breathtaking technological advancements in the last few years have made AI the focus of numerous policy discussions. AI has tremendous potential to transform society and our economy for the better and address complex national challenges. From optimizing manufacturing to developing cures for grave illnesses, AI can greatly boost productivity, enabling us to achieve our objectives more quickly and cost-effectively. Nevertheless, we also recognize that AI can be misused and lead to various types of harm.

This report highlights America's leadership in its approach to responsible AI innovation while considering guardrails that may be appropriate to safeguard the nation against current and emerging threats. You charged twenty-four members, twelve Republicans and twelve Democrats, with developing a U.S. vision for AI adoption, innovation, and governance. The AI Task Force gathered information on salient AI issues from domain experts in industry, government, civil society, and academia to provide 66 key findings and 85 recommendations. In summary, this report encapsulates a targeted approach that balances the need to promote vibrant AI innovation while safeguarding Americans from potential harms as we enter an era of widespread adoption of AI.

We thank you for establishing the AI Task Force and are eager for this report to inform future congressional policymaking.

Sincerely,

Ted W. Lieu

Jay Obernolte
CHAIRMAN

Ted W. Lieu
CO-CHAIRMAN

About the Bipartisan House AI Task Force in the 118th Congress

The bipartisan AI Task Force was created by Speaker Johnson and Democratic Leader Jeffries on February 20, 2024. The AI Task Force is led by co-chairs Jay Obernolte (R-CA) and Ted Lieu (D-CA) and comprises twenty-four members, twelve Republicans and twelve Democrats. The AI Task Force members are drawn from twenty committees to ensure comprehensive jurisdictional responsibilities over the numerous AI issues that we addressed and to benefit from a range of different insights and perspectives.

A full list of Task Force members and the committees they represent is included in **Appendix I**.

Throughout 2024, the AI Task Force convened to investigate dozens of issues at the heart of how AI intersects with numerous policy areas. The AI Task Force held multiple hearings and numerous roundtables and engaged with over one hundred experts, including business leaders, government officials, technical experts, academics, legal scholars, and other domain specialists. These experts generously offered their insights, suggestions, and comments spanning a range of viewpoints.

This approach allowed each issue to be comprehensively explored from various perspectives. A multifaceted approach to policy analysis will better prepare the decision-makers who address the complex AI challenges that confront our nation and will continue to affect public policy.

A full list of experts and a list of the events the Task Force convened is included in **Appendix II**.

Leading AI Progress: Policy Insights and a U.S. Vision for AI Adoption, Responsible Innovation, and Governance

The United States is the global leader in Artificial Intelligence, a transformative technology that also comes with risks that must be addressed responsibly. To ensure the economic and national security benefits accrue in the United States and the technology is developed and deployed responsibly here and across the world, the United States must take active steps to safeguard our current leadership position. Developed and deployed responsibly, AI has the potential to help improve Americans' quality of life, health, jobs, security, and economic prosperity for decades to come.

The United States leads the world in AI research, the number of AI companies, private sector AI investment, and industry adoption of AI. This overwhelming national advantage derives from two of our longstanding strengths: we have cultivated a thriving innovation ecosystem and a flexible sectoral regulatory framework. If maintained, these strengths will help our country remain the world's undisputed leader in the responsible design, development, and deployment of AI.

The collective experiences and insights of the bipartisan House AI Task Force are encapsulated in this report. During the second session of the 118th Congress, the Task Force engaged with over one hundred experts on dozens of salient AI issues. We consulted with business leaders, government officials, technical experts, academics, and legal scholars, all of whom offered their insights, suggestions, and comments on the varied and complex AI challenges confronting our nation.

This report articulates guiding principles, 66 key findings, and 85 recommendations, organized into 15 chapters. It is intended to serve as a blueprint for future actions that Congress can take to address advances in AI technologies. The Task Force members feel strongly that Congress must develop and maintain a thoughtful long-term vision for AI in our society. This vision should serve as a guide to the many priorities, legislative initiatives, and national strategies we undertake in the years ahead.

In considering new policies, Congress should adopt an agile approach that allows us to respond appropriately and in a targeted, achievable manner that benefits from all available evidence and insights. Supporting this agile paradigm requires continual learning and adaptation. Congress should regularly evaluate the effectiveness of its policies and update them as AI technologies and their impacts evolve. If we follow this approach and take strategic action while encouraging innovation, we can lead in AI development and drive a global vision for AI public policy.

This report is certainly not the final word on AI issues for Congress. Instead, it should be viewed as a tool for identifying and evaluating AI policy proposals. Given the breadth of AI policy opportunities and challenges, the report also includes a list of areas for future exploration.

Philosophy and Principles

In addition to evaluating specific AI issues, the Task Force adopted several high-level principles to frame this policy analysis. These principles represent high-level policy considerations that transcend specific AI issues and can help guide future congressional efforts. The principles we established are:

- **Identify AI Issue Novelty**
- **Promote AI Innovation**
- **Protect Against AI Risks and Harms**
- **Empower Government with AI**
- **Affirm the use of a Sectoral Regulatory Structure**
- **Take an Incremental Approach**
- **Keep Humans at the Center of AI Policy**

Principle: Identify AI Issue Novelty

Policymakers can avoid duplicative mandates if they consider whether issues raised by AI are truly novel and without precedent or if existing laws and regulations already address the underlying concern. For each AI issue investigated, the novelty of the issue should be identified to understand whether the issue is:

- **Truly new for AI due to capabilities that did not previously exist.** When an AI issue has emerged recently due to the nature of available AI technology, this suggests that we need to more thoroughly consider how well existing regulatory regimes address that issue.
- **An existing issue that's nature has been changed significantly by AI.** If AI is exacerbating an existing issue, this suggests that the issue merits consideration but is in concert with the existing policy paradigm. Existing approaches to the issue might be appropriate, but new approaches may also be. Existing regimes may also not be designed to address evolving technologies. Congress should strive to modernize laws and regulations to ensure they are sufficiently technology-neutral in application and enforcement.
- **An existing issue that has not been significantly changed by contemporary AI capabilities.** If AI is just one of many ways of accomplishing an old purpose with substantially the same effect, then this suggests that existing laws, regulations, and regulatory bodies are perhaps best positioned to assess and address that issue.

Principle: Promote AI Innovation

As the global leader in AI development and deployment, the United States is best positioned to responsibly enable the potential of this transformative technology for all. To maintain this leadership and enable the U.S. economy to harness the full benefits of AI, policymakers should continue to promote AI innovation.

Principle: Protect Against AI Risks and Harms

We have an obligation to protect Americans from both accidental and malicious uses of AI. Meaningful AI governance will require a combination of technical and policy solutions that seek to understand, identify, and mitigate the potential risks and harms from the development and deployment of AI systems. A thoughtful, risk-based approach to AI governance can promote innovation rather than stifle it.

Moreover, for every problem that AI creates, AI can be a candidate for helping to remediate or solve that problem. While technological solutions are not always possible, their use in AI policy should be borne in mind as policy is developed. This is especially important as AI technology continues to evolve rapidly, potentially presenting new issues more frequently than in past technological revolutions.

Principle: Government Leadership in Responsible Use

Trust is a necessary component for the widespread adoption of AI by the public and private sectors in the United States. The federal government should foster that trust by adopting responsible principles and policies that capture the benefits of AI while addressing its risks and leading by example. Powering government services with AI is also necessary, given their prominent role in our economy. We must ensure that Congress, federal agencies, courts, and other government entities utilize AI to improve their services, speed, efficiency, and quality

Principle: Support Sector-Specific Policies

For an agile and focused approach to AI policy, sector-specific regulators within federal agencies and other parts of government should use their existing authority to respond to AI use within their individual domains of expertise and the context of the AI's use. This would enable more informed and efficient engagement between federal agencies and entities utilizing AI. Agency expertise should remain focused on where it can be most effective.

Sector-specific regulators would also benefit from drawing upon a federal repository of AI resources. Examples of these resources include AI expertise, AI-ready data, computing hardware, technical resources, and evaluations that allow AI risks to be assessed safely. Additionally, coordination among federal agencies through interagency structures could improve access to such resources. Improved access and coordination could empower agencies with AI skills and allow them to share lessons learned while ensuring they continue to specialize in what they do best.

Principle: Take an Incremental Approach

AI is a rapidly evolving technology. It is unreasonable to expect Congress to enact legislation this year that could serve as its last word on AI policy. To use AI technology properly requires a carefully designed, durable policy framework. In this report, we propose a number of recommendations to begin to build this framework with the understanding that as AI capabilities continue to advance, we must remain humble and acknowledge we do not know what we do not know. Policy will likely need to adapt and evolve in tandem with advances in AI. Congress must remain vigilant and flexible in how it addresses AI in the years to come.

Principle: Keep Humans at the Center of AI Policy

AI systems reflect the principles of the people who design them and require human input to train them. The United States will also need to attract, train, and retain the talent to remain competitive in this technology. Further, the automation that AI brings will have some labor market effects. As policymakers consider laws and regulations focused on AI development and governance, they should focus on human impact and human freedom.

Bipartisan House Task Force on Artificial Intelligence: Overview of Recommendations

Government Use

Federal agencies have already begun leveraging AI to empower existing agency missions and streamline programs. While use cases vary in application and maturity, the benefits of responsible government use of AI are potentially transformative. However, irresponsible or improper use fosters risks to individual privacy, security, and the fair and equal treatment of all citizens by their government.

Key Findings

- The federal government should utilize core principles and avoid conflicting with existing laws.
- The federal government should be wary of algorithm-based decision-making.
- The federal government should provide notification of AI's role in governmental functions.
- Agencies should pay attention to the foundations of AI systems.
- Roles and associated AI knowledge and skills are unclear and highly varied across the federal workforce.
- Skills-based hiring is critical for filling the demand for AI talent in the federal workforce.

Recommendations

- Take an information and systems-level approach to the use of AI in the federal government.
- Support flexible governance.
- Reduce administrative burden and bureaucracy using AI.
- Require that agencies provide notification of AI's role in governmental functions.
- Facilitate and adopt AI standards for federal government use.
- Support NIST in developing guidelines for federal AI systems.
- Improve cybersecurity of federal systems, including federal AI systems.
- Encourage data governance strategies that support AI development.
- Congress and the government must understand the federal government's AI workforce needs.
- Support different pathways into federal service for AI talent.

Federal Preemption of State Law

Preemption of state AI laws by federal legislation is a tool that Congress could use to accomplish various objectives. However, federal preemption presents complex legal and policy issues that should be considered.

Key Findings

- Federal preemption of state law on AI issues is complex.
- Federal preemption has benefits and drawbacks.
- Preemption can allow state action subject to floors or ceilings.
- Preemption can be multifaceted.
- Definitions must be fit for purpose.

Recommendations

- Study applicable AI regulations across sectors.

Data Privacy

As AI systems amass and analyze vast amounts of data, there are increasing risks of private information being accessed without authorization. Thoughtful and effective data privacy policies and protections will support consumer confidence in the responsible development and deployment of AI systems.

Key Findings

- AI has the potential to exacerbate privacy harms.
- Americans have limited recourse for many privacy harms.
- Federal privacy laws could potentially augment state laws.

Recommendations

- Explore mechanisms to promote access to data in privacy-enhanced ways.
- Ensure privacy laws are generally applicable and technology-neutral.

National Security

Like any major dual-use technology, AI has the potential to both bolster and undermine national security. This underscores its significance in U.S. defense strategy. Currently, the U.S. national security ecosystem is both using and developing AI, but a significant proportion of research and development related to AI is occurring outside of government activities.

Key Findings

- AI is a critical component of national security.
- U.S. adversaries are adopting and militarizing AI.
- National security requires advanced cloud access and AI.
- National security requires AI for contested environments.
- AI can vastly improve DOD business processes.

Recommendations

- Focus congressional oversight on AI activities for national security.
- Support expanded AI training at DOD.
- Continue oversight of autonomous weapons policies.
- Support international cooperation on AI used in military contexts.

Research, Development, & Standards

The U.S. remains the leader in fundamental research and standards and consistently produces cutting-edge AI applications before other nations. To maintain U.S. leadership in global AI innovation and governance, Congress will need to continue federal R&D efforts, supporting AI evaluations, and bolstering U.S. standardization efforts for AI.

Key Findings

- Federal investments in fundamental research have enabled the current AI opportunity.
- Continued AI research and evaluation will promote AI advancement.
- Progress in AI R&D is closely linked to access to AI resources.
- A closed AI research ecosystem could limit U.S. competitiveness in AI.
- University AI R&D is necessary but must be paired with vibrant technology transfer activities.
- Advancing the science around AI evaluation will help advance adoption.
- The U.S. is a global leader in standard setting but faces competitors.

Recommendations

- Continually monitor and evaluate the impact of AI on different industries and the nation.
- Support fundamental R&D for continued leadership in AI innovation.
- Increase technology transfer from university R&D to market.

- Promote public-private partnerships for AI R&D.
- Promote research and standardization surrounding the evaluation and testing of AI.
- Promote the development of infrastructure and data to enable AI research.
- Continue engagement in international standards development.
- Uphold the U.S. approach to setting standards.
- Align national AI strategy with broader U.S. technology strategy.
- Explore how to accelerate scientific discovery across disciplines with AI.
- Support AI R&D by small businesses.
- Encourage international collaboration with like-minded allies and partners on R&D.

Civil Rights & Civil Liberties

Adverse effects from flawed or misused technologies are not new developments but are consequential considerations in designing and using AI systems. AI models, and software systems more generally, can produce misleading or inaccurate outputs. Acting or making decisions based on flawed outputs can deprive Americans of constitutional rights.

Key Findings

- Improper use of AI can violate laws and deprive Americans of our most important rights.
- Understanding the possible flaws and shortcomings of AI models can mitigate potentially harmful uses of AI.

Recommendations

- Have humans in the loop to actively identify and remedy potential flaws when AI is used in highly consequential decision-making.
- Agencies must understand and protect against using AI in discriminatory decision-making.
- Empower sectoral regulators with the tools and expertise to address AI-related risks in their domains.
- Explore transparency for users affected by decisions made using AI.
- Support standards and technical evaluations to mitigate flawed decision-making involving AI systems.

Education & Workforce

Despite federal and state efforts, the U.S. has a significant gap in the appropriate talent needed to research, develop, and deploy AI applications—and this gap is growing. Educating and training American learners in AI topics will be critical to continued U.S. leadership in AI technology and for America's economic and national security.

Key Findings

- AI is increasingly used in the workplace by both employers and employees.
- Fostering domestic AI talent and continued U.S. leadership will require significant improvements in basic STEM education and training.
- U.S. AI leadership would be strengthened by utilizing a more skilled technical workforce.
- AI adoption in America requires AI literacy.
- K–12 educators need resources to promote AI literacy.

Recommendations

- Invest in K–12 STEM and AI Education and Broaden Participation.
- Bolster U.S. AI skills by providing needed AI resources.
- Develop a full understanding of the AI workforce in the United States.
- Facilitate public-private partnerships to bolster the AI workforce.
- Develop regional expertise when supporting government-university-industry partnerships.
- Broaden pathways to the AI workforce for all Americans.
- Support the standardization of work roles, job categories, tasks, skill sets, and competencies for AI-related jobs.
- Evaluate existing workforce development programs.
- Promote AI literacy across the U.S.
- Empower U.S. educators with AI training and resources.
- Support NSF curricula development.
- Monitor the interaction of labor laws and worker protections with AI adoption.

Intellectual Property

Advances in generative AI technology have introduced new issues for intellectual property (IP) laws, raising questions about how the ownership, creation, and protection of art, writings, brands, songs, inventions, and other creations should be treated.

Key Findings

- It is unclear whether legislative action is necessary in some cases, and a number of IP issues are currently in the courts.
- Generative AI poses a unique challenge to the creative community.
- It is often difficult for creators to know if their copyrighted works are being used by AI developers.
- The global IP policy landscape presents challenges and opportunities to both developers and creators.
- While some use cases are legitimate and protected forms of expression, the proliferation of deepfakes and harmful digital replicas is a significant and ongoing challenge.

Recommendations

- Clarify IP laws, regulations, and agency activity.
- Appropriately counter the growing harm of AI-created deepfakes.

Content Authenticity

Generative AI systems include AI that can generate text, image, video, and audio/voice content. These systems are trained on a large set of existing written, visual, or audio data. The systems identify statistical patterns in this training data and then create novel content that matches these patterns. As generative AI systems continue to be trained with greater amounts of data and more powerful computing resources, they can produce outputs with increasing quality and realism.

Key Findings

- Synthetic content has many beneficial uses, but if used improperly it can create harms and undermine confidence in information integrity.
- There is currently no single, optimal technical solution to content authentication.
- Technical literacy would help with the content authenticity challenges but would not be sufficient.
- Digital identity technology allows a person online to verify who they are and reduces fraud.

Recommendations

- Support a risk-based, multipronged approach to content authenticity.
- Support technical solutions to content authenticity.
- Address demonstrable harms, not speculative harms of synthetic content.
- Identify the responsibilities of AI developers, content producers, and content distributors when it comes to synthetic content.
- Examine existing laws related to harmful synthetic content.
- Ensure victims have the necessary tools.

Open & Closed Systems

Despite often being characterized as either open or closed, there is in fact a continuum of different forms of AI model availability and transparency. Open models offer many benefits, including customization, transparency, and accessibility. However, there is an increased risk that malicious actors could use open models to cause harm, including perpetrating financial fraud, threatening national security, or large-scale identity theft.

Key Findings

- Open AI models encourage innovation and competition.
- There is currently limited evidence that open models should be restricted.

Recommendations

- Encourage innovation and competition in the development of AI models.
- Focus on demonstrable harms and physical threats.
- Evaluate chemical, biological, radiological, or nuclear (CBRN) threats in light of AI capabilities.
- Continue to monitor the risks from open-source models.

Energy Usage & Data Centers

Significant amounts of power are needed to create and use the most advanced AI models, and the data centers supporting AI have increased electricity demand. This creates challenges for electrical grid reliability and affordable electricity, which must be addressed for U.S. economic and national security. AI can also be a valuable tool for developing American energy supplies.

Key Findings

- AI is critical to U.S. economic interests and national security and maintaining a sufficiently robust power grid is a necessity.
- The growing demands of AI are creating challenges for the grid.
- Continued U.S. innovation in AI requires innovations in the energy sector.
- Planning properly now for new power generation and transmission is critical for AI innovation and adoption.
- AI tools will play a role in innovation and modernization in the energy sector.

Recommendations

- Support and increase federal investments in scientific research that enables innovations in AI hardware, algorithmic efficiency, energy technology development, and energy infrastructure.
- Strengthen efforts to track and project AI data center power usage.
- Create new standards, metrics, and a taxonomy of definitions for communicating relevant energy use and efficiency metrics.
- Ensure that AI and the energy grid are a part of broader discussions about grid modernization and security.
- Ensure that the costs of new infrastructure are borne primarily by those customers who receive the associated benefits.
- Promote broader adoption of AI to enhance energy infrastructure, energy production, and energy efficiency.

Small Business

Small businesses play a crucial role in maintaining the United States' lead in the AI race against other world powers. Unfortunately, small businesses often lack the understanding or resources that would allow them to meaningfully adopt this critical technology.

Key Findings

- Small businesses can lack a full understanding of how best to adopt AI.
- Small businesses can lack sufficient access to capital and AI resources.
- Small businesses face excessive challenges in meeting AI regulatory compliance.

Recommendations

- Support small business AI literacy.
- Provide resources for small business AI adoption.
- Investigate the resource challenges of small businesses adopting AI.
- Investigate the resource challenges of small AI businesses.
- Ease compliance burdens for small businesses.

Agriculture

AI has emerged as a powerful tool capable of revolutionizing agriculture. AI advancements have the potential to increase food availability, lower food prices, and bolster economic growth.

Key Findings

- AI-driven precision agriculture could enhance farm productivity and natural resource management.
- Increased AI integration could enable mechanization and automation technologies and enhance efficiency within the specialty crop industry.
- Lack of reliable network connectivity in rural and farming communities impedes AI adoption in the agricultural sector.
- AI is already a powerful tool in addressing and combating the wildfire and forest health crises.
- Greater adoption of AI at USDA could enhance delivery of numerous agriculture programs and reduce costs for farmers and others.
- The CFTC's principles-based approach allows for flexibility in addressing new technologies.

Recommendations

- Assess existing programs to identify opportunities for advancing AI in precision agriculture.
- Pursue further AI research and development to enhance efficiency in specialty crops.
- Continue to explore how research and innovation in AI technology could aid land managers in improving forest health through better planning and strategies.
- Direct USDA to better utilize AI in program delivery.
- Continue to review the application of the CFTC's principles-based framework to ensure it captures unique risks posed by AI in financial markets.

Healthcare

AI technologies have the potential to improve multiple aspects of healthcare research, diagnosis, and care delivery. AI can quickly analyze large data sets, improve diagnostic accuracy, streamline operations and automate routine tasks, all of which have the potential to improve efficiency and efficacy in treatment and reduce burdens on healthcare practitioners, freeing up more time for patient care.

Key Findings

- AI's use in healthcare can potentially reduce administrative burdens and speed up drug development and clinical diagnosis.
- The lack of ubiquitous, uniform standards for medical data and algorithms impedes system interoperability and data sharing.

Recommendations

- Encourage the practices needed to ensure AI in healthcare is safe, transparent, and effective.
- Maintain robust support for healthcare research related to AI.
- Create incentives and guidance to encourage risk management of AI technologies in healthcare across various deployment conditions to support AI adoption and improve privacy, enhance security, and prevent disparate health outcomes.
- Support the development of standards for liability related to AI issues.
- Support appropriate payment mechanisms without stifling innovation.

Financial Services

The financial services sector has employed AI technologies for decades. The ideal environment for continued growth would allow AI innovation to thrive while protecting consumers and maintaining market integrity. By focusing on fostering innovation, enhancing customer experiences, and ensuring financial inclusion, AI can significantly improve the financial sector's efficiency and accessibility.

Key Findings

- AI presents an opportunity to transform the financial services sector.
- Data quality and data security are paramount in financial service AI models.
- AI can expand access to financial products and services.
- AI technologies are already deployed across the financial services sector.
- Some regulators use AI to identify non-compliance with regulations.
- Small financial services firms can be at a disadvantage in AI adoption.

Recommendations

- Foster an environment where financial services firms can responsibly adopt the benefits of AI technology.
- Encourage and resource regulators to increase their expertise with AI.
- Maintain consumer and investor protections in the use of AI in the financial services and housing sectors.
- Consider the merits of regulatory “sandboxes” that could allow regulators to experiment with AI applications.
- Support a principles-based regulatory approach that can accommodate rapid technological changes.
- Ensure that regulations do not impede small firms from adopting AI tools.

GOVERNMENT USE

Background

Federal agencies have already begun leveraging artificial intelligence (AI) in various use cases to empower existing agency missions and streamline programs.[1] A February 2020 report from Stanford University found that "nearly half of the federal agencies studied have experimented with AI and related machine learning tools."[2] As part of Executive Order 13960, signed in 2020, the Trump Administration directed federal agencies to create an inventory of AI use cases.[3]

There are several prominent examples from the Department of State, Department of Justice, and Centers for Disease Control and Prevention.[4] Since 2020, the federal government has continued to add more AI use cases. A December 2023 report from the U.S. Government Accountability Office (GAO) found that 20 of the 23 surveyed agencies use AI and collectively reported approximately 200 instances of AI use.[5] AI.gov also publicizes AI use cases across the federal government, including a portal for professionals and students to join the national AI talent surge.**[6]**

[1] The White House. "AI Use Cases." AI.gov, https://ai.gov/ai-use-cases/.
[2] David Freeman, et al. "Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies." Stanford Law, February 2020, https://law.stanford.edu/wp-content/uploads/2020/02/ACUS-AI-Report.pdf.
[3] Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government. Federal Register, vol. 85, no. 236, December 2020, https://www.federalregister.gov/documents/2020/12/08/2020-27065/promoting-the-use-of-trustworthy-artificial-intelligence-in-the-federal-government.
[4] Lewis Kamb, "Some U.S. government agencies are testing out AI to help fulfill public records requests," NBC News, 1 Aug. 2023. https://www.nbcnews.com/news/us-news/federal-agencies-testing-ai-foia-concerns-rcna97313.
[5] U.S. GAO, "Artificial Intelligence: Agencies Have Begun Implementation but Need to Complete Key Requirements," Government Accountability Office, Dec 2023, https://www.gao.gov/products/gao-24-105980.
[6] Supra 1.

While use cases vary in application and maturity, the benefits of responsible government use of AI are potentially transformative. However, irresponsible or improper use fosters risks to individual privacy, security, and the fair and equal treatment of all citizens by their government.

Responsible Government Use of Artificial Intelligence

There is no single source of comprehensive guidance on what responsible government use of AI entails. However, multiple government entities have holistically identified and defined guiding principles for responsible and trustworthy AI, which guide the design, use, and deployment of automated systems. This evolving list of AI guiding principles for government use includes but is not limited to: [7,8]

- Accuracy
- Reliability
- Robustness
- Safety and effectiveness
- Security
- Privacy
- Transparency
- Explainability and interpretability
- Notice and explanation
- Human alternatives, considerations, and fallback
- Equity
- Mitigation of harmful bias

While consensus around AI governance principles would provide a useful starting point to meaningfully address responsible AI usage, agencies still face significant challenges in curating the approaches needed to fulfill them. Operationalizing principles across the AI lifecycle is a challenging and complex task.[9]

[7] AI Bill of Rights. The White House, https://www.whitehouse.gov/ostp/ai-bill-of-rights/.

[8] IT Modernization Centers of Excellence. “AI Guide for Government: A living and evolving guide to the application of artificial intelligence for the U.S. Federal Government. U.S. General Services Administration, https://coe.gsa.gov/coe/ai-guide-for-government/evolving-principles-and-guidelines/index.html.

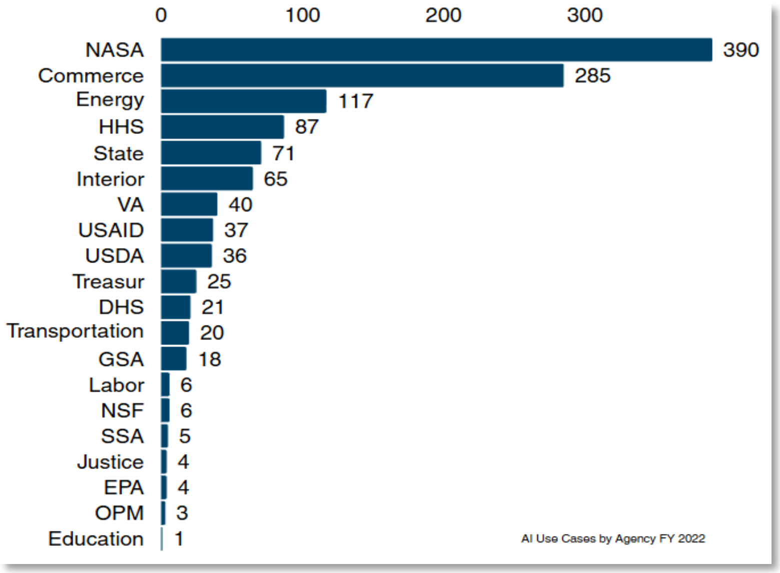
[9] Qinghua Lu, Liming Zhu, Xiwei Xu, Jon Whittle, Didar Zowghi, and Aurelie Jacquet. Responsible AI Pattern Catalogue: A Collection of Best Practices for AI Governance and Engineering. ACM Comput. Surv. 56, 7, Article 173. July 2024, 35 pages. https://doi.org/10.1145/3626234.

To address these challenges, ad-hoc guidance has been published over the years, such as the Government Accountability Office's Accountability Framework for Federal Agencies and the General Services Administration's (GSA) AI Guide for Government.[10,11]

While such resources can support agencies in developing AI governance approaches, none offer a holistic approach to implementing guiding principles. The Biden Administration issued Executive Order 14110 on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, which required the Office of Management and Budget (OMB) to establish an interagency council to coordinate and develop guidance on federal agency AI use, governance, and risk management.

In response, OMB published M-24-10 on March 28, 2024, which "establishes new agency requirements and guidance for AI governance, innovation, and risk management, including through specific minimum risk management practices for uses of AI that impact the rights and safety of the public."[12]

The memorandum requires agencies to update any existing internal AI guiding principles and guidelines to ensure consistency with the new guidance, to implement the Executive Order's minimum practices by December 1, 2024 (with extensions possible), and to stop using any AI in their operations that is not compliant with the minimum practices by that date.



Source: GAO Report - Artificial Intelligence: Agencies Have Begun Implementation but Need to Complete Key Requirements.

As the federal government continues implementing various laws, executive orders, and guidance on federal AI use, such actions should be considered in the context of the existing policies governing federal information systems, data, cybersecurity, and procurement.

[10] U.S. GAO. "Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities." U.S. Government Accountability Office, June 2021, https://www.gao.gov/products/gao-21-519sp.
[11] Supra 8.
[12] NOTE: Excludes elements of the IC and DoD. Shalanda Young. "Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence." The White House, https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf.

Furthermore, the federal government has tools and policies in place to support the procurement of more AI systems, and recent OMB guidance has set forth additional AI acquisition requirements.[13]

To reduce any potential redundancy, ambiguity, or conflicting guidance, developing new policies involving AI use and procurement should start with an analysis and understanding of how existing policies and procedures can be applied. Legislatively harmonious solutions and holistic operational resources spanning the AI life cycle are needed to enable consistently managed government AI systems.

The following section describes (non-exhaustive) information that might be documented to operationalize responsible AI principles and how that information might be shared to actualize transparency.

Federal AI Governance and Transparency

As agencies use AI today and identify use cases for the future, Congress should ensure necessary safeguards are in place to protect the public's privacy, security, civil rights, and civil liberties. The public should know that federal agencies have mature policies to leverage AI while safeguarding against the risks presented by algorithmic-based decision-making that inappropriately rely on AI systems without the necessary governance and transparency policies to ensure proper and effective use.

Governance and transparency requirements around federal agency use of AI can provide important information about government AI systems and be used to inform internal government management, Congress, the public, and stakeholders impacted by AI-informed outputs.

Transparent reporting requirements are one of many tools to inform future policymaking and provide the detailed, actionable, and timely information needed to ensure federal agencies consistently use AI responsibly and effectively.[14] Views differ on the type and breadth of information that agencies should document and disclose, but they could include:

- **Data and Metadata**: information about the data that was used to train, test, or fine-tune the model, including information about the data's sources or provenance, collection methods, sample size, procedures for cleaning the data, bias and skewness, inclusion of protected characteristics or proxy features, and ultimate integrity.

[13] Letter from The Alliance for Digital Innovation to David A. Myklegard, Deputy Federal Chief Information Officer, and Christine J. Harada, Senior Advisor on Federal Procurement Policy, Office of Management and Budget 29 April 2024, https://alliance4digitalinnovation.org/wp-content/uploads/2024/04/2024.04.29-ADI-Comments-on-Responsible-Procurement-of-Artificial-Intelligence-in-Government-RFI-Final.pdf.

[14] Alex Engler, "The AI Regulatory Toolbox: How Governments Can Discover Algorithmic Harms," Brookings, 2023. https://www.brookings.edu/articles/the-ai-regulatory-toolbox-how-governments-can-discover-algorithmic-harms/.

- **Software**: information about the software components and their origins.[15]
- **Model Development**: information about the training, tuning, validation, and testing of the AI system, who requested its development, who developed the model, communities consulted in development, the development process, tools used in development, the model's intended uses and known limitations, and metrics pertaining to the model's efficiency, performance, bias, and energy usage.
- **Model Deployment:** information about model deployment and monitoring, including metrics identified in model development, plans to provide notice and explanation of the use of AI models to members of the public impacted by the model's use, and any ongoing training, validation, and testing.
- **Model Use**: information about how the model is used, including the organizational context and design of the entire system in which the model is deployed, specific use case applications of the model, the information that a deployed model utilizes, the types of determinations or decisions the model is intended to inform, meaningful explanations of the model and its outcomes given relevant stakeholders, the policies for how to handle outputs, the risks of harm identified, and risk mitigation plans including human oversight or intervention.

It can be difficult for federal agencies to balance the desire for transparency against protections for privacy, security, proprietary information, and national security, and as a result, transparency does not always mean complete disclosure to the public.[16] In such cases, transparency may rely on documentation regarding the data collection and testing methodology rather than access to the underlying test data.[17]

Reporting and transparency policies should be designed to enable appropriate governance of different internal governmental functions, such as enabling oversight functions internal to the intelligence community, the controlled sharing of statistical information, or Congress's oversight role over executive branch agencies.

Federal Standards for AI Systems

While the U.S. government primarily plays a supportive role in developing international standards related to information technology (IT), the federal government does set its own standards for government systems. These standards are usually based on or align with international consensus standards.

[15] National Telecommunications and Information Administration. "Software Bill of Materials" Department of Commerce, National Telecommunications and Information Administration, https://www.ntia.gov/page/software-bill-materials.
[16] Olsen, Henrik Palmer, et al. "The Right to Transparency in Public Governance: Freedom of Information and the Use of Artificial Intelligence by Public Agencies." Digital Government: Research and Practice, vol. 5, no. 1, 12 March 2024, pp. 1–15, https://doi.org/10.1145/3632753.
[17] Id.

In 1995, Congress signed the National Technology Transfer and Advancement Act (NTTAA)[18] to guide federal agencies' standard-setting activities. The NTTAA directs federal agencies to adopt voluntary consensus standards wherever possible to avoid duplication of efforts. It also makes federal agencies responsible for evaluating the efficacy of adopting standards through conformity assessment activities.

In supporting or adopting standards, each agency must coordinate its activities with those of other appropriate agencies and the private sector. To provide agencies with guidance on implementing NTTAA, OMB maintains OMB Circular A-119 ("Federal Participation in the Development and Use of Voluntary Consensus Standards and in Conformity Assessment Activities").

The National Institute of Standards and Technology (NIST) is responsible for promulgating the standards that broadly underpin federal computer systems, called the Federal Information Processing Standards (FIPS), in accordance with the E-Government Act of 2002 (P.L. 107-347)[19] and the Federal Information Security Modernization Act (P.L. 113-283).[20]

In promulgating these standards, NIST must ensure the FIPS adheres to voluntary consensus standards wherever possible to avoid duplication of efforts in accordance with the NTTAA.

OMB is the agency tasked with overseeing and coordinating federal information management, including IT management. Congress has directed OMB to develop and oversee agency IT policies and practices, including leading the government-wide implementation of standards promulgated by NIST and enforcing agency policies consistent with such standards.

While OMB issues policies and guidance, NIST continues to support the implementation of such standards by providing technical support to other agencies as needed. The Cybersecurity and Infrastructure Security Agency also plays a major role in helping to facilitate the implementation of cybersecurity standards in government IT systems.

Specific to AI, the AI in Government Act (P.L. 116-260)[21] required OMB to issue government-wide guidance on agency use of AI and agency AI governance plans.

[18] Office of the Federal Register, National Archives and Records Administration. Public Law 107 - 347 - E-Government Act of 2002. U.S. Government Printing Office, 16 Dec. 2002, https://www.govinfo.gov/app/details/PLAW-107publ347.

[19] Id.

[20] Office of the Federal Register, National Archives and Records Administration. Public Law 113 - 283 - Federal Information Security Modernization Act of 2014. U.S. Government Publishing Office, 17 Dec. 2014, https://www.govinfo.gov/app/details/PLAW-113publ283.

[21] Office of the Federal Register, National Archives and Records Administration. Public Law 116 - 260 - Consolidated Appropriations Act, 2021. U.S. Government Publishing Office, 26 Dec. 2020, https://www.govinfo.gov/app/details/PLAW-116publ260.

Further, the Advancing American AI Act (P.L. 117–263)[22] required specified federal agencies to take steps to promote responsible AI acquisition and use while protecting privacy, civil rights, and civil liberties. For example, the Department of Homeland Security (DHS) must issue policies and procedures for DHS related to the acquisition and use of AI and considerations for risks and ramifications of AI-enabled systems. The Advancing American AI Act also directed OMB to require federal agencies to prepare, maintain, and make publicly available inventories of their current and planned AI use cases.

In the National AI Initiative Act (P.L. 116-283),[23] Congress directed NIST to support AI standards and develop a voluntary AI risk management framework by collaborating with stakeholders across the public and private sectors.

In July 2021, NIST launched a request for information to develop a framework to better manage AI risks to individuals, organizations, and society. After substantial collaboration with public and private sector partners, NIST released its AI Risk Management Framework and accompanying materials on January 26, 2023, to help guide the safe and responsible development and use of AI.[24] However, the first iteration of the framework only sets the theoretical baseline for identifying and mitigating AI risks by guiding readers in thinking critically about the context, measurement, and management of AI systems and so, is not a standards document.

As discussed in the **Research, Development, & Standards chapter,** AI-related standards are significantly underdeveloped. This includes standards for federal systems.

AI-Enabling Infrastructure

Effective public sector AI system governance must be managed across the AI system's entire lifecycle. Maintaining sound policies over the federal IT and data that support AI systems will enable better governance over the AI ultimately used by agencies, especially policies that focus on maintaining high-quality data, ensuring data governance, and fostering technical capability within the public sector workforce.[25]

[22] Office of the Federal Register, National Archives and Records Administration. Public Law 117 - 263 - James M. Inhofe National Defense Authorization Act for Fiscal Year 2023. U.S. Government Publishing Office, 22 Dec. 2022, https://www.govinfo.gov/app/details/PLAW-117publ263.

[23] Office of the Federal Register, National Archives and Records Administration. Public Law 116 - 283 - William M. (Mac) Thornberry National Defense Authorization Act for Fiscal Year 2021. U.S. Government Publishing Office, 31 Dec. 2020, https://www.govinfo.gov/app/details/PLAW-116publ283.

[24] National Institute of Standards and Technology, AI Risk Management Framework, 2024, www.nist.gov/itl/ai-risk-management-framework.

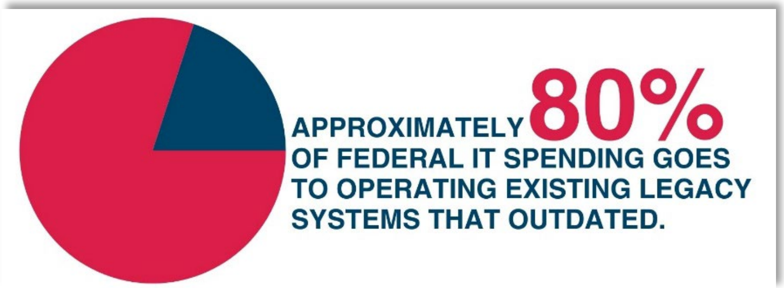
[25] D O'Toole, K., C. Turbes, and A. Freeman, Data Policy in the Age of AI: A guide to using Data for Artificial Intelligence, Data Foundation, 28 Aug. 2024, https://datafoundation.org/news/ai/301/301-Data-Policy-in-the-Age-of-AI-A-guide-to-using-Data-for-Artificial-Intelligence.

Several federal laws already provide the foundation for effective data governance policies, including the Foundations for Evidence-based Policymaking Act of 2018 (P.L. 115 – 435)[26] and the Privacy Act of 1974 (P.L. 93-579)[27] (discussed in more detail below). Enabling the responsible use of AI requires removing barriers to developing safe and effective AI systems. Agencies should take steps to remove barriers to responsible use of AI with the following considerations in mind.

Modernization of Federal Information Systems

Each year, the federal government spends over $100 billion on information technology and cybersecurity.[28] Approximately 80% of this spending goes to operating existing legacy systems that are typically outdated and underpinned by archaic software and hardware components.[29] These legacy systems create security and operational risks and are costly to maintain and remediate when incidents occur.[30]

Addressing this problem and modernizing legacy IT will require significant resources. These projects can take several years, require substantial upfront financial investment, and depend on the technical expertise of engineers experienced in both legacy and contemporary technology.



Source: GAO - Information Technology: Agencies Need to Develop and Implement Modernization Plans for Critical Legacy Systems

With appropriate oversight and accountability safeguards in place, the federal government may be able to use AI to take steps toward modernizing legacy federal IT, including by "automating the migration of legacy software to more flexible cloud-based applications or accelerating mainframe application modernization."[31]

[26] Office of the Federal Register, National Archives and Records Administration. Public Law 115 - 435 - Foundations for Evidence-Based Policymaking Act of 2018. U.S. Government Publishing Office, 13 Jan. 2019, https://www.govinfo.gov/app/details/PLAW-115publ435.

[27] Office of the Federal Register, National Archives and Records Administration. Public Law No. 93 - 579 - The Privacy Act of 1974, U.S. Government Publishing Office, 21 Dec. 1974, www.govinfo.gov/content/pkg/STATUTE-88/pdf/STATUTE-88-Pg1896.pdf.

[28] U.S. Government Accountability Office., GAO-21-524T, "Agencies Need to Develop and Implement Modernization Plans for Critical Legacy Systems", U.S. Government Accountability Office, 27 April 2021, https://www.gao.gov/products/gao-21-524t.

[29] Id.

[30] Id.

[31] Florian Breger and Cristina Caballe Fuguet, "What can AI and generative AI do for governments", IBM, 2024. https://www.ibm.com/blog/what-can-ai-and-generative-ai-do-for-governments/

Further, AI may be used to improve data processing and extraction and to create documentation related to modernization efforts. For example, Google Cloud and Amazon Web Services have both released AI tools that can transform unstructured data from documents or tables into structured data. These AI tools allow organizations to use their unstructured data in new ways or automate their cumbersome manual processes for cleaning data.

Legacy IT systems and unmanaged data in federal agencies are also significant barriers to more rapid adoption of and realization of the benefits of modern AI applications. Investments in AI system technology and reforms to accommodate AI adoption in the public sector need to consider corresponding IT resources and legacy IT modernization projects.

In the Modernizing Government Technology Act of 2017,[32] Congress created the Technology Modernization Fund, allowing agencies to apply for technology modernization funding assistance outside the annual appropriations process.

In July 2024, the GSA, which administers the fund, announced it was partnering with OMB to "harness AI's potential and mitigate its risks in line with" Executive Order 14110 by investing in innovative projects and modernizing outdated, legacy systems.[33]

Improving Federal Cybersecurity

As information communications technology systems become increasingly complex and interconnected, they collect and use greater amounts of data. This compels organizations to interact with a growing number of external systems and users, which adds to the challenges facing federal cybersecurity teams. Security teams must address an expanding attack surface, complex infrastructure, complicated permissions regimes, and growing data storage and access requirements.

As sophisticated attackers seek to use AI offensively to exploit complex systems, security teams must use AI defensively to improve cybersecurity resiliency. The interest in using AI to improve cybersecurity is underscored by projections that the AI cybersecurity market will reach $60.6 billion by 2028.[34]

[32] Office of the Federal Register, National Archives and Records Administration. Public Law 115 - 91 - National Defense Authorization Act for Fiscal Year 2018. U.S. Government Publishing Office, 11 Dec. 2017, https://www.govinfo.gov/app/details/PLAW-115publ91

[33] Laurence Bafundo, "How TMF is Helping Agencies Harness Artificial Intelligence.", General Services Administration, 10 July 2024, https://www.gsa.gov/blog/2024/07/10/how-tmf-is-helping-agencies-harness-artificial-intelligence.

[34] FN Media Group LLC. How Artificial Intelligence (AI) in Cybersecurity is Generating a Billion-Dollar Revenue Opportunity for Tech Industry. GlobeNewswire, 16 Sept. 2024 https://www.globenewswire.com/news-release/2024/09/16/2946676/0/en/How-Artificial-Intelligence-AI-In-Cybersecurity-is-Generating-a-Billion-Dollar-Revenue-Opportunity-for-Tech-Industry.html.

AI already has a growing role in identifying, mitigating, and responding to threat actors and cybersecurity incidents.[35] This is largely due to AI's ability to rapidly process large datasets, detect subtle patterns, and adapt to new threats, resulting in "a powerful level of efficiency and continuous learning that complements human capabilities" and acts as a force multiplier.[36]

Executive Order 14110 directed DHS to take several actions to improve security, resilience, and incident response to AI-related cybersecurity threats to critical infrastructure.[37] It also directed an AI cybersecurity challenge to develop AI tools to find and fix vulnerabilities in critical software.[38]

DHS's Cybersecurity and Infrastructure Security Agency also released a Roadmap for Artificial Intelligence that is "focused at the nexus of AI, cyber defense, and critical infrastructure, sets forth an agency-wide plan to promote the beneficial uses of AI to enhance cybersecurity capabilities; ensure AI systems are protected from cyber-based threats; and deter the malicious use of AI capabilities to threaten the critical infrastructure Americans rely on."[39]

While AI systems offer the prospect of enhanced cybersecurity defense and vulnerability detection, we do not yet know the efficacy of these tools. For example, fuzzing tools have long been used to find vulnerabilities in software,[40] and AI systems have demonstrated the capacity to expand the coverage of these tools.[41] As a result, some cybersecurity challenges may give significant benefit to asymmetric attacks against federal IT systems, while others may support better defense of these systems.[42] More research and testing are necessary to better understand and mitigate these risks while facilitating AI-enabled defenses.

[35] Nikki Henderson, "Feds Weigh Generative AI Use in Cybersecurity, Data Analysis." GovCIO Media, 6 Aug. 2024, https://govciomedia.com/feds-weigh-generative-ai-use-in-cybersecurity-data-analysis/#:~:text=Federal%20officials%20see%20generative%20AI,information%20for%20driving%20better%20decisions.

[36] Lucia Stanham, "The Role of Artificial Intelligence in Cybersecurity." CrowdStrike, 10 May 2024, https://www.crowdstrike.com/cybersecurity-101/artificial-intelligence/.

[37] The White House. Fact Sheet: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence. The White House, 30 Oct. 2023, https://whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/.

[38] Id.

[39] DHS Cybersecurity and Infrastructure Security Agency Releases Roadmap for Artificial Intelligence. U.S. Cybersecurity and Infrastructure Security Agency, 14 Nov. 2023, https://www.cisa.gov/news-events/news/dhs-cybersecurity-and-infrastructure-security-agency-releases-roadmap-artificial-intelligence#:~:text=%E2%80%9COur%20Roadmap%20for%20AI%2C%20focused%20at%20the%20nexus,the%20critical%20infrastructure%20Americans%20rely%20on%20every%20day.%E2%80%9D.

[40] Takanen, A., Demott, J. D., and Miller, C. Fuzzing for Software Security. Artech House Publishers, 2018, https://us.artechhouse.com/Fuzzing-for-Software-Security-Testing-and-Quality-Assurance-Second-Edition-P1930.aspx

[41] Dongge Liu, et al., AI-Powered Fuzzing: Breaking into Bug Hunting with Machine Learning. Google Security Blog, 16 Aug. 2023, https://security.googleblog.com/2023/08/ai-powered-fuzzing-breaking-bug-hunting.html.

[42] Adversarial Threat Landscape for Artificial-Intelligence Systems (ATLAS). MITRE Corporation, https://atlas.mitre.org/matrices/ATLAS.

Securing Government AI Systems

Like all software integrated into federal information systems, AI systems can pose various security risks, including vulnerabilities created by poor systems configuration, malicious data manipulation or poisoning, and increased risks of data breaches and leaks through large-scale automated attacks.

Deploying an AI system can require detailed configuration to ensure it does not inadvertently introduce vulnerabilities into federal information systems.[43] When deploying AI systems, federal agencies must properly define and secure the boundaries between sensitive federal information systems and AI systems. Further, federal agencies must continuously monitor and improve the configurations of the deployment environment to protect the host federal information system. Similarly, when an AI system is connected to a federal information system to support or augment human decision-making, the data ingested by or output from the AI must be resistant to tampering.

While many cybersecurity standards and processes can address cybersecurity vulnerabilities in any information technology system, including AI systems, some may need to be updated to address novel or changing threats to AI systems. NIST's National Vulnerability Data (NVD) is a repository of standards-based vulnerability management data that removes known vulnerabilities from IT systems, including federal systems.[44] While this repository already processes and includes discovered software vulnerabilities resulting from AI systems, additional activities may be required to update definitions and standards for AI-related vulnerabilities. Because more advanced AI systems can be difficult to change once launched and certain types of attacks aimed at AI systems are hard to mitigate against, it may not always be wise to widely disclose discovered vulnerabilities, and sharing practices and norms may need to adapt over time.[45]

Artificial Intelligence and Data Privacy

Artificial intelligence systems have the potential to significantly innovate and improve social services, but these systems are also often fueled by data about individuals. Because these systems are often complex or a part of a complex ecosystem, individuals may not know the potential consequences of how their information is used when interacting with products and services or larger IT ecosystems.[46]

[43] Deploying AI Systems Securely: Best Practices for Deploying Secure and Resilient AI Systems. U.S. Department of Defense, April 2024, https://media.defense.gov/2024/Apr/15/2003439257/-1/-1/0/CSI-DEPLOYING-AI-SYSTEMS-SECURELY.PDF.

[44] National Vulnerability Database (NVD). National Institute of Standards and Technology, 20 Sept. 2022 https://nvd.nist.gov/.

[45] Matt Burgess. "Generative AI's Biggest Security Flaw Is Not Easy to Fix." Wired, 2023, https://www.wired.com/story/generative-ai-prompt-injection-hacking/.

[46] NIST AI Risk Management Framework. National Institute of Standards and Technology, 16 Jan. 2020, https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.01162020.pdf.

The federal government led the way in promoting data privacy in government services through the Privacy Act of 1974,[47] as amended. Ensuring data privacy protections while enabling innovation in government services will be critical to advancing AI-enabled innovation in federal services and beyond. As such, further updates to how the government protects data privacy may be needed.

Some AI systems present larger data privacy challenges to adoption in the federal government than others. Government agencies, especially law enforcement, use and help develop a variety of technologies that enhance and expand surveillance capabilities that can directly impact people's lives. When misused, these technologies can present significant privacy harms for individuals.

For example, police wrongfully arrested a man in Michigan after a facial recognition search misidentified him.[48] To protect against misuse, federal agencies will need to ensure their AI-enabled products have robust controls, including privacy-by-design, collection and use limitations, risk mitigation, and more.[49] Allowing users of government services to understand and control how their data is used will also provide data privacy benefits.

Facilitating access to federal government-controlled data, including sensitive data, through privacy-protective means could open significant data resources for AI development. There are several ongoing efforts to support open government data and controlled access to sensitive data.

First, the federal government is implementing the Foundations for Evidence-based Policymaking Act of 2018,[50] which created several federal activities to support access to non-sensitive federal data and govern controlled access to statistical data.[51] As authorized in the CHIPS and Science Act (P.L. 117 – 167),[52] the National Science Foundation (NSF) also launched the National Secure Data Service pilot project in 2022, which will test a governmentwide effort to strengthen data linkage and data access infrastructure for statistical data.[53]

[47] Supra 27.

[48] New York Times. "A False Arrest in Detroit Spotlights the Risks of Facial Recognition Technology." The New York Times, June 2024, https://www.nytimes.com/2024/06/29/technology/detroit-facial-recognition-false-arrests.html#:~:text=Williams%20said.-,Mr.,she%20was%20eight%20months%20pregnant.

[49] Supra 37

[50] Supra 26.

[51] Id.; see also: "Actions - H.R.1770 - 115th Congress (2017-2018): OPEN Government Data Act." Congress.gov, Library of Congress, 29 March 2017, https://www.congress.gov/bill/115th-congress/house-bill/1770/all-actions.; see also: Confidential Information Protection and Statistical Efficiency Act (CIPSEA). CIO.gov, https://www.cio.gov/handbook/it-laws/cipsea/.

[52] Office of the Federal Register, National Archives and Records Administration. Public Law 117 - 167 - An act making appropriations for Legislative Branch for the fiscal year ending September 30, 2022, and for other purposes. U.S. Government Publishing Office, 8 Aug. 2022, https://www.govinfo.gov/app/details/PLAW-117publ167.

[53] National Secure Data Service Demonstration. National Center for Science and Engineering Statistics, National Science Foundation, https://ncses.nsf.gov/initiatives/national-secure-data-service-demo.

Finally, many agencies have launched several efforts related to research, development, and demonstration of privacy-enhancing technologies, which have the potential to provide access to sensitive data with minimal privacy risk.[54]

Artificial Intelligence and the Federal Workforce

The federal government is the nation's largest employer and is likely to be the biggest user of AI systems. Agencies will require data scientists and other technical experts to develop, purchase, and maintain AI systems. Other employees will need to be upskilled to ensure that AI is used effectively within federal agencies.

To address this challenge, in 2020, the Office of Personnel Management (OPM) was tasked by Congress under the AI in Government Act (P.L. 116-260) with determining the existing state of the federal government's AI workforce and what its capacities need to be in the coming years.[55]

Following this directive, in July 2023, OPM published general technical competencies to assist agencies in "targeting AI skills needed to fill positions to expand AI capabilities governmentwide."[56]

Then, on November 1, 2023, OMB supplemented this work by clarifying the distinction between different skill sets required for different jobs in the AI workforce in a memo with draft guidance, noting:

> *"When identifying and filling workforce needs for AI, agencies should include both technical roles, such as data scientists and engineers, and non-technical roles, such as designers, behavioral scientists, contracting officials, managers, and attorneys, whose contribution and competence with AI are important for successful and responsible AI outcomes. Agencies should provide resources and training to develop such AI talent internally and should also increase AI training offerings for Federal employees, including opportunities that provide Federal employees pathways to AI occupations and that assist employees affected by the application of AI to their work."* [57]

[54] Supra 32.; see also: U.S. and U.K. Launch Innovation Prize Challenges in Privacy-Enhancing Technologies to Tackle Financial Crime and Public Health Emergencies. Office of Science and Technology Policy, The White House, 20 July 2022, https://www.whitehouse.gov/ostp/news-updates/2022/07/20/u-s-and-u-k-launch-innovation-prize-challenges-in-privacy-enhancing-technologies-to-tackle-financial-crime-and-public-health-emergencies/.
[55] Supra 21.
[56] AI in Government Act of 2020 – Artificial Intelligence Competencies. Chief Human Capital Officers Council, 2023, https://chcoc.gov/content/ai-government-act-2020-%E2%80%93-artificial-intelligence-competencies.
[57] Draft Memorandum: Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence. Office of Management and Budget, Nov. 2023, https://www.whitehouse.gov/wp-content/uploads/2023/11/AI-in-Government-Memo-draft-for-public-review.pdf.

To further implement the AI in Government Act, in April 2024, OPM issued a competency model for AI, data, and technology talent “to assist agencies in identifying key skills and competencies needed for AI professionals,” as well as a competency model for civil engineers to ensure “adequate AI expertise and credentials in the Federal Government reflect the increased use of AI in critical infrastructure.”[58,59]

Additionally, after partnering with the Office of Science and Technology Policy (OSTP) and other stakeholder groups to study whether it should “establish an occupational series, or update and improve an existing occupational series, to include positions the primary duties of which relate to artificial intelligence,” as required by the Act, OPM determined that the multidisciplinary nature of AI work warranted the creation of policy guidance providing flexibility to align existing occupational job series with agencies’ varied AI needs.

OPM issued this guidance in April 2024 and continues to support federal agencies at their various stages of implementation as they assess AI uses and programs to inform their workforce needs.[60]

OPM also estimated the number of AI employees in positions related to AI by agency and prepared a 2- and 5-year forecast of the number of federal employees in positions related to AI that each agency will need to employ, as required by the Act.

Upskilling and new educational pathways can be leveraged to equip the federal workforce with AI-ready workers and AI practitioners. These pathways reflect broader workforce considerations discussed in the **Education & Workforce** chapter.

However, some considerations are specific to the federal workforce, including unique federal educational pathways and position or hiring practice requirements. OPM’s ongoing efforts with federal agencies to implement the April 2024 guidance should continue to clarify AI work and roles to support the development of education and hiring pipelines for the federal AI workforce.

[58] Skills-Based Hiring Guidance and Competency Model for Artificial Intelligence Work. Chief Human Capital Officers Council, April 2024, https://www.chcoc.gov/content/skills-based-hiring-guidance-and-competency-model-artificial-intelligence-work.

[59] Artificial Intelligence (AI) Competency Model for Civil Engineering (0810). Chief Human Capital Officers Council, April 2024, https://www.chcoc.gov/content/artificial-intelligence-ai-competency-model-civil-engineering-0810.

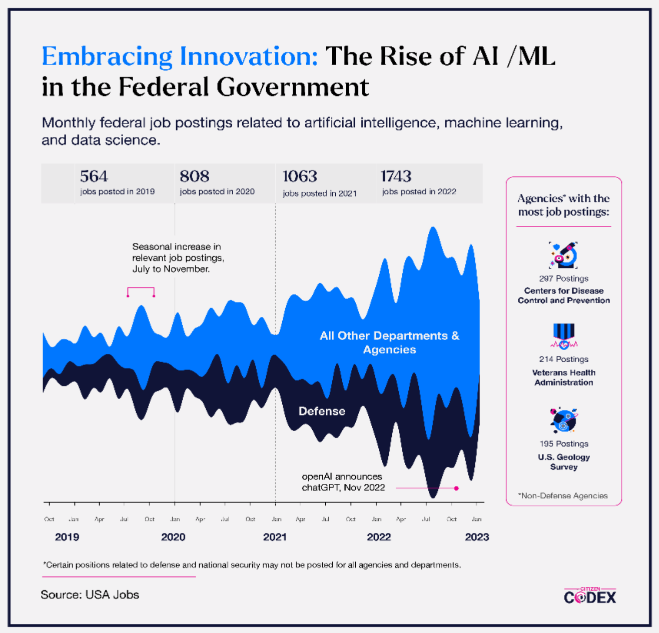
[60] Artificial Intelligence Classification Policy and Talent Acquisition Guidance: AI in Government. Chief Human Capital Officers Council, April 2024, https://www.chcoc.gov/content/artificial-intelligence-classification-policy-and-talent-acquisition-guidance-ai-government.

AI Upskilling and Educational Pathways for Government Work

The Cybersecurity Scholarship for Service program, known as CyberCorps, is an NSF and OPM supported program that boosts the federal cybersecurity workforce by offering scholarships to students in relevant degree programs in exchange for performing a tour of federal service in positions related to cybersecurity.[61]

The CHIPS and Science Act expanded the CyberCorps program to include AI-related degrees supporting cybersecurity missions.[62] These may include the use of AI for cybersecurity and the security of AI systems.

The knowledge, skills, and abilities relevant to CyberCorps programs and graduates are defined by the National Initiative for Cybersecurity Education (NICE) Workforce Framework for Cybersecurity, a framework developed by NIST to create a common lexicon to describe cybersecurity work and workers.[63]



Source: Citizen Codex – AI in the Federal Workforce

The Act also directed NSF to submit to Congress a report and implementation plan to develop an AI scholarship-for-service program to augment the federal AI workforce.

The report was released in May 2024.[64] A key challenge to developing a similar program to educate future federal AI workers is the lack of a foundational framework for AI skills similar to the NICE Framework. These efforts, and the subsequent AI-trained workforce, will be crucial to maintaining U.S. economic competitiveness and national security.

Congress has also enacted legislation requiring federal acquisition professionals to receive AI training to make informed, strategic purchasing decisions on behalf of their agencies and taxpayers. While the government may need new workforce entrants with specialized skills, many existing federal employees will have opportunities to augment

[61] CyberCorps®: Scholarship for Service (SFS). U.S. Office of Personnel Management, https://sfs.opm.gov/.
[62] Supra 52.
[63] Rodney Petersen, et al., National Initiative for Cybersecurity Education (NICE) Workforce Framework for Cybersecurity (SP 800-181 Revision 1). National Institute of Standards and Technology, November 2020, https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-181r1.pdf.
[64] Artificial Intelligence Scholarship for Service Initiative: Need, Feasibility, and Implementation. National Science Foundation, 2024, https://www.nsf.gov/edu/Pubs/2024SFSAIReport.pdf.

their skills by incorporating AI into their portfolio through training, reskilling, and upskilling efforts.

Hiring for Government AI Workers

Industry positions in AI have significantly different hiring pathways from those in the federal government. For example, while the industry may shift its hiring approaches and incentives to temporarily surge its hiring of AI workers, similar flexibilities in hiring for federal agencies require approval. These flexibilities might include special rates, critical pay, waivers of recruitment, relocation, retention, or incentive payment limits, education requirements, and altered position titles or descriptions to attract the right talent.

OPM has provided agencies with a host of tools to more flexibly recruit, hire, and retain candidates with valuable AI-related skill sets. For example, in May 2022, OPM released guidance encouraging agencies to focus on skills-based hiring, which will broaden hiring pathways for all Americans by evaluating candidates based on their skills rather than strictly relying on proxies for skills, such as college degrees.[65]

In 2024, OPM also released memos supporting pay flexibility, incentive pay, and leave and workforce flexibility programs for AI-enabling employees and issued guidance on skills-based hiring and competencies for AI work.[66] Continuing to respond to federal agency challenges in hiring AI practitioners is critical to maintaining a robust federal AI workforce.

While OPM developed and released guidance for agencies to more effectively hire AI and AI-enabling technologists into new roles, existing technologist hiring pathways quickly pivoted to support the AI Talent Surge launched in Executive Order 14110.

The U.S. Digital Service (USDS), a team within the Executive Office of the President consisting of over 200 senior-level technologists to help agencies build critical public-facing government services, has received over 3,000 new applications since the surge was announced and has hired over 30 AI and AI-enabling experts.[67]

Additionally, the GSA's Presidential Innovation Fellows (PIF) and U.S. Digital Corps (USDC) programs, which recruit and place senior and early career tech talent at federal agencies, received over 2,500 applications for AI and AI-enabling roles. In 2024, these programs collectively hired over 50 fellows focused on AI.[68]

[65] Office of Personnel Management, “Press Release: OPM Releases Skills-Based Hiring Guidance” 19 May 2022, https://www.opm.gov/news/releases/2022/05/release-opm-releases-skills-based-hiring-guidance/.

[66] Office of Personnel Management, “Press Release: ICYIMI: OPM Highlights Key Actions Supporting AI Talent Surge to Recruit and Hire AI Professionals”, https://www.opm.gov/news/releases/2024/05/opm-highlights-key-actions-supporting-ai-talent-surge-to-recruit-and-hire-ai-professionals/.

[67] AI and Tech Talent Task Force, “Increasing AI Capacity Across the Federal Government” April 2024, https://ai.gov/wp-content/uploads/2024/04/AI-Talent-Surge-Progress-Report.pdf.

[68] Id.; see also: General Services Administration, “Presidential Innovation Fellows launches first cohort exclusively focused on Artificial Intelligence” 17 June 2024, https://www.gsa.gov/about-us/newsroom/news-releases/presidential-innovation-fellows-launches-first-cohort-focused-exclusively-on-artificial-intelligence-06172024.; see also: General

Key Findings

The federal government should utilize core principles and avoid conflicting with existing laws.

Policies governing the federal government's use of AI should be consistent with a centralized set of core principles to ensure harmonization across agencies regarding how to approach and use AI. Guidance to implement these principles should be built upon existing law and established policy requirements for data and information records, federal information technology systems and cybersecurity, procurement and acquisition, workforce management, and performance and accountability. However, agencies should have the flexibility to set policies governing their use of AI that best meet their needs, consistent with the core principles. AI governance should also be commensurate with the complexity and risk profile of the AI system or use case. For example, simple process automation and workflow tools should not require the same level of system governance controls as large language models with access to sensitive public sector data sets.

Additionally, policies governing agency use of AI should provide holistic, operations-focused guidance spanning the AI lifecycle to enable efficient agency implementation. AI systems can have multiple applications for mission- or programmatic-specific use cases. Government-wide policies governing agency use of AI should be designed in the context of existing federal information policy requirements governing federal agency data and IT systems, such as established privacy and cybersecurity policies. Agencies should determine what restrictions for specific AI use cases are necessary, consistent with applicable and existing legal and regulatory requirements, and appropriate levels of risk consideration.

The federal government should be wary of algorithm-based decision-making.

Policymakers should be cautious of "algorithmic-based" decision-making taking hold within the federal government. Instead, the government should pursue "algorithmic-informed" decision-making supporting missions and programs. Algorithmic-informed decision-making entails proper governance of AI systems and a deliberate policy design that accounts for the inherent limitations of AI systems within certain use cases. Congress and agencies should consider the level of human involvement required for algorithmic-based decision-making across use cases.

Services Administration, "GSA announces new cohort of U.S. Digital Corps fellows" 13 August 2024, https://www.gsa.gov/about-us/newsroom/news-releases/gsa-announces-new-cohort-of-us-digital-corps-fellows-08132024.

The federal government should provide public notification of AI's role in governmental functions.

As AI is implemented to support federal agency workflows and inform decision-making, it is critical that individuals and entities that are substantively and meaningfully affected by an agency decision or determination are provided the proper notifications of AI involvement and appropriate recourse for appeals and human review.

Agencies should pay attention to the foundations of AI systems.

Congress and federal agencies pushing to adopt AI technologies should be mindful of cybersecurity, privacy, and data and IT infrastructure needs when adopting AI technologies. Adopting the fundamentals will be critical to ensuring the responsible adoption and use of federal AI technologies.

Roles and associated AI knowledge and skills are unclear and highly varied across the federal workforce.

Understanding the AI roles needed in the federal workforce will require a standard taxonomy or workforce framework in AI. Currently, AI-related roles are difficult to track due to their evolving nature and non-standardized definitions. This also leads to challenges recruiting and hiring AI practitioners with skillsets well-aligned to relevant positions. Defining the knowledge and skills needed for AI-focused roles in the federal government is critical. A standard taxonomy or workforce framework can enable better alignment of AI training programs to workforce needs, more aligned hiring pathways for the federal government, and improved analytics on the federal government's supply and demand of AI skills.

Skills-based hiring is critical for filling the demand for AI talent in the federal workforce.

Fulfilling the AI training needs of the American workforce will require developing and establishing a pipeline that meets public sector needs by pursuing an “all of the above” approach. Training and credentialing in technical fields like cybersecurity are increasingly done via nontraditional routes like boot camps and certificate programs while relying less on traditional academic degrees. Similarly, many AI work roles will likely be filled by those trained or upskilled via non-traditional education pathways. Candidates from non-traditional education backgrounds can fulfill increasing demands for AI work, and it will be critical to ensure that the correct hiring pathways and policies—such as OPM's May 2022 guidance on skills-based hiring—are in place to enable this.[69]

[69] Supra 65.

Recommendations

Recommendation: Take an information and systems-level approach to the use of AI in the federal government.

In developing new laws for government use of AI, Congress should consider existing federal laws pertaining to federal information policy, information security and cybersecurity, public sector data management and privacy, and procurement.

Existing laws and policies may need to be adopted to address new concerns raised by the federal government's use of AI. AI policy that does not engage with existing information system level requirements across these policy domains will exist outside the current management structure for federal information systems and could lead to confusion, inefficiency, undue administrative or industry burden, and ultimately overlapping or competing legal or policy requirements.

For instance, legislation governing federal agency use of AI should build upon existing areas of law governing federal information policy and security (Chapter 35, Title 44, U.S. Code) or the acquisition of information technology (Chapter 113, Title 40, U.S. Code). Where necessary, repetitive or conflicting requirements and definitions should be repealed or harmonized.

Further, when approaching policy decisions about the collection of private-sector information or the handling of the public's data, such policy should be designed in the context of established law such as the 1967 Freedom of Information Act (FOIA) (5 U.S.C. 552) or the Privacy Act of 1974 (5 U.S.C. 552a) respectively.

Lastly, this policy approach should be considered for properly balancing and coordinating the emerging role of Chief AI Officers with existing federal agency Chief Information Officers, Chief Data Officers, Chief Acquisition Officers, Chief Privacy Officers, etc., who maintain well-established roles over their respective policy domains.

Recommendation: Support flexible governance of AI.

AI will continue to evolve quickly. Guidance pertaining to government use of AI must be flexible enough to adjust to new advancements. All new laws and policies should adhere to core principles and prioritize providing resources that support agencies' ongoing ability to implement evolving standards efficiently, establishing clear authorities across the government and within agencies to make decisions about future AI use, and fostering communication between relevant decision-makers to ensure harmonization where practicable.

Recommendation: Reduce administrative burdens and bureaucracy using AI.

The federal government, like private industry, can achieve enormous efficiency gains by leveraging AI to modernize legacy IT systems and outdated processes. Government functions rich in well-structured, high-quality data will likely be among the first to achieve the productivity gains promised by adopting AI. Such functional domains could include geospatial analysis, financial management and accounting, regulatory compliance, and procurement workflows—all domains with established and controlled workflows and standards for related data management.

Congress should encourage agencies to prioritize the adoption of AI within functional domains that have the requisite open and structured data (i.e., accessible, explainable, controlled, and interpretable data) to make the application of AI most successful. Further, Congress should provide agencies with sufficient resources to invest in data management strategies consistent with applicable laws that allow more functions to have the data necessary to facilitate appropriate and accurate use of AI systems.

Recommendation: Require that agencies provide notification of AI's role in governmental functions.

In considering legislation governing AI use in government, Congress should require that agency policies include plain language notification and appeals processes for individuals or entities impacted by federal agency determinations that were substantively and meaningfully augmented by an AI system. Such policies could be established in conformity with existing statutory requirements regarding the protection of and handling of agency controlled Personally Identifiable Information (PII) records (Privacy Act of 1974 (5 U.S.C. 552a)) and require recourse to pursue alternative, independent human review of agency determinations across federal financial assistance and public benefits adjudication, regulatory enforcement and analysis, and public engagement and service delivery.

Recommendation: Facilitate and adopt AI standards for federal government use.

Congress should provide resources for federal agencies to engage in and support international standards development for AI-related standards related to their missions by participating in those standards activities and adopting consensus standards where practicable, as described in NTTAA and OMB Circular A-119. NIST should continue to ensure the standards it promulgates for federal systems align with international standards.

Recommendation: Support NIST in developing guidelines for federal AI systems.

Congress should support NIST in developing additional voluntary guidance and resources for federal agencies adopting AI systems and promulgating AI-related standards consistent with its authority to develop and apply the Federal Information Processing Standards (FIPS) for federal computing systems. For example, NIST should consider developing a risk profile of the AI risk management framework for specific federal systems. Such standards should follow existing federal processes and align with international standards to the furthest extent practicable.

Recommendation: Improve cybersecurity of federal systems, including federal AI systems.

Cybersecurity will continue to be a critical policy area for federal information systems as the government adopts AI systems. Federal agencies should explore bug bounty programs[70] that enable white hat hackers to find and report vulnerabilities in federal systems, including AI systems. In some cases, AI may even help to improve federal security. Congress should examine legislation requiring OMB to issue guidance on agencies' use of AI to improve the cybersecurity of information systems.

Recommendation: Encourage data governance strategies that support AI development.

Access to federal data will enable AI innovation and facilitate continued U.S. leadership in AI. To accomplish this, agencies' policies must be consistent with applicable and existing legal and regulatory requirements. Congress should provide continued support for public access to federal data, including statistical data. Congress should investigate opportunities to support OMB in implementing the Evidence-Based Policymaking Act of 2018 and expand access to data for AI development.

For example, Congress could explore how to support federal agencies in making AI-ready data repositories, ensuring standardized documentation and formatting, and following applicable laws. Additionally, depending on the results of the National Secure Data Service pilot project authorized by the CHIPS and Science Act, Congress should explore legislation to support the full implementation of the Service.

Finally, Congress should explore legislation to support developing and demonstrating privacy-enhancing technologies across different use cases.

[70] A method of compensating individuals for reporting software errors, flaws, or faults ("bugs") that might allow for security exploitation or vulnerabilities. See: Kim Schaffer, et al., "Recommendation for Federal Vulnerability Disclosure Guidelines." National Institute of Standards and Technology, May 2023, https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-216.pdf.

Recommendation: Congress and the government must understand the federal government's AI workforce needs.

As agencies work to implement OPM's Artificial Intelligence Classification Policy and Talent Acquisition Guidance under the *AI in Government Act*, Congress should remain apprised of their progress and ensure that OPM is aware of and responding to agency AI workforce needs.[71]

To improve the understanding of skills and competencies for cybersecurity roles, Congress passed the *Cybersecurity Enhancements Act of 2014* (P.L. 113-274) and subsequently the *HACKED Act of 2020* (P.L.116-283, included as part of the William M. Thornberry National Defense Authorization Act of Fiscal Year 2021). Among other cybersecurity education-related activities, these bills directed NIST to create and update a workforce framework for cybersecurity that helped to create standardized work categories, roles, tasks, and competency areas.[72] The Science, Space, and Technology Committee should explore legislation to improve the understanding and standardization of AI-related roles, especially those focused on AI governance, including supporting career pathways in the federal government.

Recommendation: Support different pathways into federal service for AI talent.

Congress should investigate ways to incentivize AI workers to join the federal workforce, including by supporting the CyberCorps Scholarship for Service program, creating an AI Scholarship for Service program, and continuing to support the development of AI roles in federal agencies.

[71] Supra 60.

[72] National Initiative for Cybersecurity Careers and Studies (NICCS): NICE Framework. Cybersecurity and Infrastructure Security Agency, March 2024, https://niccs.cisa.gov/workforce-development/nice-framework.

FEDERAL PREEMPTION OF STATE LAW

Background

Preemption influences the distribution of powers between the federal government and the states. Federal-state preemption originates in the Supremacy Clause of the United States Constitution: Article VI, Clause 2.[1] The Supremacy Clause establishes that the Constitution, federal laws made pursuant to it, and treaties made under its authority are collectively the supreme law of the land. Consequently, federal law takes precedence over any conflicting state laws.

Preemption of state AI laws by federal legislation is a tool that Congress could use to accomplish various objectives. However, federal preemption presents complex legal and policy issues that should be considered.

Legal Issues in Preemption

Since preemption delineates the boundaries at which federal authority supersedes state legislation, legal analysis is required to understand the circumstances under which preemption would take effect and the extent to which state laws would be superseded. The requisite legal analysis can involve considerations of the specific text in the federal legislation, the intent of Congress, and the interplay between the applicable federal and state regulatory regimes.

[1] "This Constitution, and the Laws of the United States which shall be made in Pursuance thereof; and all Treaties made, or which shall be made, under the Authority of the United States, shall be the supreme Law of the Land; and the Judges in every State shall be bound thereby, any Thing in the Constitution or Laws of any State to the Contrary notwithstanding." The Constitution of the United States: A Transcription. National Archives, www.archives.gov/founding-docs/constitution-transcript.

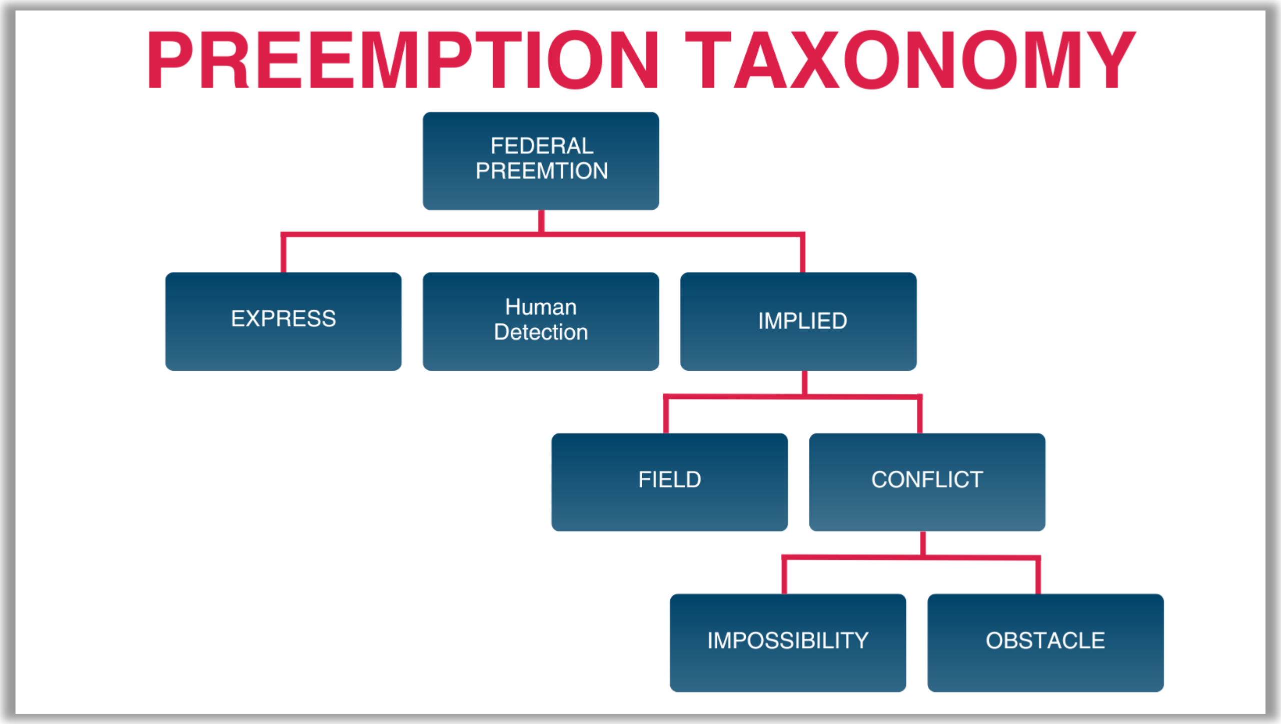
Numerous court decisions have shaped how federal laws can preempt state laws and the circumstances under which preemption applies. Other legal issues surrounding federal-state preemption involve the balance of power between the federal government and state governments, as well as the potential conflicts and inconsistencies that can arise between different levels of government.

Applicability of Preemption

Preemption by federal statute can be either express or implied. When Congress exercises its constitutional authority to regulate in a particular area, it can explicitly state in legislation whether it intends for federal law to preempt or override state laws on the same subject. Moreover, even if legislation does not explicitly state its intent to preempt, federal preemption might be implied by the structure and purpose of the legislation or by court interpretations of the laws.

The supremacy of federal law over state law takes many forms. Under the doctrine of impossibility preemption, if it were impossible to comply with both the federal and state laws, the federal law would preempt the conflicting state law provisions.[2]

Similarly, obstacle preemption preempts any state laws that would serve to frustrate the objectives of federal legislation.



Source: Congressional Research Service – Federal Preemption: A Legal Primer

[2] Jay Sykes and Nicole Vanatko, Congressional Research Service, "Federal Preemption: A Legal Primer," Congressional Research Service, 23 July 2019, https://crsreports.congress.gov/product/pdf/R/R45825/1.

Extent of Preemption

The extent to which an area subject to legislation has been preempted from state law can be unclear. The precise scope of preemption is especially ambiguous when federal legislation has not explicitly spoken on preemption. When no express preemption clause exists in a federal statute, courts must determine if Congress intended to implicitly preempt state law. Determining the scope of federal preemption and resolving conflicts between federal and state laws can be legally complex and time-consuming, requiring interpretation of numerous statutes, regulations, and court decisions.

Moreover, the language chosen by Congress or state legislatures where preemption is explicit can leave the extent of preemption unclear. For example, courts have grappled with interpreting phrases such as state laws “related to” or “covering” certain subjects or preempting state “requirements” versus “laws.”

Finally, there is no universal legal definition of AI.[3] If Congress chooses to preempt state AI laws, then the preempting legislation should precisely define AI in a manner that represents the intended scope of preemption.

> ***“There is no universal definition of AI. If Congress chooses to preempt state AI laws, then the preempting legislation should precisely define AI in a manner that represents the intended scope of preemption.***

Authority to Regulate

One significant issue is determining which level of government has the authority to regulate the specified area represented by legislation. The Constitution enumerates areas for congressional lawmaking, and the Tenth Amendment reserves powers not delegated to the federal government to the states or the people. When both federal and state governments assert their authority over the same issue, it can precipitate jurisdictional conflicts and legal disputes necessitating judicial resolution under applicable legal precedents.

Policy Issues in Preemption

Prominent policy issues surrounding federal-state preemption revolve around Congress's objectives and the specific legislative approach taken to accomplish those objectives.

[3] See Appendix on Definitional Challenges of AI

Uniformity

Federal preemption can promote uniformity in laws and regulations across the country. Uniformity can benefit businesses and individuals operating in multiple states since it establishes one consistent approach to compliance. Absent a uniform approach, different states can adopt different, ambiguous, or even conflicting regulatory requirements that are enforced by disparate state agencies. This disjointed system could create obstacles that businesses across the country struggle to comply with as they are forced to engage with numerous state legislators and agencies. The benefits of uniformity can apply even if no applicable state laws have been promulgated. For example, the possibility of several conflicting state regulations could deter businesses from making investments that could be rendered obsolete by later regulations.

Flexibility and Customization

A disadvantage of federal preemption is its national applicability. State regulations may have been created in response to the different needs and preferences of their populus. Uniformity by federal preemption can limit a state’s ability to tailor laws to its popular political will. Allowing the states to create legislation on a given topic beyond a federal standard can also allow for flexibility in response to changing or unforeseen circumstances.

State Experimentation

When comprehensive information on the relative advantages, drawbacks, and costs of different regulatory approaches is unknown or unclear, it can be acquired by allowing different states to adopt different approaches. These state-level “experiments” could more quickly provide this information than the federal government cycling through a sequence of different regulatory approaches. However, state-level experimentation might be burdensome, costly, or not happen quickly enough.

Relevant Expertise

Information is valuable in creating informed and comprehensive policies. Due to their respective levels of expertise, federal or state governments may be better equipped to understand the ramifications of regulating an area. Federal agencies may possess the experience and resources that states lack. For example, states lack the Department of Defense's expertise on national security issues. Conversely, state governments may be better informed due to long-standing experience or relevant industry presence in the state.

Moratorium During a Learning Period

Another approach to information scarcity is for the federal government to enact a moratorium prohibiting state activity until necessary information is acquired during a learning period. Learning periods are not without risks, and in some cases, they have extended indefinitely to become, in effect, federal preemption of the applicable domain.

Establishing Floors and Ceilings for State Regulation

Federal preemption can establish both floors and ceilings that permit some state regulation. Federal laws can impose a “floor” by requiring minimum standards and baseline protections that states must meet while permitting states to choose to exceed these minimums. This ensures that all states meet basic standards while allowing flexibility in adopting more stringent regulations.

An example of legislation that imposes a floor is the Health Insurance Portability and Accountability Act (HIPAA).[4] Federal legislation can likewise impose a “ceiling” by preventing states from imposing any stricter requirements than specified by federal law. An example of legislation that implements a regulatory ceiling is the E-SIGN Act,[5] which imposed a standard for all digital signatures, or Europe’s General Data Protection Regulation (GDPR).[6]

While legislation like the E-SIGN Act sets forth Congress’ vision of how an area should be regulated, federal legislation can impose a ceiling even if Congress does not regulate that area. Federal legislation could merely preempt specified types of state regulation without providing any accompanying regulation (e.g., during a learning period moratorium).

Authority to Enforce

Federal laws that preempt states from creating their own laws may still allow those states’ attorneys general (AG) to enforce the provisions in the federal law. For example, the Child Online Privacy Protection Act (COPPA) preempts state and local governments from passing laws that create additional liability for activities regulated under the Act.

However, COPPA enables state AGs to seek damages or other relief under COPPA rules in their jurisdictions.[7]

Federal Preemption of Internet Technology as an Analog

It may be helpful to consider the history of federal preemption of another significant technology: the internet. The Federal Communications Commission (FCC) has attempted to preempt state broadband laws under its authority over interstate communications. However, the precise boundaries of this preemption power are still actively litigated.

[4] Office of the Federal Register, National Archives and Records Administration. Public Law 104 - 191 - Health Insurance Portability and Accountability Act of 1996. U.S. Government Printing Office, 20 Aug. 1996, https://www.govinfo.gov/app/details/PLAW-104publ191.

[5] U.S. House of Representatives. U.S. Code Title 15: Commerce and Trade, Chapter 96—Protection of Intellectual Property Rights, Govinfo, uscode.house.gov/view.xhtml?path=/prelim@title15/chapter96&edition=prelim.

[6] "General Data Protection Regulation Information Portal." GDPR-Info.eu, https://gdpr-info.eu/.

[7] U.S. House of Representatives. U.S. Code Title 15: Commerce and Trade, Chapter 91: Privacy Protection. GovInfo, https://uscode.house.gov/view.xhtml?path=/prelim@title15/chapter91&edition=prelim.

The FCC has grappled with the scope of its authority to preempt state laws regarding broadband internet access services based on how such services are classified under federal law.

When the services are classified as a lightly regulated "information service" under Title I of the Communications Act,[8] the FCC has argued this allows federal preemption of state laws that could impose greater regulation. However, uncertainty remains. For example, courts such as the D.C. Circuit in Mozilla v. FCC[9] questioned whether a Title I classification provides a sufficient statutory basis for preempting state net neutrality laws.

In addition, the FCC had attempted to preempt state laws restricting municipal broadband networks. Section 253 of the Telecommunications Act prohibits state laws that may restrict any entity's ability to provide "telecommunications services." In Nixon v. Missouri Municipal League,[10] the Supreme Court narrowly interpreted Section 253 as not preempting state laws restricting municipalities from providing broadband services directly to consumers.

Finally, the FCC has sought to accelerate the deployment of 5G wireless infrastructure by preempting state and local regulations viewed as prohibitive. In 2018, the FCC restricted the fees localities could charge for small cell deployments and imposed "shot clocks" for approvals, all of which preempted conflicting state laws under its statutory authority over interstate services.[11, 12]

[8] 47 U.S.C. §§1-646. U.S. House of Representatives. U.S. Code Title 47: Communications https://uscode.house.gov/view.xhtml?path=/prelim@title47&edition=prelim
[9] Mozilla Corp. v. FCC, 940 F. 3d 1 (D.C. Cir., 2019).
[10] Nixon v. Missouri Municipal League, 541 U.S. 125 (2004).
[11] Chris Linebaugh, U.S. Congressional Research Service. "Overview of Legal Challenges to the FCC's 5G Order on Small Cell Siting," Feb. 2019, https://crsreports.congress.gov/product/pdf/LSB/LSB10265.
[12] United States v. Vasquez-Alvarez, 9th Cir. (2020).

Key Findings

Federal preemption of state law on AI issues is complex.

Preemption raises many legal and policy issues that should be considered and addressed so that Congress effectively implements its intended policies. The context in which an AI is deployed is critical to its governance. An AI system's functional purpose, how it was developed, how it is deployed, and who interacts with it will all affect the rules and regulations that governments set to minimize harm. A generally applicable foundation model used in one sector may require different regulations than the same systems deployed in another. As such, Congress will need to weigh many different factors as it considers preemption in any law targeting artificial intelligence or related technologies.

Federal preemption has benefits and drawbacks.

Federal preemption of state law can bring uniformity and clarity, reduce compliance burdens, and otherwise implement Congress' policy objectives. However, state-level regulation has the advantages of flexibility, customization to different state populations, preservation of state authority, and experimentation that provides information relevant to policy choices.

Preemption can allow state action subject to floors or ceilings.

Federal preemption can allow states to pass laws that either meet federal minimums or that do not exceed federal maximums. This type of preemption can be established with or without a corresponding federal regulatory regime in the same area.

Preemption can be multifaceted.

Preemption of state AI regulation can be extremely multifaceted. For example, the federal government could preempt some, but not other, types of state regulation of a domain. Likewise, the federal government could explicitly permit some, but not other, types of state regulation of a domain.

Definitions must be fit for purpose.

AI has no universal definition and is occasionally seen as a general-purpose category of technology present in many sectors. Defining covered "artificial intelligence" too broadly or too narrowly could either exclude high-risk systems from regulation or accidently sweep in commonplace technologies, such as spreadsheets and spellcheckers. If Congress chooses to preempt state AI laws, the preempting legislation should precisely define AI to represent the intended scope of preemption.

Recommendations

Recommendation: Study applicable AI regulations across sectors.

To better understand the effects of law on this general-purpose technology, Congress should commission a study to analyze the applicable federal and state regulations and laws that affect the development and use of AI systems across sectors. Such a study should analyze which existing laws and legislative and administrative policies are technology-neutral but cover AI systems. Further, such a study could help policymakers better understand existing regulations and preemptive provisions.

DATA PRIVACY

Background

As AI systems amass and analyze vast amounts of data, there are increasing risks of private information being accessed without authorization. Training algorithms identify patterns within the data and produce a set of instructions or a model that can be used with new data.[1] AI models are often trained on diverse datasets that include text from books, websites, and other digital sources, some of which may contain personal or sensitive information. AI systems can also be deployed in sensitive contexts, including healthcare settings, that rely on sensitive data. When users interact with some AI systems, especially generative AI systems, they can inadvertently reveal private or confidential information stored and processed by the AI. Each of these situations has provoked significant concerns regarding the data privacy challenges associated with AI.

Thoughtful and effective data privacy policies and protections will support consumer confidence in the responsible development and deployment of AI systems. While the House AI Taskforce has endeavored to examine data privacy in the context of AI, further exploration of this issue is warranted. Committees with jurisdiction over data privacy should continue to invest time and resources in examining these problems and proposing solutions for the American people.

Advanced AI Systems Require Increasing Amounts of Data

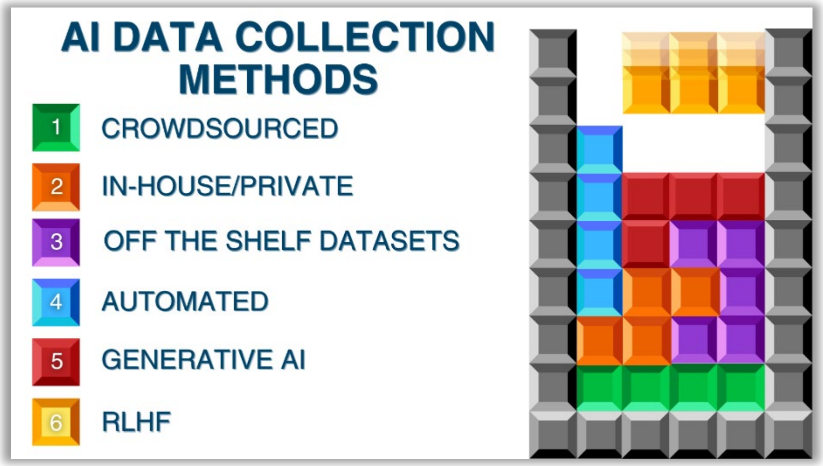
If the data used for training is too small or of poor quality, the model may perform suboptimally. Using large quantities of data from multiple diverse sources generally allows the trained models to perform better.

[1] Tursman, Eleanor, et al. "AI 101 - Aspen Digital." Aspen Digital, 20 June 2023, www.aspendigital.org/report/ai-101/#section2.

According to some research, the performance of algorithms benefits significantly from larger training datasets.[2,3] In the decades since, the amount of digital data that could be used for training algorithms has increased dramatically.[4] To be sure, a growing volume of research suggests smaller datasets may also play a role in improving AI performance.[5]

Data available to train AI models is collected and licensed in various ways. Some companies use a combination of internal and external data.[6] Other firms, such as those deploying large language models (LLMs) and foundation models, mainly rely on data acquired ("scraped") from the internet.

Web scraping is a process by which data is copied from the internet. Some companies package, process, and label the scraped data for sale, while others release open-source data sets.[7] There is a voluntary standard used by many websites to indicate that they should not be scraped, and other companies have added such a stipulation to their terms of service. Unfortunately, these clearly stated requests are often ignored[8] and there are a growing number of disputes and litigation over scraping issues involving AI companies.



Source: AI Multiple Research - Top 6 Data Collection Methods for AI & Machine Learning

[2] Banko, Michele, and Eric Brill. "Scaling to Very Very Large Corpora for Natural Language Disambiguation." ACL Anthology, 2001, aclanthology.org/P01-1005.pdf.

[3] Kaplan, Jared, et al. "Scaling Laws for Neural Language Models." arXiv, 23 Jan. 2020, https://arxiv.org/abs/2001.08361.

[4] Roded, Tal, and Peter Slattery. "What Drives Progress in AI? Trends in Data." FutureTech, 19 March 2024, futuretech.mit.edu/news/what-drives-progress-in-ai-trends-in-data.

[5] Li, Kangming, et al. "Exploiting redundancy in large materials datasets for efficient machine learning with less data." Nature Communications, 10 Nov. 2023, https://www.nature.com/articles/s41467-023-42992-y.

[6] Brown, Sara. "Why External Data Should Be Part of Your Data Strategy." MIT Sloan School , 18 Feb. 2021, mitsloan.mit.edu/ideas-made-to-matter/why-external-data-should-be-part-your-data-strategy.

[7] Newman, Marissa, and Aggi Cantrill. "A High School Teacher's Free Image Database Powers Ai Unicorns." Bloomberg, 24 Apr. 2023, www.bloomberg.com/news/features/2023-04-24/a-high-school-teacher-s-free-image-database-powers-ai-unicorns.

[8] Paul, Katie. "Exclusive: Multiple AI companies bypassing web standard to scrape publisher sites, licensing firm says." Reuters, 21 June 2024, https://www.reuters.com/technology/artificial-intelligence/multiple-ai-companies-bypassing-web-standard-scrape-publisher-sites-licensing-2024-06-21/

Companies are also turning to their own users' data to train AI systems. Google allegedly scraped Google Docs and Gmail for data to train AI tools.[9] Users may also transmit their personal or company data via queries provided to AI models that are hosted or otherwise controlled by a third party, like an AI company. Meta and X have changed their privacy policies to allow for training AI models on the platforms' data.[10,11]

More companies are updating their privacy policies in order to permit the use of user data to train AI models.[12] Meta faces legal challenges in eleven European countries over its plans to use users' personal data to train AI models.[13]

The Federal Trade Commission (FTC) has addressed the matter, stating that it may be unfair or deceptive for a company to adopt more permissive data practices but only inform consumers of this change through an amendment to its privacy policy.[14]

In response to these concerns, some companies are turning to privacy-enhancing technologies, which seek to protect the privacy and confidentiality of data when sharing it. For example, Apple has used a privacy-preserving technology called differential privacy to analyze Apple users without sharing individuals' information.[15]

Similarly, the AI company Anthropic recently partnered with the UK Safety Institute and the PET company OpenMined to test how to utilize secure computation to allow multiple parties to access advanced models and nonpublic data.[16]

The growth in data widely available to AI companies may be reaching a plateau.[17] It is unclear how AI developers and researchers will satisfy the need for additional training data. Some are exploring synthetic data, which is created artificially through computer simulations or algorithms.

[9] Morrison, Sara. "The Tricky Truth about How Generative AI Uses Your Data." Vox, 27 July 2023, www.vox.com/technology/2023/7/27/23808499/ai-openai-google-meta-data-privacy-nope.

[10] Mearian, Lucas. "Meta's Privacy Policy Lets It Use Your Posts to Train Its AI." Computerworld, 21 June 2024, www.computerworld.com/article/2264949/metas-privacy-policy-lets-it-use-your-posts-to-train-its-ai.html.

[11] Perez, Sarah. "Elon Musk's X Is Changing Its Privacy Policy to Allow Third Parties to Train AI on Your Posts." TechCrunch, 17 Oct. 2024, techcrunch.com/2024/10/17/elon-musks-x-is-changing-its-privacy-policy-to-allow-third-parties-to-train-ai-on-your-posts/.

[12] Hays, Kali. "A Long List of Tech Companies Are Rushing to Give Themselves the Right to Use People's Data to Train AI." Business Insider, 13 Sept. 2023, www.businessinsider.com/tech-updated-terms-to-use-customer-data-to-train-ai-2023-9.

[13] Woollacott, Emma. "Meta Faces Legal Complaints Over New AI Training Data Plans." Forbes, 10 June 2024, www.forbes.com/sites/emmawoollacott/2024/06/10/meta-faces-legal-complaints-over-new-ai-training-data-plans/.

[14] Staff in the Office of Technology and The Division of Privacy and Identity Protection. "AI (and Other) Companies: Quietly Changing Your Terms of Service Could Be Unfair or Deceptive." Federal Trade Commission, 13 Feb. 2024, www.ftc.gov/policy/advocacy-research/tech-at-ftc/2024/02/ai-other-companies-quietly-changing-your-terms-service-could-be-unfair-or-deceptive.

[15] "Differential Privacy Overview." Apple, 2 Nov. 2017, images.apple.com/privacy/docs/Differential_Privacy_Overview.pdf.

[16] "Interviewing Andrew Trask on How Language Models Should Store (and Access) Information." Interconnects, 10 Oct. 2024. www.interconnects.ai/p/interviewing-andrew-trask.

[17] Xu, Tammy. "We Could Run out of Data to Train AI Language Programs ." MIT Technology Review, 24 Nov. 2022, www.technologyreview.com/2022/11/24/1063684/we-could-run-out-of-data-to-train-ai-language-programs/.

Synthetic data can be used as an alternative or supplement to real-world data, particularly when real-world data of the appropriate form is unavailable or has already been exhausted.[18] Synthetic data allows for exploring new possibilities because such data can be designed to represent hypothetical situations beyond what existing real-world data represents.[19] Synthetic data also offers privacy-enhancing benefits, given that it does not include information on real individuals. Unfortunately, since it does not truly represent actual measurements, synthetic data may lack the complexity and nuances of real-world data. Accordingly, models trained on synthetic data may be unable to perform well in various real-world scenarios, and an overreliance on synthetic data may lead to technical complications like model collapse.[20]

Privacy Harms From AI

Americans are vulnerable to several privacy harms. The full breadth of privacy harms is difficult to estimate because they are so varied and can encompass different but related concerns. Nevertheless, to clarify the policy issues, the following types of privacy harms are frequently referenced by stakeholders:

- *Physical harms* result in bodily injury or death. For example, a man purchased personal data about Amy Boyer from New Hampshire, including the address of Boyer's employer. The man fatally shot her where she worked.
- *Economic harms* involve monetary losses or other losses of value. Identity thieves steal personal data and use it to conduct fraudulent transactions in victims' names, including opening credit card accounts and accruing debt that damages the victims' credit history.

  Increasingly, these thieves target public school districts and steal the identities of children. The credit records of minors can be exploited for years before the victims even discover it. Celeste Gravatt is one of the thousands of parents who had her children's data stolen as part of a Minneapolis Public Schools cyberattack. She locked their credit accounts but remains worried.
- *Emotional harms* result from emotional distress from information being released about someone without their knowledge or consent. These harms form the basis of many privacy torts, such as intrusion upon seclusion, trespass, and more.

[18] IBM. "What Is Synthetic Data?" IBM, 2024, www.ibm.com/topics/synthetic-data.
[19] Bozzella, Kim. "The Pros And Cons Of Using Synthetic Data For Training AI." Forbes, 20 Nov. 2023, www.forbes.com/councils/forbestechcouncil/2023/11/20/the-pros-and-cons-of-using-synthetic-data-for-training-ai/.
[20] Shumailov, Ilia, et al. "AI models collapse when trained on recursively generated data." Nature, 24 July 2024, https://www.nature.com/articles/s41586-024-07566-y.

- *Reputational harms* involve injuries to an individual's reputation and standing in the community, such as lost business, employment, or social status. For example, Murray Dowey was the target of "sextortion," online blackmail based on the threat of exposing his intimate images. Dowey tragically took his own life earlier this year.[21]

- *Discrimination harms* involve disadvantaging people based on characteristics like sex, race, age, religion, or political affiliation. They can thwart people's ability to obtain jobs, secure insurance, and find housing.

- *Autonomy harms* involve subverting or impairing an individual's autonomy. For example, some bad actors use "dark patterns" or design features used to deceive or manipulate users.

There are many examples of AI systems exacerbating privacy harms. Synthetic content can duplicate someone's likeliness without their consent. Facial recognition systems can enable pervasive tracking of people in public places. Advanced AI systems, such as LLMs, have been found to inadvertently leak personally identifiable information if not properly configured or protected.[22] Further, AI systems have been shown to infer sensitive information about someone,[23] even from legally obtained and deidentified data,[24] in some cases inadvertently revealing personal attributes such as political views or sexual orientation. In one case, a major retailer's system predicted a shopper was pregnant and accidentally revealed that information to her father.[25]

American's Privacy Protections Vary

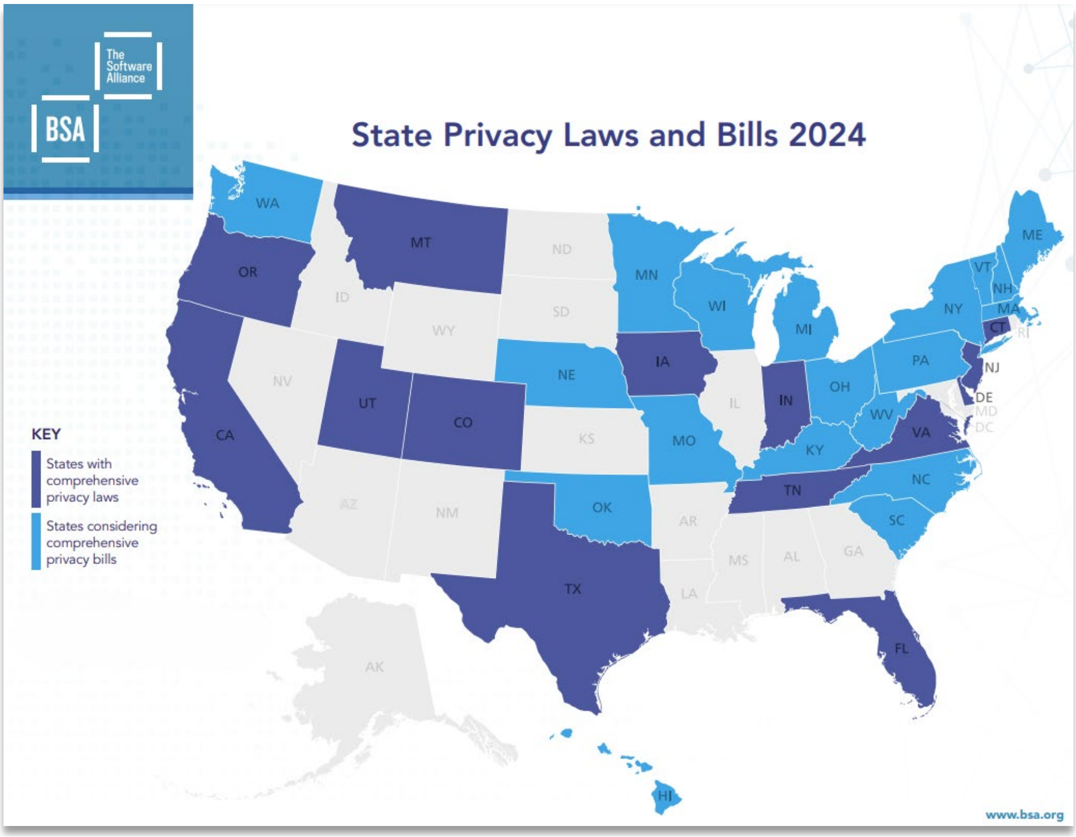
Currently, there is no comprehensive U.S. federal data privacy and security law. However, there are several federal privacy laws focused on various sectors or use cases, such as child privacy or health information. States have also acted. To date, nineteen U.S. states have enacted their own state privacy laws with varying standards.[26]

[21] Chigozie Ohaka, et al,. "Stop terrorizing children with sextortion, say parents." *BBC*, November 2024. https://www.bbc.com/news/articles/cz6jywx37dlo
[22] Yan, Biwei, et al. "On Protecting the Data Privacy of Large Language Models (LLMs): A Survey." arXiv, 8 March 2024, arxiv.org/abs/2403.05156.
[23] Creţu, Ana-Maria, et al. "Interaction data are identifiable even across long periods of time." Nature Communications, 25 Jan. 2022, https://www.nature.com/articles/s41467-021-27714-6
[24] Na, Lingyuan, et al. "Feasibility of Reidentifying Individuals in Large National Physical Activity Data Sets From Which Protected Health Information Has Been Removed With Use of Machine Learning." JAMA Network Open, 21 Dec. 2018, https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2719130
[25] Hill, Kashmir. "How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did." Forbes, 11 Aug. 2022, www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/.
[26] Kibby, C. "U.S. State Privacy Legislation Tracker." Resource Center, iapp, 4 Nov. 2024, iapp.org/resources/article/us-state-privacy-legislation-tracker/. Some states, such as New York and Colorado, have enacted AI-related legislation as well, while others like California are seeking to use their state data privacy and security laws to regulate AI.

One salient example of state action is data breach notification laws; in the absence of a federal standard, each state moved forward to create its own.[27] State laws have created a patchwork of rules and regulations with many drawbacks. Consumers can be confused about the extent of privacy protections, gaps exist in data privacy protections, and businesses can face increased compliance burdens and uncertainty.

Companies that fail to comply with each existing privacy law risk a myriad of lawsuits from state regulators and individuals. As such, private companies conducting business in multiple states must track and monitor changes to state laws, which is often a challenging task for smaller businesses with fewer resources. If businesses are forced to comply with multiple state laws, it can inhibit business expansion, hiring, and the ability to develop and deploy new technologies.

Federal legislation that preempts state data privacy laws has advantages and disadvantages.[28] The complexities of federal preemption are discussed further in the chapter on **Federal Preemption of State Law**.



Source: BSA | The Software Alliance - State Privacy Bills and Laws Map

[27] "Data Breach Notification Laws by State." IT Governance, www.itgovernanceusa.com/data-breach-notification-laws.
[28] Mulligan, Stephen P., and Chris D. Linebaugh. "Data Protection and Privacy Law: An Introduction." Congressional Research Service, 12 Oct. 2022, crsreports.congress.gov/product/pdf/IF/IF11207.

Key Findings

AI has the potential to exacerbate privacy harms.

AI is inherently linked to issues of data: how to obtain large amounts of data, how to analyze data for patterns, and how to use those patterns to make predictions. Developers and users of AI can intentionally or unintentionally cause or exacerbate data privacy harms related to each of these facets.

Americans have limited recourse for many privacy harms.

Many businesses are generally unrestricted in the types of sensitive information they can collect from Americans, how they can use that information, who they can transfer or sell it to, and how long they can retain it. While state laws have started to address these concerns, many Americans have limited rights or recourse when faced with encroachments on their privacy.

Federal privacy laws could potentially augment state laws.

Federalism has been a controversial issue for federal data privacy laws because of its complexity. Congress could adopt a comprehensive system for data protection by expressly preempting state laws related to data privacy. Alternatively, Congress could preserve state laws in some ways but preempt them in others. Another option is for Congress to pass a law that preempts state legislation but still enables states to enforce the federal standard. Congress also has the option to leave state schemes intact in conjunction with a federal scheme.

Recommendations

Recommendation: Explore mechanisms to promote access to data in privacy-enhanced ways.

Access to privacy-enhanced data will continue to be critical for AI development. The government can play a key role in facilitating access to representative data sets in privacy-enhanced ways, whether through facilitating the development of public datasets or the research, development, and demonstration of privacy-enhancing technologies or synthetic data. Congress can also support partnerships to improve the design of AI systems that consider privacy-by-design and utilize new privacy-enhancing technologies and techniques.

Recommendation: Ensure privacy laws are generally applicable and technology-neutral.

Congress should ensure that privacy laws in the United States are technology-neutral and can address many of the most salient privacy concerns with respect to the training and use of advanced AI systems. Congress should also ensure that general protections are flexible to meet changing concerns and technology and do not inadvertently stymie AI development.

NATIONAL SECURITY

Background

Like any major dual-use technology, AI has the potential to both bolster and undermine national security. This underscores its significance in U.S. defense strategy. Currently, the U.S. national security ecosystem is both using and developing AI, but a significant proportion of research and development related to AI is occurring outside of government activities.

The Department of Defense (DOD) has been involved with AI since its earliest days, and its research underpins the flourishing AI ecosystem we see today. DOD is also on the leading edge of operationalizing and deploying AI. As early as 1991, DOD used AI to solve logistical challenges and reduce expenses.[1] Today, the Pentagon is experimenting with AI in logistics, business operations, and vehicle autonomy, as well as conducting research and cooperative efforts with partners and allies across the globe.

In 2018, Congress mandated the establishment of the National Security Commission on Artificial Intelligence (NSCAI)[2] in the National Defense Authorization Act (NDAA) for Fiscal Year 2019.[3] The Commission's work resulted in 543 recommendations spanning actions across the Executive Branch and Legislative Branch, organized along several lines of effort, including defense against AI-enabled threats, risk management for AI-enabled and autonomous systems, talent management, and oversight and confidence-building.[4]

[1] S. R. Hedberg, "DART: revolutionizing logistics planning," in IEEE Intelligent Systems, vol. 17, no. 3, pp. 81-83, April 2005, https://ieeexplore.ieee.org/document/05635.

[2] National Security Commission on Artificial Intelligence, "Final Report", National Security Commission on Artificial Intelligence 2021, https://reports.nscai.gov/final-report/.

[3] U.S. Congress. Public Law No. 115-232: John S. McCain National Defense Authorization Act for Fiscal Year 2019. GovInfo, 13 Aug. 2018, https://www.govinfo.gov/app/details/PLAW-115publ232.

[4] Supra 2.

Congress implemented more than 100 legislative recommendations of the NSCAI report, including Section 235 of the FY21 NDAA (P.L. 116-283),[5] *Acquisition of Ethically and Responsibly Developed Artificial Intelligence Technology*; Section 236 of the FY21 NDAA (P.L. 116-283), *Steering Committee on Emerging Technology*; and Section 1118 of the FY22 NDAA (P.L. 117-81),[6] *Occupational Series for Digital Career Field*.

Select History of Artificial Intelligence in the Department of Defense

Defense Advanced Research Projects Agency (DARPA)

DARPA has conducted AI research since the 1960s.[7] Its research is fundamental to a wide range of technologies, such as autonomous vehicles and natural language processing (NLP), and forms the basis of commonly used AI-enabled applications as Apple's Siri (the PAL, Personalized Assistant that Leans, program).[8]

In 2018, DARPA launched AI Next, a $2 billion, multi-year campaign focused on automating DOD business practices, improving the robustness and reliability of AI, and enhancing AI security.[9]

DARPA currently has multiple ongoing AI research programs, including the AI Cyber Challenge (AIxCC), a prize challenge conducted in collaboration with multiple AI companies such as Open AI, Anthropic, Google, and Microsoft, as well as with civil society groups such as the Linux Foundation, the Open Source Software Foundation, DEFCON, and Black Hat USA.

AIxCC finalists develop open-source AI that can be used to secure AI from various cyberattacks and improve the use of AI in defending against cyberattacks.[10]

[5] Office of the Federal Register, National Archives and Records Administration. Public Law 116 - 283 - William M. (Mac) Thornberry National Defense Authorization Act for Fiscal Year 2021. U.S. Government Publishing Office, 31 Dec. 2020, https://www.govinfo.gov/app/details/PLAW-116publ283.

[6] Office of the Federal Register, National Archives and Records Administration. Public Law 117 - 81 - National Defense Authorization Act for Fiscal Year 2022. U.S. Government Publishing Office, 26 Dec. 2021, https://www.govinfo.gov/app/details/PLAW-117publ81.

[7] Defense Advanced Research Projects Agency (DARPA). Defense Advanced Research Projects Agency. https://www.darpa.mil.

[8] Id.

[9] Defense Advanced Research Projects Agency (DARPA). "AI Next Campaign." Defense Advanced Research Projects Agency, https://www.darpa.mil/work-with-us/ai-next-campaign.

[10] Defense Advanced Research Projects Agency, "DARPA AI Cyber Challenge Proves Promise of AI-Driven Cybersecurity." Defense Advanced Research Projects Agency, 11 Aug. 2024, https://www.darpa.mil/news-events/2024-08-11.

Responsible AI

DOD issued its first responsible use policy for AI in 2012, and additional Ethical Principles for AI were adopted thereafter.[11] In 2021, DOD released a Responsible AI (RAI) Strategy and Implementation Pathway that details a framework to harness AI in line with deployment guidelines and use standards.[12] This document also outlines how DOD will develop and deploy AI with appropriate levels of human oversight and intervention. The five RAI principles establish that AI should be responsible, equitable, traceable, reliable, and governable.[13]

Project Maven

Often considered DOD's first large-scale use of AI, Project Maven automated the analysis of full-motion video collected by intelligence, surveillance, and reconnaissance platforms.[14] Despite some obstacles in the early years, there has been a robust maturation of Maven-related technologies. At the start of FY 2023, DOD transitioned Project Maven, including the management and responsibility for labeled data, AI algorithms, test & evaluation capabilities, and the platform itself, to the National Geospatial Intelligence Agency (NGA).[15]

Chief Digital and Artificial Intelligence Officer

The Chief Digital & AI Officer (CDAO), established in 2021, serves as DOD's senior official responsible for strengthening and integrating data, AI, and digital solutions.[16] The CDAO assumed responsibility for the Joint AI Center (JAIC), Defense Digital Service (DDS), and the Chief Data Officer (CDO). The office's stated mission is to "accelerate DOD adoption of data, analytics, and artificial intelligence from the boardroom to the battlefield to enable decision advantage."

[11] Department of Defense. "DOD Adopts Ethical Principles for Artificial Intelligence." U.S. Department of Defense, 24 Feb. 2020, https://www.defense.gov/News/Releases/release/article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/.
[12] U.S. Department of Defense. Department of Defense Responsible Artificial Intelligence Strategy and Implementation Pathway. June 2022, https://media.defense.gov/2022/Jun/22/2003022604/-1/-1/0/Department-of-Defense-Responsible-Artificial-Intelligence-Strategy-and-Implementation-Pathway.PDF.
[13] Supra 11.
[14] U.S. Deputy Secretary of Defense Memorandum, "Establishment of an Algorithmic Warfare Cross-Functional Team (Project Maven)," April
2017, https://www.govexec.com/media/gbc/docs/pdfs_edit/establishment_of_the_awcft_project_maven.pdf.
[15] Nathan Strout. "Intelligence Agency Takes Over Project Maven, the Pentagon's Signature AI Scheme." C4ISRNET, 27 April 2022, https://www.c4isrnet.com/intel-geoint/2022/04/27/intelligence-agency-takes-over-project-maven-the-pentagons-signature-ai-scheme/.
[16] U.S. Department of Defense. Artificial Intelligence (AI). https://www.ai.mil.

As of 2023, CDAO's stated priorities were:

- Advancing ADVANA, the Department's big-data analytic and visualization platform.
- Improving data quality and data labeling.
- Driving Artificial Intelligence/Machine Learning (AI/ML) or "Enterprise Scaffolding."
- Enabling a joint All-Domain Command & Control (JADC2) Data Integration Layer. CDAO is responsible for the data fabric necessary to exchange data among Combatant Commands and foreign military partners.
- Planning a Digital Talent Management Pilot for establishing a Defense Digital Corps, a cadre of digital experts assigned to digital positions across DOD and managed as a unified cohort.

The CDAO structures its efforts around five key functions:

- Leading and overseeing DOD's strategy development and policy formulation for data, analytics, and AI.
- Working to break down barriers to the adoption of AI and data services within appropriate DOD institutional processes.
- Creating digital infrastructure and services that support components' development and deployment of data, analytics, AI, and digitally enabled solutions.
- Selectively scaling proven digital and AI-enabled solutions for enterprise and joint use cases.
- Surging digital services for rapid response to crises and emergent challenges.

As of January 2024, the CDAO's total military and civilian combined onboard strength was 177, with authorization for up to 197 civilian full-time equivalents (FTEs). Additionally, total military on-board strength stood at 49, of which 26 were filled by active-duty military billets, with the remainder filled by reservists and National Guard members.

CDAO is Working to Scale AI Across DOD

The CDAO has been working on multiple initiatives to build the foundations for AI across the Department. This includes, most notably, the ADVANA platform, which enhances enterprise analytics, enables common data models, and facilitates natural language discovery.[17] CDAO has also played an important role in the Combined Joint All Domain Command & Control (CJADC2) effort to improve data transport.

[17] Grace Lin. "Meet Advana: How the Department of Defense Solved Its Data Interoperability Challenges." Government Technology Insider, 7 April 2021, https://governmenttechnologyinsider.com/meet-advana-how-the-department-of-defense-solved-its-data-interoperability-challenges/.

Another CDAO effort, Tradewinds, is a suite of services designed to accelerate adoption of AI/ML across DOD. The Tradewinds Solutions Marketplace (launched in November 2022) serves as a marketplace for opportunities to work with DOD on AI/ML, data, and digital projects.

CDAO also recently announced a new initiative called Open DAGIR (Open Data and Applications Government-owned Interoperable Repositories).[18] This multi-vendor ecosystem will enable industry and government to integrate data platforms and development tools/environments while preserving government data ownership and industry IP. These and other ongoing efforts at CDAO remain key to establishing the necessary underpinnings for scaling AI/ML across DOD.

National Security Memorandum on AI

In October 2024, the Biden Administration released a national security memorandum (NSM) on AI.[19] The NSM makes several policy changes, including making competitor nations' operations against domestic AI companies a top intelligence priority, activities to support the evaluation of AI systems, and activities to support the use of AI systems in service of the national security mission in ways that align with U.S. values and human rights.

Artificial Intelligence Research and Development

Department of the Army

The Army leverages and integrates AI/ML to identify, assess, and prioritize threats and facilitate rapid decision-making. The Army intends to integrate AI/ML to enhance intelligence operations, predictive maintenance, talent management, and command and control, among other use cases. The Army's Aided Threat Recognition from Mobile Cooperative and Autonomous Sensors (ATR-MCAS) program designs and develops AI/ML mobility algorithms to allow autonomous ground and air vehicles to perceive the environment, to enable collaborative teaming between vehicles and to aggregate and distribute large amounts of target data during reconnaissance missions.[20]

[18] U.S. Department of Defense. "CDAO Announces New Approach to Scaling Data, Analytics, and AI Capabilities." U.S. Department of Defense, 30 May 2024, https://www.defense.gov/News/Releases/Release/Article/3791829/cdao-announces-new-approach-to-scaling-data-analytics-and-ai-capabilities/.

[19] White house, "Memorandum on Advancing the United States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence." White House. 24 Oct. 2024, https://www.whitehouse.gov/briefing-room/presidential-actions/2024/10/24/memorandum-on-advancing-the-united-states-leadership-in-artificial-intelligence-harnessing-artificial-intelligence-to-fulfill-national-security-objectives-and-fostering-the-safety-security/.

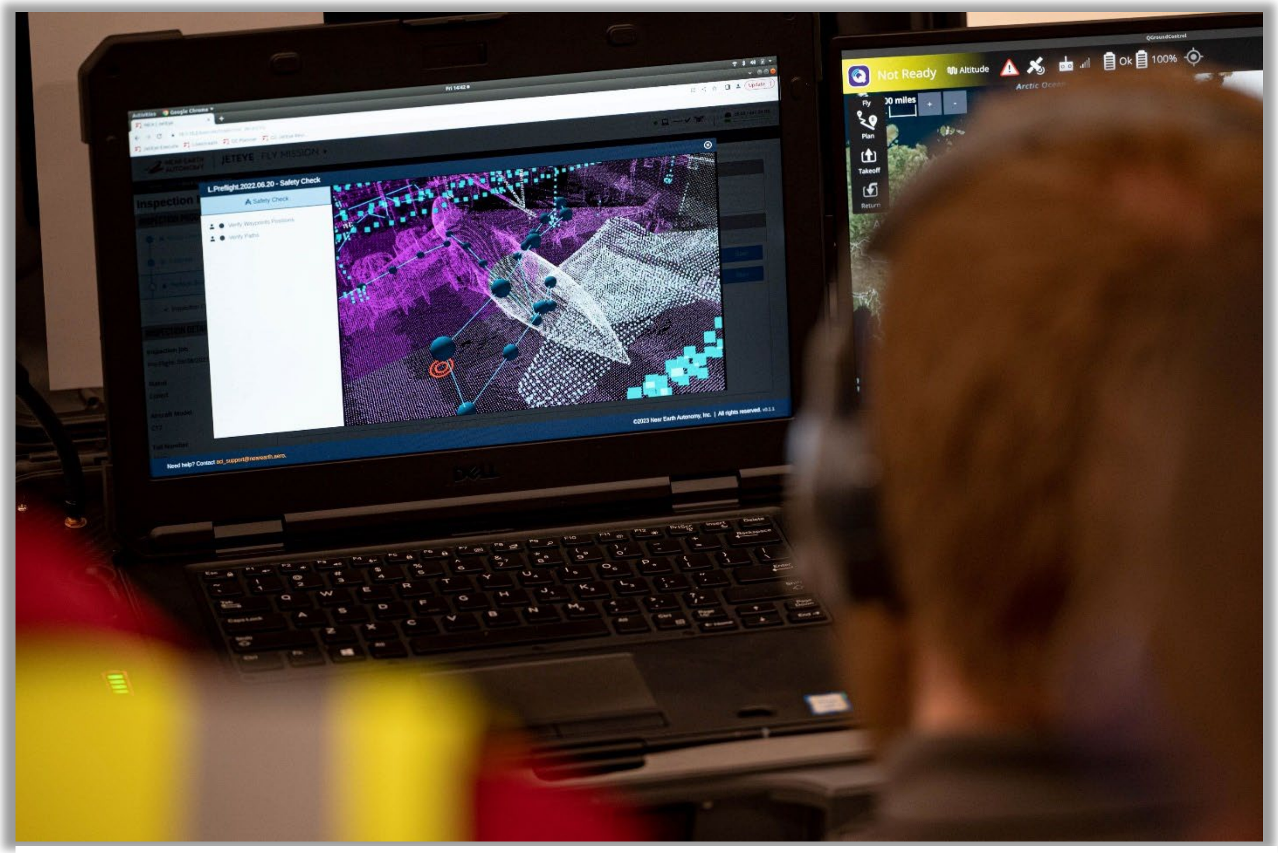
[20] Patrick Ferraris, "Aided Detection on the Future Battlefield." U.S. Army, https://www.army.mil/article/232074/aided_detection_on_the_future_battlefield.

Department of the Navy

The Office of Naval Research (ONR) is integrating AI to optimize mission planning, monitor Navy platforms, and forecast possible enemy courses of action.[21] ONR also incorporates AI to enable intelligent autonomous systems (IAS) for distributed maritime operations. The goal is to demonstrate IAS maneuverability and teaming and to assess kill chains over tactically relevant ranges without relying on vulnerable command and control systems.

The Minerva program uses AI to optimize the assignment and location of kill chains (including sensors, ships, and weapons) to better balance the Navy's offensive and defensive posture.

The U.S. Marine Corps is leveraging AI to develop a common operating picture, enhance situational awareness, and inform and support command decision-making by inferring adversarial intent and recognizing patterns in common intelligence and tactical scenarios.



A ground control station renders a flight path for an autonomous inspection drone at Pittsburgh International Airport Air Reserve Station, Pa., Sept. 8, 2023. *Source: DARPA Aims to Develop AI, Autonomy Applications Warfighters Can Trust*

[21] U.S. Navy, Office of Naval Research. Office of Naval Research. https://www.onr.navy.mil/.

Department of the Air Force

The Air Force's basic research program investigates scientific principles and algorithms that underlie intelligent, human-machine decision-making to enable machine-based future battlespace networks. The Air Force (USAF) and Space Force (USSF) are also seeking to leverage AI/ML to support faster material development and characterization modeling for materials science problems; automate multi-sensor data exploitation, information processing, and data fusion; develop networked collaborative autonomous weapons technology; and support and perform operations in complex adversarial environments. The Fight Tonight program seeks to develop and demonstrate AI-based military planning capabilities to build, assess, and adopt combat power that possesses the speed and scale necessary to achieve a decisive advantage in a peer conflict in highly contested environments.

In June 2024, the Department of the Air Force's Chief Information Officer, in partnership with the Air Force Research Laboratory, announced a series of initiatives to responsibly experiment with Generative AI within the confines of predetermined safeguards. These efforts, collectively known as "NIPRGPT," are designed to build capabilities organic to USAF and USSF and provide a proving ground for private sector partners to demonstrate new technologies and new use cases with military applications.[22]

DARPA

DARPA currently has several initiatives focused on developing and integrating AI for mission areas, including space domain awareness/data fusion, the autonomous maneuver of unmanned systems in military-relevant environments, defense against adversarial AI, and algorithms to detect AI-generated content ("deepfakes"). In addition, DARPA has also coordinated with the Air Force and industry to use AI to fly an F-16 fighter jet autonomously.

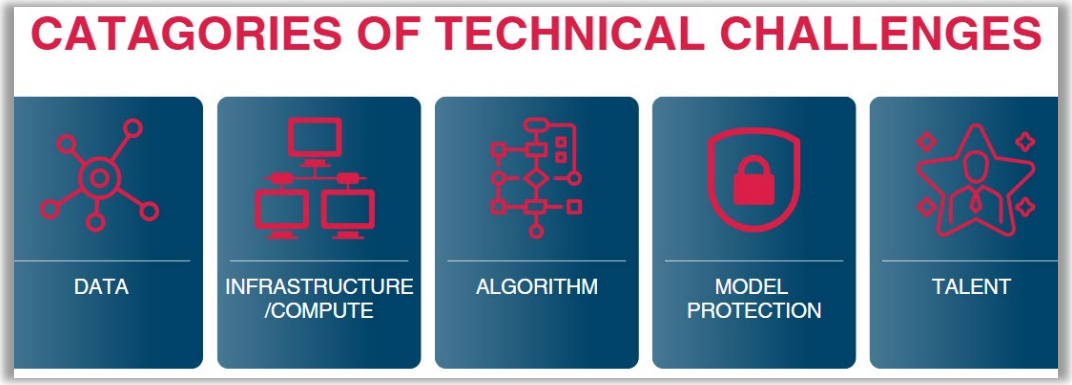
In 2023, DARPA launched AI Forward, which aims to explore new directions for AI to ensure trustworthy military systems.[23] AI Forward brings together the private sector, academia, and government to identify future DARPA AI focus areas and exploration opportunities. This initiative also leverages unique funding opportunities and streamlined contracting processes to rapidly address emergent AI research challenges and ensure trustworthiness in current and future AI systems.

Issues

The development of AI in the national security arena faces both technical and non-technical impediments.

[22] VADM Robert Sharp, "GEOINT: The Foundation of Intelligence". 25 April 2022. https://www.nga.mil/news/Remarks_as_prepared_for_delivery_by_Vice_Adm_Rober.html.
[23] Defense Advanced Research Projects Agency (DARPA). "AI Forward." Defense Advanced Research Projects Agency, https://www.darpa.mil/work-with-us/ai-forward.

The technical challenges fall into the following categories: data, infrastructure/compute, algorithm and model protection, and talent. All should be viewed as distinct from the additional DOD and Executive Branch institutional issues requiring a commensurate effort to ensure AI is properly utilized.



Data

Data, analytics, and AI play a role in all four DOD priorities outlined in the National Defense Strategy.[24] They constitute core capabilities that underpin DOD's operational and business analytics, as well as its decision-making in support of the Secretary of Defense's (SECDEF) priorities to "Defend the Nation," "Take Care of Our People," and "Succeed Through Teamwork." Data, analytics, and AI are also core capabilities of executing joint warfighting functions, especially CJADC2. However, specific obstacles to increased data usage include the following:

- DOD still uses a significant number of legacy systems. Some of these systems are in locations that make replacements or upgrades difficult for various reasons. For example, some legacy systems are located on ships or satellites, and others are employed in critical pieces of technology with a low tolerance for risk. Legacy systems may use data in antiquated formats not amenable to use by AI.
- Data is frequently unlabeled, rendering it unable to be leveraged to train AI algorithms.
- Data ownership is unclear, or DOD does not own relevant data. Contractors may own data and not make it available to train algorithms.
- Many programs across DOD and the services are siloed for programmatic or classification reasons, hindering both the accessibility and usability of the data for AI applications.
- Data is susceptible to attack through poisoning or purposeful insertion or deletion that maliciously alters algorithms or models.

[24] U.S. Department of Defense. 2022 National Defense Strategy, Nuclear Posture Review, and Missile Defense Review. October 2022, https://media.defense.gov/2022/Oct/27/2003103845/-1/-1/1/2022-NATIONAL-DEFENSE-STRATEGY-NPR-MDR.pdf.

Infrastructure/Compute

Infrastructure must satisfy certain operational criteria to support DOD's AI efforts. To deploy AI across the DOD enterprise, especially operationally and at the edge, sufficient infrastructure is necessary to transport data at the speed required for operational relevance. In conflict, especially in the Indo-Pacific area of operation, warfighters' use of AI could be limited by a lack of proper infrastructure or compute power or by an inordinate reliance on infrastructure (e.g., undersea cables or satellites) susceptible to attacks.

AI Algorithm and Model Protection

Algorithms and models must resist various attacks that could render them ineffective or harmful when deployed by DOD. Algorithms and models can be purposefully or accidentally manipulated to render them ineffective, useless, or inaccurate, potentially endangering human lives.

The CDAO, in close coordination with data owners and end users of AI systems, is developing guiding principles to ensure DOD components and vendors manage the entire lifecycle of AI training data, algorithms, and trained models. This control will help make the resulting AI systems secure and resilient to attacks, manipulation, or misuse by malicious state and non-state actors.

DOD's supply chain assurance and test and evaluation activities are essential for protecting AI data, algorithms, and models. Vulnerabilities in hardware and software supply chains, including our vendor base, can provide adversaries with attack vectors against DOD's AI capabilities.

The CDAO is working with the Office of the Under Secretary of Defense for Acquisition and Sustainment (OUSD A&S), DARPA, the Defense Logistics Agency (DLA), the Services, and the Military Department Counterintelligence Organizations to prototype AI supply chain risk evaluation and establish best practices for leveraging data to identify and prioritize supply chain threats. Hardware and software supply chains are secured from threats across the risk spectrum defined by the DOD Supply Chain Risk Management Taxonomy.

Talent

Realizing the full benefits of AI requires a workforce skilled in diverse technical and analytic skills. DOD must compete with the private sector for the technical talent required to develop, use, and deploy AI for military applications. DOD must also ensure adequate training of its workforce to understand how to best utilize AI for military use cases, ensure ethical applications of AI across the force, and integrate AI-enabled systems into legacy platforms.

Institutional

DOD faces organizational challenges in acquiring and using AI. Acquisition professionals, senior leaders, and warfighters often hesitate to adopt new, innovative technologies and their associated risk of failure.

DOD must shift this mindset to one more accepting of failure when testing and integrating AI and other innovative technologies. It also needs an acquisition workforce that understands AI and its impact on military systems. This workforce must understand the most effective acquisition pathways for AI systems and software.

It can take years to procure and scale a new capability across DOD, whereas AI and other software-centric technologies generally iterate in much shorter timeframes. To avoid technology obsolescence issues, DOD must accelerate its planning, programming, budgeting, and execution (PPBE) cycles and its procurement timelines to ensure that AI technologies, once procured, remain relevant and up to date. Simultaneously, AI is evolving faster than the development and adoption of associated policies for its responsible use. DOD must ensure that it updates policies appropriately to reflect the changing nature of technology and its impact on warfare.

> ***“It can take years to procure and scale a new capability across DOD, whereas AI and other software-centric technologies generally iterate in much shorter timeframes***

Non-technical challenges

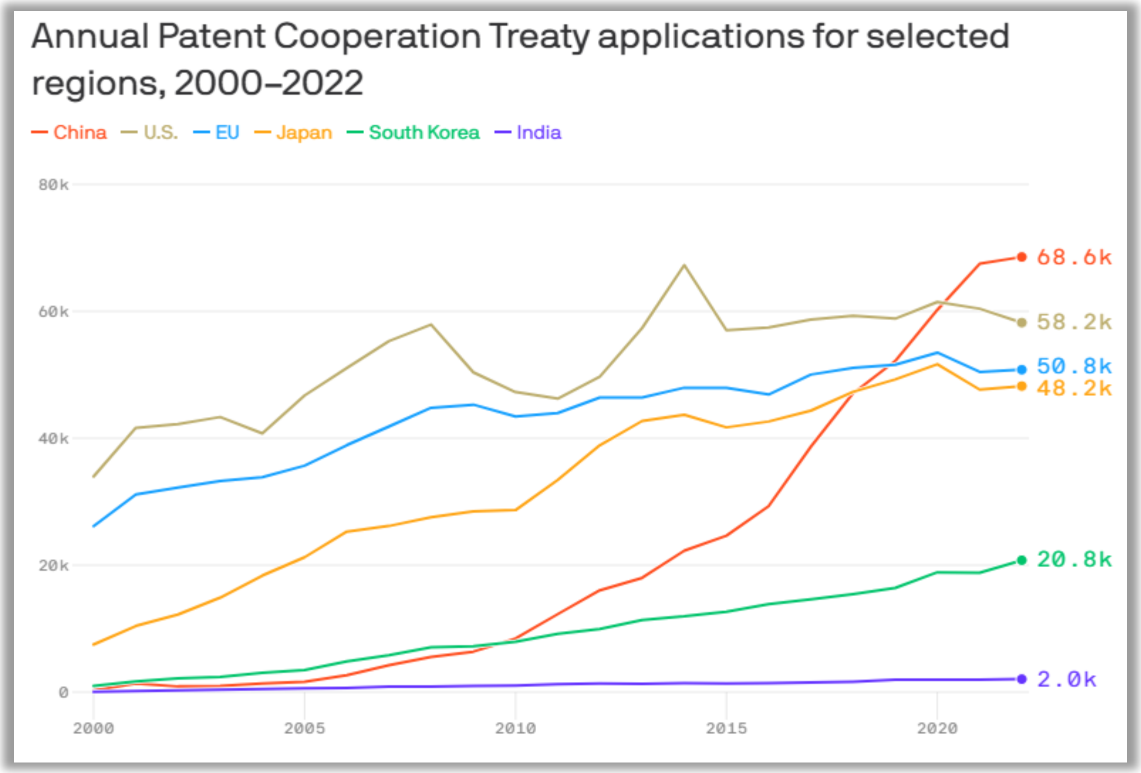
Non-technical challenges include the risks that stem from AI development, especially AI with advanced capabilities. Since DOD does not control the course of AI development by third parties, the resulting AI assumptions, reliability thresholds, use cases, purposes, and capabilities may not be appropriate for DOD uses.

External AI development also limits DOD’s ability to protect the technology from adversaries, including nation-states and individual actors. Although the United States is committed to developing AI ethically and responsibly, many other parties are not. Finally, the novelty of contemporary AI’s capabilities makes it more challenging to predict and address how it might be used for harm.

Dual-Use

The AI marketplace is primarily driven by commercial interests outside the traditional defense industrial base. DOD must take advantage of this burgeoning private sector and the associated capital markets to incentivize industry to develop solutions that benefit the warfighter. This will require DOD to engage in more timely procurement of AI solutions at scale with nontraditional defense contractors. At the same time, industry must also work to ensure adequate protection of DOD-relevant data and algorithms.

Many AI applications are dual-use in that their versatility and wide-ranging capabilities permit them to be used in both civilian and military capacities. Since the commercial sector widely uses such dual-use technologies, it is extremely difficult for DOD to limit adversaries' acquisition of these technologies. Consequently, DOD must be capable of both using and defending against dual-use AI.



Source: Axios - Patent applications from Chinese inventors pass U.S. for first time

Chinese Investment & Ethical AI Standards

U.S. adversaries are developing extremely advanced AI models, tools, and applications and are participating extensively in global AI research and publications. The Chinese Communist Party (CCP) has stated that China intends to become the world leader in AI by 2030.[25]

China has been ranked in the top three countries for global AI vibrancy.[26] It was ahead of the United States in several categories, including the number of AI patent applications, journal publications, and journal citations. It is estimated that China's burgeoning AI sector could create over $600 billion in economic value annually.[27]

[25] Mozur, Paul. "China's Artificial-Intelligence Boom." The New York Times, 20 July 2017, https://www.nytimes.com/2017/07/20/business/china-artificial-intelligence.html.
[26] "Tortoise Intelligence: Global AI." Tortoise Media, https://www.tortoisemedia.com/intelligence/global-ai/.
[27] Shen, Kai, et al. "The next Frontier for AI in China Could Add $600 Billion to Its Economy." McKinsey & Company, 7 June 2022, www.mckinsey.com/capabilities/quantumblack/our-insights/the-next-frontier-for-ai-in-china-could-add-600-billion-to-its-economy.

In addition to these economic effects, the CCP has outlined an innovation strategy for the People's Liberation Army (PLA). A component of this strategy is to militarize AI for next-generation warfighting, thereby obtaining a military advantage over countries like the United States.

China's success in AI is partly due to the large data sets available to the government under its autocratic regime. The country has exceptionally large swaths of labeled data sets available to continuously train AI models and algorithms.[28] For the United States to lead in AI development over China, it must develop better methods to label large amounts of data at scale.

While the United States holds high standards for the ethical development and use of AI, our adversaries do not. The CCP, in particular, is not bound by the same standards of AI development and deployment as the United States. This raises grave concerns about the way China and the PLA will leverage AI systems and the underlying data as a weapon against the U.S.

Autonomy

AI and autonomy are inextricably linked. Autonomous systems can provide great value to the military for intelligence, surveillance, reconnaissance, logistics, base protection and defense, and more. However, autonomous weapons leveraging AI also raise concerns about ethics, decision-making, and governance. Autonomous systems must be carefully and continuously reviewed and aligned with DOD policy on autonomous weapons.

Standards Adoption

AI standards are rooted in the initiatives, frameworks, guidance, and ethical guidelines outlined above. Given how quickly AI technologies evolve, prematurely enshrining formal enterprise AI standards risks constraining DOD to approaches or processes that cannot keep pace with changes in the state of AI technology and commercial offerings.

Because of this, CDAO has adopted a dual approach to AI standards. In areas where a common enterprise approach is critical, such as those outlined in DOD AI Ethical Principles or the Responsible AI Strategy and Implementation Pathway, CDAO pursues early implementation of relevant AI standards.

In contrast, CDAO has adopted a more gradual, iterative approach to AI standards development in areas where it is less critical to have a single unified enterprise standard. This more measured pace of AI standards development enables experimentation and ensures DOD can keep pace with rapidly changing technologies.

[28] Data, Analytics, and AI Adoption Strategy. U.S. Department of Defense, 27 June 2023, https://media.defense.gov/2023/Nov/02/2003333300/-1/-1/1/DOD_DATA_ANALYTICS_AI_ADOPTION_STRATEGY.PDF.

This approach to AI standards is reflected in the process CDAO outlined in DOD's Data, Analytics, and AI Adoption Strategy[29] and the Federated AI Construct Implementation Plan.[30]

Through this approach, in collaboration with DOD Components and conversation with industry, CDAO iteratively assesses areas where common best practices, approaches, technical patterns, or standards may be appropriate for adoption and scaling across the enterprise. This approach, outlined in the Federated AI Construct Implementation Plan, was developed based on inputs from the Service Chief Data Officers, the Joint Staff J2, ODNI's Automating Intelligence Using Machines (AIM) Initiative, and NGA's Project Maven.

[29] Id.

[30] U.S. Government Accountability Office, "Artificial Intelligence: DOD Needs Department-Wide Guidance to Inform Acquisitions." U.S. Government Accountability Office, 29 June 2023, https://www.gao.gov/products/gao-23-105850.

Key Findings

AI is a critical component of national security.

AI technology and expertise with AI, encompassing both offensive and defensive use capabilities, is essential to our national security ecosystem. Congress should explore ways that the U.S. national security apparatus can safely adopt and harness the benefits of AI.

U.S. adversaries are adopting and militarizing AI.

U.S. adversaries are developing extremely advanced AI models, tools, and applications and are participating extensively in global AI research and publications.

National security requires advanced cloud access and AI.

In a conflict, the U.S. military will be required to operate over vast distances. Transporting data across these immense spaces and oceans will require effective use of the cloud, automation, and AI.

National security requires AI for contested environments.

In a conflict, all systems, including AI-enabled autonomous systems, will be required to operate in contested, denied, and degraded environments. DOD is actively exploring requirements for compute at the edge in order to ensure that AI remains relevant in contested environments and that autonomous systems remain secure and trustworthy.

AI can vastly improve DOD business processes.

AI can be deployed today to simplify back-office tasks and functions, such as auditing, financial management, and other business processes. This can result in increased cost savings and reduced bureaucracy.

Recommendations

Recommendation: Focus congressional oversight on AI activities for national security.

Congress should exercise its oversight function through briefings, hearings, letters, and other opportunities. Such oversight would ideally include hearing from those inside and outside the government. These interactions must recognize that Members of Congress and staff have varying levels of understanding of AI. The House Armed Services Committee (HASC) and other committees of jurisdiction should ensure that briefings and hearings on AI activities for national security encompass a variety of perspectives and are suitable for different levels of expertise.

Recommendation: Support expanded AI training at DOD.

DOD is expanding employee training on AI for acquisition professionals, warfighters, senior leaders, and others serving in the Department. However, Congress must ensure DOD vigorously embraces these efforts.

Recommendation: Continue oversight of autonomous weapons policies.

In 2023, DOD updated Directive 3000.09 Autonomy in Weapon Systems, which applies to autonomous and semi-autonomous weapon systems, including those that incorporate AI.[31] This policy establishes that "autonomous and semi-autonomous weapon systems will be designed to allow commanders and operators to exercise appropriate levels of human judgment over the use of force." The directive establishes guiding principles for designing, developing, acquiring, testing, fielding, and employing autonomous and semi-autonomous weapon systems. Congress should continue robust oversight of this and any relevant successor policies and support DOD's policy of requiring meaningful human control over the launch of nuclear weapons.

Recommendation: Support international cooperation on AI used in military contexts.

International cooperation will be key to addressing the broader security concerns posed by AI in military contexts. Encouraging global norms and agreements that enshrine oversight over AI in military contexts, such as through the Political Declaration on Responsible Military Use of AI, could strengthen global security efforts.[32]

[31] U.S. Department of Defense, DOD Directive 3000.09: Autonomy in Weapon Systems. U.S. Department of Defense, 25 Jan. 2023, https://media.defense.gov/2023/Jan/25/2003149928/-1/-1/0/DOD-DIRECTIVE-3000.09-AUTONOMY-IN-WEAPON-SYSTEMS.PDF.

[32] Bureau of Arms Control, Deterrence, and Stability. "Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy." U.S. Department of State, 9 Nov. 2023, https://www.state.gov/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy-2/.

RESEARCH, DEVELOPMENT, & STANDARDS

Background

AI has been transformative across the scientific, economic, and defense realms.[1] There is a worldwide race to lead in fundamental AI research and commercial applications. The U.S. remains the leader in fundamental research and standards and consistently produces cutting-edge AI applications, such as ChatGPT, before other nations. However, according to the latest National Science Board's (NSB) Science and Engineering Indicators report, adversarial nations like the People's Republic of China (PRC) are quickly outpacing the world in highly cited and collaborative research and development.[2] To maintain U.S. leadership in global AI innovation and governance, Congress will need to continue federal R&D efforts, supporting AI evaluations, and bolstering U.S. standardization efforts for AI.

The field of artificial intelligence is not new. The U.S. government has been investing in artificial intelligence research since the 1950s. However, in the last 10 years, the number of AI scientific publications began to accelerate, and in 2016, the U.S. government released its first initiative on AI.[3] Simultaneously, the private sector has been accelerating its investments in AI research and development (R&D), with most of it heavily concentrated in a few large companies. As a result, there is an increased focus on public-private partnerships in AI R&D to accelerate different areas of AI and broaden opportunities for smaller companies and institutions of higher education to contribute to their development.

[1] Eric Schmidt. "AI, Great Power Competition & National Security." Daedalus, 2022, https://www.amacad.org/publication/daedalus/ai-great-power-competition-national-security.

[2] National Science Board. The State of U.S. Science and Engineering 2024: Advancing the Future of the Nation's Workforce and Innovation. National Science Foundation, 2024, https://ncses.nsf.gov/pubs/nsb20243.

[3] National Science Board. "Artificial Intelligence and the Future of U.S. Technology and Workforce Development." National Science Foundation, 2020, https://ncses.nsf.gov/pubs/nsb20205/artificial-intelligence-technology.

AI R&D involves developing methods for learning from data, representing knowledge, and performing reasoning to build computer systems capable of performing tasks that typically require human intelligence. Fundamental research to advance AI systems seeks to improve methods for learning, reasoning, problem-solving, planning, knowledge representation, language understanding, and visual perception, as well as to understand how systems operate in the real world to assess and address risks, including safety, security, and bias.[4]

Understanding and mitigating potential risks while capturing benefits will be important to AI adoption across all domains and use cases, especially for safety-critical cases or cases in which AI-based decisions significantly affect individuals, communities, the environment, or society at large. Research in this area is needed to solve several major outstanding science and technology challenges, including improving transparency mechanisms, such as explainability and interpretability; evaluating the capabilities and limitations of AI systems; developing technical mitigations and defense mechanisms for highly capable AI systems; improving the abilities of AI systems or agents to perceive and act; and developing scalable, general-purpose AI systems that can be deployed across real and virtual environments.

AI R&D also involves the application of AI systems to advance research across fields of science and engineering. The following is a small sampling of scientific fields for which machine learning and AI have been helping to drive discovery for years:

- Biology: personalized genetic medicine; biomolecule structure prediction; microscopy image analysis; synthetic biology
- Geology: data interpretation of air, space, and marine sensors; geomodelling
- Astronomy: celestial body identification
- Chemistry: property prediction, sustainable chemistry
- Physics: analysis of complex systems; materials prediction; quantum information science

U.S. AI Research and Development

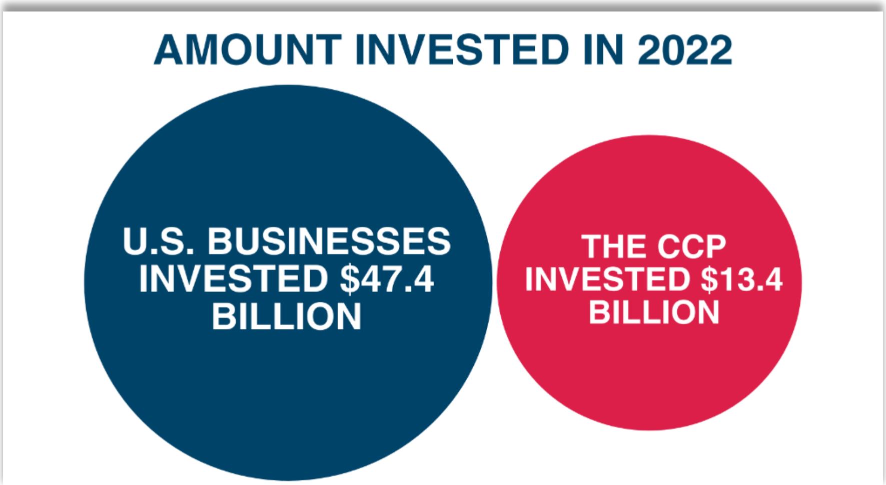
The United States has maintained its AI leadership largely due to continued and consistent federal investments in AI R&D over decades. Federal investments have enabled key discoveries that have driven the technology forward. For example, the National Science Foundation (NSF) investments supported key research that led to the development of neural networks, which are widely used by industry and academia today and whose progenitor was awarded the 2024 Nobel Prize in Physics.[5]

[4] Horvitz, Eric and Mitchell, Tom. “Scientific Progress in Artificial Intelligence: History, Status, and Futures.” February 2024, http://erichorvitz.com/AI_Overview_History_Status_Futures_February_2024.pdf.

[5] National Science Foundation. "NSF Congratulates the 2024 Nobel Prize in Physics Laureates." National Science Foundation, 2024, https://new.nsf.gov/news/nsf-congratulates-laureates-2024-nobel-prize-physics.

The United States' total AI R&D growth over the last decade has greatly benefited from contributions made by industry. While nonprofits and academia have contributed to this growth, building cutting-edge AI systems requires large amounts of data, computing power, and financial resources and industry members, especially large firms, tend to have greater access to these resources. According to the Stanford University AI Index Report 2023, private U.S. businesses invested $47.4 billion in 2022, roughly 3.5 times the amount invested by the next highest country's private industry (China, $13.4 billion).[6]

The U.S. is also the global leader in newly funded AI companies, with almost double the number of new companies in the European Union and United Kingdom combined, which is 3.4 times more than those in China.[7]



Source: Stanford University - The AI Index 2023 Annual Report

Federal Agency Investment in AI R&D

The most recent estimate of total U.S. federal R&D spending on AI was $2.9 billion in 2023.[8] While the private sector funds and performs the majority of U.S. R&D, the federal government has been the leading source of support for basic research. It funds R&D in areas where the industry lacks incentives to invest, which is critical for national security, public health, weather prediction, and other societal needs. Federal spending on non-defense AI R&D has increased from $560 million in fiscal year 2018 to $2.1 billion in 2023.[9] All federal science agencies are making substantial investments in fundamental AI R&D.

[6] N. Maslej, et al. "The AI Index 2023 Annual Report," Stanford University, April 2023. https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf.
[7] Id.
[8] National Information Technology Research and Development. "AI R&D Investments." Networking and Information Technology Research and Development, https://www.nitrd.gov/apps/itdashboard/ai-rd-investments/.
[9] Supra 6.

Federal R&D Expenditures for AI by Select Agencies (U.S. dollars, in millions)[10]

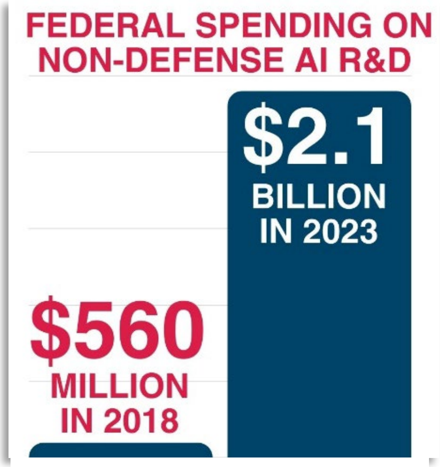
Agency	FY 2022 Actual	FY 2023 Enacted	FY 2024 Requested
DOE	224.9	246.4	240.6
NASA	5.7	6.6	12.8
NIH	860.4	919	925.7
NIST	31.5	36.2	36.4
NOAA	3.5	6.4	5
NSF	720	637.6	757.6
USDA	135.9	141.3	156.9
DARPA	429.8	400.5	322.1

The most significant step Congress has taken to enhance AI R&D was including the National Artificial Intelligence Initiative Act of 2020 (NAIIA) in Division E of the 2021 National Defense Authorization Act.[11] NAIIA takes a multifaceted approach to AI innovation. It prioritizes AI R&D through U.S.-sponsored research, strengthening research infrastructure, facilitating public-private partnerships, modernizing governance and technical standards for AI technologies, utilizing AI technologies for government services, promoting international engagement on AI, and providing AI education, including workforce R&D and re-skilling education.[12]

In January 2021, OSTP implemented a requirement in NAIIA and established the National AI Initiative Office (NAIIO).[13] The Office oversees and implements the United States' national AI strategy and serves as the central hub for federal coordination and collaboration in AI research and policymaking across the government and with the private sector, academia, and other stakeholders.

The AI initiative is further coordinated through various subcommittees and working groups under the National Science and Technology Council (NSTC), an executive-level council of advisors for the President primarily focused on coordinating science and technology policy.[14] The different interagency groups address different aspects of federal AI efforts, from coordinating fundamental research to monitoring developments in the private sector and internationally to carrying out specific NAIIA requirements, including reporting requirements.

[10] Supra 8.
[11] House of Representatives, Congress. 15 U.S.C. 9411 - National Artificial Intelligence Initiative. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title15/USCODE-2023-title15-chap119-subchapI-sec9411
[12] U.S. House of Representatives, Committee on Science, Space, and Technology. AI One-Pager: H.R. 6216, the National Artificial Intelligence Initiative Act of 2020. https://republicans-science.house.gov/_cache/files/b/1/b1cdbff1-29fb-42eb-9b68-62549079d797/29C8210F379D9896460FA4171986B70B.ai-one-pager-6216.pdf.
[13] The White House. "White House Launches National Artificial Intelligence Initiative Office." The White House, 2021, https://trumpwhitehouse.archives.gov/briefings-statements/white-house-launches-national-artificial-intelligence-initiative-office/.
[14] Office of Science and Technology Policy. "National Science and Technology Council (NSTC)." The White House, https://www.whitehouse.gov/ostp/ostps-teams/nstc/.



Source: NITRD - AI R&D Investments. FY19 – FY25.

The NSTC has released several reports on different aspects of the federal role in AI, including the first National AI R&D Strategic Plan in 2016.[15,16] The strategic plan calls for long-term investments in AI research to keep the United States a world leader in AI. It has been updated twice, in 2019 and 2023.[17,18] The plan identified knowledge gaps regarding fundamental AI questions, human-AI teaming, AI governance, measuring, testing, and standards for AI, robotics, and more. It included recommendations for research, research infrastructure, international cooperation, developing shared public data sets, fostering public-private partnerships, and AI R&D workforce needs.

Highlights of Federal AI R&D and Infrastructure Investments

As an illustration of the breadth of AI R&D, every research directorate at the NSF contributes to its total portfolio of AI R&D.[19] It is worth noting that NSF is our government's largest computer science research funder.

In addition, because of its Social, Behavioral, and Economic Sciences Directorate, NSF also has an outsized role in advancing research on ethical, legal, and social implications of AI and the science of AI governance more broadly, including by supporting interdisciplinary research teams that include social scientists and technology ethicists. As the U.S. faces intensifying global competition in science and technology, NSF's investments in emerging technology research are imperative for scaling innovation and commercializing basic research. Created in 2022, the Technology, Innovation, and Partnership (TIP) Directorate builds on NSF's longstanding leadership in science and engineering research and education. NSF recently released a 3-year roadmap for TIP, identifying AI and machine learning (ML) as the top priority.

[15] Office of Science and Technology Policy. "Preparing for the Future of Artificial Intelligence." The White House, October 2016, https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.

[16] Office of Science and Technology Policy. "Artificial Intelligence, Automation, and the Economy." The White House, December 2016, https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/documents/Artificial-Intelligence-Automation-Economy.PDF.

[17] "National Artificial Intelligence Research and Development Strategic Plan: 2019 Update". National Coordination Office for Networking and Information Technology Research and Development, 2019, https://www.nitrd.gov/pubs/National-AI-RD-Strategy-2019.pdf.

[18] "National Artificial Intelligence Research and Development Strategic Plan: 2023 Update". National Coordination Office for Networking and Information Technology Research and Development, 2023, https://www.nitrd.gov/pubs/National-Artificial-Intelligence-Research-and-Development-Strategic-Plan-2023-Update.pdf.

[19] National Science Foundation. "What is Artificial Intelligence?" National Science Foundation, https://new.nsf.gov/focus-areas/artificial-intelligence#what-is-artificial-intelligence-695.

The United States Department of Energy (DOE) and its National Laboratory complex have invested in AI research and development for decades. During this time, they have been responsible for developing cutting-edge AI tools and high-performance computing for open science and classified research needs.

Frontier at Oak Ridge National Laboratory is the most powerful supercomputer in the world.

Within the federal complex, DOE has unique computational resources, including the world's top two most powerful supercomputers[20] and physical and digital infrastructure supporting both governmental and non-governmental scientific research.

This has enabled them to rapidly advance the development of complex AI models, which could potentially accelerate discoveries in several fundamental science fields, such as materials science, chemistry, and biology.

The scope of these efforts is outlined in a May 2023 report, *Advanced Research Directions on AI for Science, Energy, and Security,* in which DOE National Lab representatives propose a strategy for further developing AI tools for scientific innovation.[21]

DOE also recently introduced a proposal titled "Frontiers in Artificial Intelligence for Science, Security, and Technology" (FASST), highlighting key areas within their jurisdiction where AI could rapidly be leveraged.[22] These include increasing the electricity grid's reliability, developing more effective cancer screenings and treatments, and improving stewardship of the nuclear weapons stockpile. In July 2024, DOE announced the roadmap for the FASST Initiative, which will harness the Department's advanced supercomputing, research infrastructure, and the vast repositories of scientific data produced by its facilities.[23]

[20] Keumars Afifi-Sabet. "Top 7 Most Powerful Supercomputers in the World Right Now." Live Science, March 2024, https://www.livescience.com/technology/computing/top-7-most-powerful-supercomputers-in-the-world-right-now.
[21] U.S. Department of Energy. "Advanced Research Directions on AI for Science, Energy, and Security." U.S. Department of Energy, May 2023, https://www.anl.gov/sites/www/files/2024-05/AI4SESReport-2023-v7.pdf.
[22] U.S. Department of Energy. "FASST: Artificial Intelligence and Machine Learning for Science and Technology." U.S. Department of Energy, https://www.energy.gov/fasst.
[23] U.S. Department of Energy. "DOE Announces Roadmap for New Initiative on Artificial Intelligence for Science, Security, and Energy." U.S. Department of Energy, 2024, https://www.energy.gov/articles/doe-announces-roadmap-new-initiative-artificial-intelligence-science-security-and.

Several other federal agencies play critical roles in our nation's AI R&D. While this report falls far short of capturing the breadth of federal R&D, it does include additional information in other chapters. Please see the **Healthcare** chapter for a summary of research at the National Institutes of Health (NIH), the **National Security** chapter for a summary of research at the Department of Defense (DOD), and the **Agriculture** chapter for a summary of research at the United States Department of Agriculture (USDA).

National Artificial Intelligence Research Institutes

As part of the National AI Initiative Act of 2020, Congress directed the creation of a network of AI Institutes, coordinated through NSF, that any federal department or agency can fund to create partnerships between academia and the public and private sectors to accelerate AI research.[24, 25] Participating agencies include USDA, the National Institute of Standards and Technology (NIST), and the Department of Veteran's Affairs. Since enactment, twenty-seven institutes have been established, with $500 million in total investments, focusing on areas ranging from agriculture to veterans' affairs to education.[26] NSF plans to award up to three more in FY 2025.[27] Each awardee comprises one or more research institutions across the U.S. and seeks to promote foundational AI research.

National Artificial Intelligence Research Resource

The National AI Initiative Act of 2020 created a task force to examine the merits of establishing a National AI Research Resource (NAIRR), connecting capable but under-resourced researchers to dependable computational data, software, training models, and other resources needed to advance AI research.[28, 29]

The NAIRR task force published its final report in January 2023, issuing final recommendations to strengthen and democratize the U.S. AI innovation ecosystem.[30]

[24] National Science Foundation. "National Artificial Intelligence Research Opportunities." National Science Foundation, https://new.nsf.gov/funding/opportunities/national-artificial-intelligence-research.

[25] National Science Foundation. "National Artificial Intelligence Research Institutes", National Science Foundation. September 2023, https://nsf-gov-resources.nsf.gov/2023-09/AI_Institutes_Hill_Day_Booklet.pdf?VersionId=pw2q_TeAvI05kmyLqqr.59M1IPocA84w.

[26] "National Artificial Intelligence Research Resource Pilot." National Science Foundation, 24 Jan. 2024, new.nsf.gov/focus-areas/artificial-intelligence/nairr.

[27] National Science Foundation. Fiscal Year 2025 National AI Research Institutes Strategic Plan. September 2023, https://nsf-gov-resources.nsf.gov/files/59_fy2025.pdf?VersionId=InUkjV8YETHGc0DM73jRVKA_m_CDL4s6.

[28] National Science Foundation. "National Artificial Intelligence Research Resource." National Science Foundation, https://new.nsf.gov/focus-areas/artificial-intelligence/nairr.

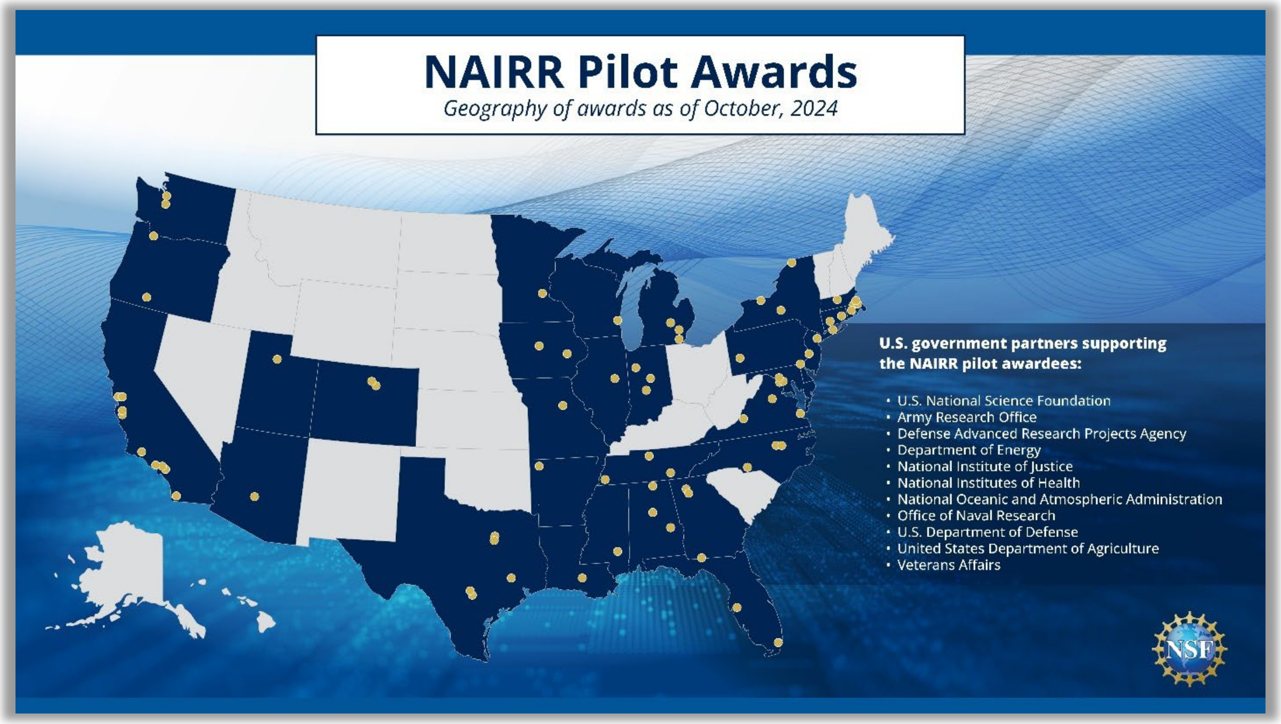
[29] Supra 11.

[30] "Strengthening and Democratizing the U.S. Artificial Intelligence Innovation Ecosystem," National AI Research Resource Task Force, January 2023. https://www.ai.gov/wp-content/uploads/2023/01/NAIRR-TF-Final-Report-2023.pdf.

One year after the NAIRR task force published its recommendations, NSF launched a two-year pilot program for the NAIRR in partnership with ten other federal agencies and more than 25 private sector, non-profit, and philanthropic organizations.[31] NSF invested $30 million in the NAIRR pilot in FY24 and included an additional $30 million in the President's budget request for FY25. Thirty-five initial awards for computing time were announced in May 2024, and as of today, more than 100 total awards represent principal investigators across 25 states, and additional awards are expected.

The NAIRR pilot leverages the Department of Energy's computational assets, including Oak Ridge National Laboratory's Summit supercomputer.[32] DOE will also be assisting with access to several different National Laboratory projects, including Argonne National Laboratory's Leadership Computing Facility's AI Testbed[33] and Oak Ridge's CITADEL framework[34] for protected data.

Moreover, the pilot will include a program called "NAIRR Secure," a multi-agency partnership comprised of DOE, NSF, and NIH that will assemble privacy and security-preserving resources to enable research involving sensitive data.



NAIRR Pilot Awards

Geography of awards as of October, 2024

U.S. government partners supporting the NAIRR pilot awardees:

- U.S. National Science Foundation
- Army Research Office
- Defense Advanced Research Projects Agency
- Department of Energy
- National Institute of Justice
- National Institutes of Health
- National Oceanic and Atmospheric Administration
- Office of Naval Research
- U.S. Department of Defense
- United States Department of Agriculture
- Veterans Affairs

Source: NSF - U.S. NAIRR pilot brings cutting-edge AI resources to researchers and educators across the nation

[31] Supra 27.
[32] Summit was initially slated for decommissioning at the end of 2023, but DOE extended its life until October 2024 to assist this program.
[33] Argonne National Laboratory. "ALCF AI Testbed." Argonne National Laboratory, https://www.alcf.anl.gov/alcf-ai-testbed.
[34] Oak Ridge National Laboratory. "Citadel User Guide." Oak Ridge National Laboratory, https://docs.olcf.ornl.gov/systems/citadel_user_guide.html.

Participants of the NAIRR pilot hope it will serve as a proof-of-concept model for connecting researchers and educators to resources and demonstrate NAIRR's ability to advance novel, transformative, and public-interest AI research. H.R. 5077, the CREATE AI Act of 2023, is a bill introduced by members of the House AI Caucus,[35] Representatives Eshoo, McCaul, Beyer, and Obernolte, that would authorize the development of a complete NAIRR.[36] Implementation and evaluation of the NAIRR pilot should be monitored in preparation for a possible full-scale NAIRR.

Academic Research

Academic research advances foundational knowledge in science and technology. Fundamental science research in universities drives discoveries that can lead to new or improved technologies while teaching and training the next generation of researchers. University research is also the source of thousands of spin-off companies contributing to regional economic development and job creation.[37]

Such spin-offs are primarily clustered near the university, which drives regional innovation. However, AI research at universities continues to be limited by access to data and computational power. Even the most well-resourced academic institutions do not have the resources to train AI systems of comparable complexity as the most advanced AI models. Institutions that serve rural areas and minority-serving institutions face even deeper challenges in accessing these resources. As a result, many researchers have left academia entirely for industry.

AI R&D would be bolstered by providing researchers greater access to advanced computational resources. This especially benefits those facing significant financial, capacity, or logistical challenges when participating in the AI research ecosystem. Science agencies could facilitate access to computational resources from a wide range of advanced computing technologies, from traditional on-site computing to cloud computing to emerging computing paradigms, such as edge computing and quantum computing. Because of resource constraints, such cyberinfrastructure would have natural tradeoffs between providing resources for a large number of smaller AI research projects and a small number of frontier research projects that need significant resources.

[35] Artificial Intelligence Caucus of the U.S. House of Representatives. Rep. Anna Eshoo, https://artificialintelligencecaucus-eshoo.house.gov/.
[36] "H.R.5077 - 118th Congress (2023-2024): CREATE AI Act of 2023." Congress.gov, Library of Congress, 11 Sept. 2024, https://www.congress.gov/bill/118th-congress/house-bill/5077.
[37] Prokop, D., Huggins, R., & Bristow, G. (2019). "The survival of academic spinoff companies: An empirical study of key determinants." International Small Business Journal, 37(5), 502-535. https://doi.org/10.1177/0266242619833540.

The development of shared public AI-ready datasets would also bolster R&D. Fundamental research, including manual laboratory measurements, automated laboratories, curation of existing data, and new experimental methods, are all critical for generating the data needed by AI models.[38] Further, consistent federal policies, guidelines, and tools that make data findable, accessible, interoperable, and reusable can significantly advance AI research across disciplines.[39]

Federal, academic, and industry researchers often lack streamlined access to existing data, which is critical for efficiently discovering, developing, and translating discoveries to industry across R&D.[40,41] This issue is outsized in the biotechnology space.

Although the U.S. has historically funded some biological database development and management at NIH, most biological data assets are fragmented across several governmental and nongovernmental repositories without considerations for interoperability. Likewise, these repositories can substantially differ in metadata curation and quality, further hindering the integration of this data to train sophisticated AI models. This may partially be explained by the lack of consistent standards for how biotechnology data and metadata are reported and stored for access.[42]

Federal investment in open-source software libraries and toolkits can also support AI R&D. Federal activities, such as the NSF's Pathways to Enable Open-Source Ecosystems program, have long supported open-source environments to spur innovation in critical and emerging technologies.[43] Federal agencies can develop open software libraries or contribute to them in areas where industries do not have the market incentive to develop these tools for government or other sectors.[44]

Scientific innovation requires the sharing of experimental design, results, and data. By sharing this information, researchers encourage more transparency, reproducibility of results, and dissemination of findings among practitioners. Many of the seminal papers that allowed AI development to accelerate over the last few years were openly published, such as Google's Transformers paper,[45] which underpins systems like ChatGPT.

[38] Id.

[39] U.S. National Security Commission on Emerging Biotechnology (NSCEB). "Leveraging Biological Data." National Security Commission on Emerging Biotechnology, 2024, https://www.biotech.senate.gov/press-releases/leveraging-biological-data/.

[40] The White House. "Visions, Needs, and Proposed Actions for Data for the Bioeconomy Initiative." The White House, 2023, https://www.whitehouse.gov/wp-content/uploads/2023/12/FINAL-Data-for-the-Bioeconomy-Initiative-Report.pdf.

[41] Supra 31.

[42] Supra 40.

[43] National Science Foundation. "POSE: Pathways to Enable Open-Source Ecosystems." National Science Foundation, https://new.nsf.gov/funding/opportunities/pose-pathways-enable-open-source-ecosystems.

[44] Supra 19.

[45] Jakob Uszkoreit, et al. "Transformer: A Novel Neural Network Architecture for Language Understanding." Google Research Blog, 2017, https://research.google/blog/transformer-a-novel-neural-network-architecture-for-language-understanding/.

However, in recent years, many industrial labs have been pulling back on the number of AI-related academic papers they publish or present at conferences. While trade secrets and competitive advantage are legitimate considerations in what to publish, there is a risk that a closed AI research ecosystem could limit U.S. competitiveness in AI.

To promote AI innovation, federal and nongovernmental organizations should collaborate to find a more appropriate balance between openness on the one hand and IP, safety, and national security risks on the other. Many observers have described a "valley of death" between basic research conducted at U.S. universities and the commercialization activities typically carried out by industry.[46] Universities generally lack the means of production necessary to take initial research results and generate marketable products. According to the American Academy of Arts and Sciences, the pace of American translation of discoveries and inventions from laboratory research to products must accelerate for the U.S. to remain competitive.[47]

The Task Force finds closer cooperation among industry, government, and academia could increase technology transfer, stimulate innovation, lead to new products and processes, and expand markets.

Research on AI Evaluations and Testing

Many AI systems consistently produce positive, predictable, and intended results, but some fail or intermittently produce unintended results. As a result, AI systems routinely fail to meet performance, safety, or reliability expectations.[48] With the recent rise in the capabilities of AI systems, there has been a corresponding growth of interest in managing the risks of AI systems.[49] Some of this discussion has focused on addressing theoretical safety concerns or aligning the development of AI systems with certain values,[50] while other researchers have focused on sociotechnical challenges, looking at AI risks through the complex human, organizational, and technical factors involved in AI design, development, and use.[51] Determining good and bad results and the impact of

[46] Charles Wessner, et al. "Driving Innovations Across the Valley of Death." ResearchGate, 2005, https://www.researchgate.net/publication/263062453_Driving_Innovations_Across_the_Valley_of_Death.
[47] Moore, J., & Wilson, I. 2021. "Decades of basic research paved the way for today's Covid-19 vaccines." https://www.statnews.com/2021/01/05/basic-research-paved-way-for-warp-speed-covid-19-vaccines/.
[48] Inioluwa Deborah Raji, I. Elizabeth Kumar, Aaron Horowitz, and Andrew Selbst. 2022. "The Fallacy of AI Functionality. In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)." Association for Computing Machinery, https://doi.org/10.1145/3531146.3533158.
[49] Dario Amodei, Chris Olah, Jacob Steinhardt, Paul Christiano, John Schulman, and Dan Mané. Concrete problems in AI safety. Dan Hendrycks, Mantas Mazeika, and Thomas Woodside. An Overview of Catastrophic AI Risks, September 2023. http://arxiv.org/abs/2306.12001.
[50] Iason Gabriel. Artificial Intelligence, Values, and Alignment. Minds and Machines, 30(3):411–437, September 2020. https://link.springer.com/10.1007/s11023-020-09539-2.
[51] Miranda Bogen and Amy Winecoff. "Applying Sociotechnical Approaches to AI Governance in Practice." Center for Democracy & Technology, 2024, https://cdt.org/insights/applying-sociotechnical-approaches-to-ai-governance-in-practice/#:~:text=Applying%20a%20sociotechnical%20lens%20to,of%20deployment%20in%20unexpected%20ways.

any given AI system will require effective and consistent testing and evaluation of AI systems.

Developers currently use several common practices to assess their AI systems.[52] The first is benchmark tests, which quantitatively evaluate the accuracy of an AI model. Developers can also use multidisciplinary teams of model developers, relevant subject matter experts, experts in responsible AI development, legal experts, and others to evaluate generative AI models before deployment. Developers also utilize post-deployment monitoring of their projects to detect improper use or violations of their terms of service.

Finally, developers commonly deploy AI red teaming, a structured process that leverages outside experts to test AI systems for vulnerabilities and flaws. In many cases, tests and evaluations are developed alongside the technology itself by private industry. As a result, there are often limited mechanisms for public qualitative evaluation and testing.

Unfortunately, many of these forms of AI evaluation currently lack rigorous scientific methodologies.[53] Several major scientific challenges underpin the evaluation of advanced AI systems that reduce reproducibility or have limited utility.[54] For example, current performance-oriented evaluations may be good at evaluating individual models but are bad at comparing capabilities between different models.[55] Further, some forms of evaluation may be insufficient by themselves. For example, red teaming is limited because developers must know all the risks and variables before testing.[56]

Federal agencies have proposed several activities to improve the evaluation of AI systems. The CHIPS and Science Act directed NIST to establish testbeds to evaluate AI systems.[57] Executive Order 14110, titled “Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence,” also established evaluation and testing activities at NIST, DOE, and the Department of Homeland Security.[58]

[52] U.S. Government Accountability Office. “Artificial Intelligence: Opportunities and Challenges in the U.S. Government.” GAO-25-107651, 2023, https://www.gao.gov/assets/gao-25-107651.pdf.

[53] John Burden, "Evaluating AI Evaluation: Perils and Prospects." Leverhulme Centre for the Future of Intelligence, Centre for the Study of Existential Risk, University of Cambridge: arXiv, 2024, https://arxiv.org/html/2407.09221v1#bib.bib7.

[54] Narayanan and Kapoor . "Evaluating Large Language Models: A Minefield." Princeton University, https://www.cs.princeton.edu/~arvindn/talks/evaluating_llms_minefield/#/9.

[55] Id.

[56] Sorelle Friedler et al., “AI Red-Teaming Is Not a One-Stop Solution to AI Harms: Recommendations for Using Red-Teaming for AI Accountability,” Data and Society, 25 Oct. 2023, https://datasociety.net/library/ai-red-teaming-is-not-a-one-stop-solution-to-ai-harms-recommendations-for-using-red-teaming-for-ai-accountability/.

[57] Office of the Federal Register, National Archives and Records Administration. Public Law 117 - 167 - An act making appropriations for the Legislative Branch for the fiscal year ending September 30, 2022, and for other purposes. U.S. Government Publishing Office, 8 Aug. 2022, https://www.govinfo.gov/app/details/PLAW-117publ167.

[58] The White House. "Fact Sheet: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence." The White House, 30 Oct. 2023, https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/.

NIST also established an AI Safety Institute to help support AI standardization related to AI safety and evaluation-related challenges.[59] The AI Safety Institute has also recently created an associated public-private consortium of over 200 organizations.[60]

U.S. Leadership in Standards Development for AI

The strength of the United States in international standards development will be instrumental to its global technological leadership in the development and governance of artificial intelligence.[61] At its core, a standard is a repeatable, harmonized, agreed-upon, and documented way of doing something.[62]

> **“*The strength of the U.S. in international standards development will be instrumental to its global technological leadership in the development and governance of artificial intelligence.***

Standards contain technical specifications, requirements, guidelines, or characteristics that can be used to ensure that materials, products, processes, and services are fit for purpose. Adherence to standards is usually voluntary. Technical standards can become mandatory for the private sector when governments adopt them as a requirement in legislation or regulation. Governments may also impose internal standards on their own agencies, as the United States does with cybersecurity. International standards facilitate international trade, enable U.S. competitiveness, and support consumer protection, security, public health, and safety.

The primary agency supporting standardization is NIST, which supports U.S. competitiveness by advancing measurement science, standards, and technology.[63] NIST is responsible for providing and maintaining many inputs and tools that support standard-setting activities. NIST offers standard reference materials, data, and instrumentation to help users verify the accuracy of specific measurements.

[59] National Institute of Standards and Technology. “Artificial Intelligence Standards and Innovation (AISI).” National Institute of Standards and Technology, https://www.nist.gov/aisi.

[60] National Institute of Standards and Technology. "Artificial Intelligence Safety Institute Consortium (AISIC)." National Institute of Standards and Technology, https://www.nist.gov/aisi/artificial-intelligence-safety-institute-consortium-aisic.

[61] Testimony Before the Subcommittee on Research and Technology: Advancing U.S. Leadership in Artificial Intelligence and Quantum Computing. U.S. House of Representatives, 17 March 2022, https://republicans-science.house.gov/_cache/files/4/f/4fbb3c44-3dd2-4c2b-bcbd-2d2c21029ce3/FA18AABDD2217B0C4777C04D3CCA18C6.2022-03-17-testimony-olthoff.pdf.

[62] National Institute of Standards and Technology. "Learn More About Standards." National Institute of Standards and Technology, https://www.nist.gov/standardsgov/learn-more-about-standards.

[63] The National Bureau of Standards Organic Act of 1901, 57th Congress, P.L. 56-177. https://www.nist.gov/system/files/documents/2017/05/09/NIST-Organic-Act.pdf.

The CHIPS and Science Act directed NIST to lead information exchange and coordination among federal agencies and communication from federal agencies to the private sector to support international standards development.[64] NIST acts as a convener, bringing together industry, academia, and government stakeholders to facilitate the development of standards that meet national priorities. Further, NIST facilitates coordination between federal, state, and local governments for standards engagement, adoption, and conformity assessment activities.[65] NIST also directly engages in standards setting bodies and tracks U.S. representation.[66]

While most countries worldwide have a top-down approach to setting standards, the United States has long maintained an industry-led, bottom-up approach to most standard setting.[67]

The U.S. standards system protects against poor standards by enabling vibrant deliberation and competition and ensuring that technical merit prevails. The government plays a supportive role by providing technical inputs to enable the standard setting, supporting scientific R&D, facilitating an open investment climate, promoting a rules-based standards system, and adopting consensus standards wherever possible.

As a result of this system, the United States has long held global leadership in standard setting. Federal coordination, tracking of federal participation, and increasing federal participation in standard setting can help promote continued U.S. leadership in international standards bodies.

In May 2023, a National Standards Strategy on Critical and Emerging Technologies was released, outlining the federal's role in supporting industry-led standards development.[68] NIST published the implementation roadmap for this strategy in July 2023.[69] This plan and implementation roadmap are designed to supplement, not supplant, the American National Standards Institute's U.S. Standards Strategy, which was updated in 2020.[70]

U.S. leadership in international standards setting is at risk, however. The increasing pace of technological change and globalization combined with the rise of strategic competitors has created challenges to our standards leadership.

[64] Supra 57.

[65] National Institute of Standards and Technology. "Interagency Committee on Standards Policy (ICSP)." National Institute of Standards and Technology, https://www.nist.gov/standardsgov/interagency-committee-standards-policy-icsp.

[66] Department of Defense Appropriations Act, 1979, Pub. L. No. 93-144, 93 Stat. 144, 1979, https://www.govinfo.gov/content/pkg/STATUTE-93/pdf/STATUTE-93-Pg144.pdf.

[67] National Institute of Standards and Technology. "Setting Standards to Strengthen U.S. Leadership in Technical Standards." National Institute of Standards and Technology, 2023, https://www.nist.gov/speech-testimony/setting-standards-strengthening-us-leadership-technical-standards#:~:text=In%20the%20U.S.%2C%20our%20voluntary,system%20in%20the%20United%20States.

[68] The White House. "U.S. Government National Standards Strategy," The White House, 2023. https://www.whitehouse.gov/wp-content/uploads/2023/05/US-Gov-National-Standards-Strategy-2023.pdf.

[69] The White House. U.S. Government National Standards Strategy: Implementation Roadmap, Version 7. 2024, https://www.whitehouse.gov/wp-content/uploads/2024/07/USG-NSSCET_Implementation_Rdmap_v7_23.pdf.

[70] American National Standards Institute. U.S. Standards Strategy 2020 Edition. 2020, https://share.ansi.org/Shared%20Documents/Standards%20Activities/NSSC/USSS-2020/USSS-2020-Edition.pdf.

Seeing international standards as another tool to gain competitive advantage for domestic industries, some countries have taken steps in clear contradiction of the core principles of international standard-setting that the United States and its like-minded partners have maintained for decades.

As a result of efforts by both allies and adversaries like the Chinese Communist Party (CCP),[71] there is a renewed focus in policy circles on technical standards, how they are set, and who sets them.

In recent years, U.S. stakeholders have cited increasing Chinese competition in standards setting.[72] There has been a rise in the number of Chinese companies participating in standards development organizations (SDOs), the number of proposals and submissions from by Chinese companies, and the number of Chinese nationals taking leadership positions in these organizations.[73] However, the number of participants, proposals, and leadership positions a nation holds does not equate to effectiveness by international standards.[74]

While stakeholders submit proposals of varying quality to these bodies, only the ones with the most technical merit are adopted. Standards experts from the U.S. industry argue that a standard's “success” can be better measured by the degree to which it is adopted in the marketplace because it meets a market need or opportunity.[75]

Policies undermining the U.S. approach of a bottom-up, rules-based, multistakeholder process for setting standards could disadvantage American companies and embolden our adversaries to ostracize U.S. firms from their domestic standards processes. For example, in 2019, the Bureau of Industry and Security (BIS) put Chinese company Huawei on the Entity List, a register of foreign individuals and organizations that pose a national security concern to the U.S. and are subject to export restrictions and licensing requirements.[76]

[71] Matt Sheehan, Marjory S Blumenthal, and Michael R. Nelson. "Three Takeaways From China's New Standards Strategy" Carnegie Endowment for International Peace. 2021, https://carnegieendowment.org/research/2021/10/three-takeaways-from-chinas-new-standards-strategy?lang=en.

[72] Mark Montgomery and Theo Lebryk, “China's Dystopian “New IP” Plan Shows Need for Renewed US Commitment to Internet Governance,” Just Security, 13 April 2021. https://www.justsecurity.org/75741/chinas-dystopian-new-ip-plan-shows-need-for-renewed-us-commitment-to-internet-governance/.

[73] Xirui Li and Dingding Chen, “Should the West Fear China's Increasing Role in Technical Standard Setting?” The Diplomat, 15 April 2021. https://thediplomat.com/2021/04/should-the-west-fear-chinas-increasing-role-in-technical-standard-setting/#:~:text=Internationally%2C%20from%202011%20to%202020,percent%20and%2067%20percent%2C%20respectively.

[74] Exovera. A New “Great Game?”: China's Role in International Standards for Emerging Technologies. Center for Intelligence Research and Analysis, 2022, https://cira.exovera.com/wp-content/uploads/2022/08/NIST-Final-Report.pdf.

[75] Nigel Cory. "America's National Security Concerns Over China Shouldn't Imperil Its Leadership in Technical Standards Development." Information Technology and Innovation Foundation, 20 Jan. 2023, https://itif.org/publications/2023/01/20/americas-national-security-concerns-over-china-shouldnt-imperil-its-leadership-in-technical-standards-development/.

[76] Entity List. U.S. Department of Commerce, Bureau of Industry and Security, https://www.bis.gov/entity-list.

BIS did not create an exception for U.S. companies to continue participating in standards development activities in which Huawei was also participating.[77] As a result, American companies were severely disadvantaged in certain international standards setting activities (until BIS later issued clarification).[78]

The greatest strength of the U.S. approach to voluntary consensus standards is its bottom-up, rules-based, multistakeholder process in which technical merit wins the day.

Federal policies regarding voluntary consensus standards, trade, or strategic competitors that deviate from this approach can risk our national interests. The federal government can work with allies to uphold the U.S. open, rules-based approach in international standards bodies.

AI Standards

For many reasons, AI standards are much less mature than other information technology standards. Some AI standardization needs are novel or more complex than other information technology standards, and there are significant gaps in the underlying science to inform standards development.

Standards for AI systems will most likely need to be context-dependent. For example, an AI system deployed in the aviation context will need technical standards different from those of an AI chatbot. Further, elements of AI standards are sociotechnical in a way that technical standards bodies lack experience addressing. All of this is in a context where technology is changing rapidly.

> **“*The greatest strength of the U.S. approach to voluntary consensus standards is its bottom-up, rules-based, multistakeholder process.***

Mature standards for AI will be critical for effective governance of the development and use of AI systems, whether through regulation or some kind of incentive-based system. Standards bodies have been developing related standards for years, but these processes are slow and limited by all the challenges described above.[79]

[77] Federal Register Notices: 2019 Regulations. U.S. Department of Commerce, Bureau of Industry and Security, https://www.bis.doc.gov/index.php/federal-register-notices/17-regulations/1541-federal-register-notices-2019.

[78] Note, for this reason, the policy was later lifted after years of being in effect.
U.S. Federal Register. "Release of Technology to Certain Entities on the Entity List in the Context of Standards." Federal Register, 18 June 2020, https://www.federalregister.gov/documents/2020/06/18/2020-13093/release-of-technology-to-certain-entities-on-the-entity-list-in-the-context-of-standards.

[79] See IEEE 7000 series on AI. “AUTONOMOUS AND INTELLIGENT SYSTEMS (AIS),” https://standards.ieee.org/initiatives/autonomous-intelligence-systems/.

In the NAIIA, Congress directed NIST to support AI standards development.[80] In *Executive Order 14110,* the President directed NIST to carry out a series of activities to support standardization and evaluation related to AI systems.[81] For example, NIST established a plan for global engagement on AI standards after stakeholder comments.[82]

Given that technical standards processes take a significant amount of time, the federal government can help organizations responsibly adopt AI systems and AI governance by developing or distributing existing guidance and best practices for AI systems.

For example, NIST developed a voluntary AI risk management framework through collaboration with stakeholders across the public and private sectors.[83] NIST developed the framework (published in January 2023) through a widely applauded, consensus-driven, open, transparent, and collaborative process.

NIST also published a draft playbook to help organizations with implementation. However, the first iteration of the framework only sets the theoretical baseline for identifying and mitigating AI risks by guiding readers to think critically about the context, measurement, and management of AI systems. Similarly, the agency produced a specific generative AI risk profile under the risk management framework.[84] These documents can also make their way into standards processes over time.

Several other federal agencies will play key roles in supporting AI-related standardization. The Department of State advises the President on foreign policy issues and leads on behalf of the United States in treaty-based international standards bodies, such as the International Telecommunication Union. Because the State Department lacks the technical expertise to engage in many technical standards, it sometimes delegates leadership to other expert agencies. Mission-oriented agencies that focus on specific sectors, such as the Department of Energy, may also engage with international standards organizations that set AI-related standards relevant to their missions.

[80] Supra 11.
[81] Supra 59.
[82] National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework. 2023, https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-5.pdf.
[83] National Institute of Standards and Technology. AI Risk Management Framework. National Institute of Standards and Technology, https://www.nist.gov/itl/ai-risk-management-framework.
[84] National Institute of Standards and Technology. Generative AI Profile: NIST AI Risk Management Framework. National Institute of Standards and Technology, https://airc.nist.gov/docs/NIST.AI.600-1.GenAI-Profile.ipd.pdf.

Key Findings

Federal investments in fundamental research have enabled current AI opportunities.

Investments in fundamental research across several agencies, such as the National Science Foundation, National Institutes of Health, Department of Defense, and Department of Energy, have provided the critical knowledge base and much of the technical workforce to advance AI opportunities today.

Continued AI research and evaluation will promote AI advancement.

Many challenges facing AI development and adoption are rooted in unanswered scientific or technical questions. These questions can be widely applicable to many industries and use cases.

Progress in AI R&D is closely linked to access to AI resources.

Researchers need significant computational and data resources to continue progress in AI R&D. This is especially apparent at the frontier of AI development, where developing the most advanced AI models requires costly computational resources and data sets.

A closed AI research ecosystem could limit U.S. competitiveness in AI.

Like many scientific fields, AI R&D has a rich history of open research, collaboration, and shared findings. However, AI researchers at some private firms could reduce the openness of their work for competitive reasons, limiting the AI knowledge that others can build upon.

University AI R&D is necessary but must be paired with vibrant technology transfer activities.

Universities generally lack the means of production necessary to translate their initial research results into marketable products. There is often a wide gap between the basic research conducted at universities and the commercialization activities carried out by industry.

Advancing the science around AI evaluation will help advance adoption.

To deploy AI technologies, users will need to ensure they behave reliably. Different types of evaluation can measure how the AI system performs in specific contexts and use cases.

The U.S. is a global leader in standard setting but faces competitors.

The United States employs an industry-led, bottom-up approach to most standard setting. However, the U.S. faces challenges from nations that could use standards as a tool to gain a competitive advantage for their domestic industries.

Recommendations

Recommendation: Continually monitor and evaluate the impact of AI on different industries and the nation.

The U.S. should establish mechanisms to monitor and evaluate how AI affects different industries and society. This information would be valuable in informing AI policies with the most up-to-date information available. This, in turn, would allow such policies and funding to align with our national objectives and priorities. For example, it might become evident that additional research would be valuable in solving certain problems or addressing new concerns.

Recommendation: Support fundamental R&D for continued leadership in AI innovation.

Fundamental science research in universities drives discoveries that can lead to new or improved technologies while teaching and training the next generation of researchers. Sustained, strategic federal investments in fundamental AI R&D, including through public-private partnerships, will be critical to maintaining U.S. leadership in AI technologies and applications. A disproportionate amount of AI R&D exists within a few large companies. Therefore, federal investments are critical to continuing the fundamental AI research that will ultimately benefit commercial development and the public.

Congress should continue to support the National Science Foundation, Department of Energy, National Institutes of Health, and other science agencies that make grants to universities for AI R&D, including AI-enabled science in the STEM field. Further, to ensure the United States remains the global leader in standardization, Congress should continue R&D investments in critical and emerging technologies such as AI, particularly metrology science, that underpin technical standards.

Recommendation: Increase technology transfer from university R&D to market.

It is difficult for universities to carry their basic research to commercialization. Universities generally lack the means of production necessary to take initial research results and generate marketable products. The pace of American translation of discoveries and inventions from laboratory research to products must accelerate for the U.S. to remain competitive. Closer cooperation among industry, government, and academia could increase technology transfer, stimulate innovation, lead to new products and processes, and expand markets.

Recommendation: Promote public-private partnerships for AI R&D.

Companies conduct a significant amount of America's AI R&D and reasonably keep certain developments as trade secrets that are key to their competitive strategy. However, public-private partnerships for AI R&D and commercialization are an advantage of our national innovation ecosystem. The U.S. should also build off its long history of partnerships between the government (including National Labs), universities, and the private sector to collaborate on finding an appropriate balance between open and closed research. This can, among other things, help avoid federal funding of redundant research with similar work in industry. Congress should support initiatives that support and expand these effective partnerships.

Recommendation: Promote research and standardization surrounding the evaluation and testing of AI.

The science surrounding AI evaluations is rapidly evolving. Evaluations for AI deployed in one use case may not apply to others. Purely technical evaluations of an AI system might not address all challenges, such as sociotechnical problems from the uses of the AI system. The context will be critical to improving evaluations of AI systems deployed in a given industry or for a particular use case. Voluntary standards developed for AI evaluation should consider particular AI use cases. Similarly, organizations and existing sectoral regulators should explore evaluation regimes that are most appropriate for their contexts or situations. Congress should consider supporting activities to improve and standardize evaluations.

Recommendation: Promote the development of infrastructure and data to enable AI research.

Researchers require significant computational and data resources to make continued progress in AI R&D. Resource constraints often impede the ability of academia, small businesses, and others to conduct AI research and utilize state-of-the-art AI systems. Developing a shared public infrastructure of computational resources, data resources, shared testing resources, and software would promote AI R&D in the United States. Federal science agencies should facilitate access to their computational resources and promote greater availability of their data. Federal investment in open-source software libraries and toolkits would also support AI development. Finally, Congress should examine how the NAIRR could provide these critically needed AI resources.

Recommendation: Continue U.S. engagement in international standards development.

The United States is a global leader in standard setting with its industry-led, bottom-up approach. However, the U.S. faces challenges from nations that could use standards as a tool to gain a competitive advantage for their domestic industries. Federal coordination, tracking of federal participation, and increasing federal participation in standard setting can help promote continued U.S. leadership in international standards bodies. Congress should also explore mechanisms to improve U.S. stakeholder engagement in international standard setting, such as grants for small businesses and addressing barriers to convening stakeholder meetings in the United States.

Recommendation: Uphold the U.S. approach to setting standards.

The United States has a long history of setting standards led by multiple industry stakeholders. The U.S. approach protects against poor standards by enabling vibrant deliberation and competition so that the standards with the most technical merit prevail.

Policies that undermine the bottom-up, rules-based, multistakeholder process for standard setting can put American companies at a disadvantage and embolden our adversaries to ostracize U.S. firms from the standards processes. Federal policies concerning voluntary consensus standards, trade, and strategic competitors should not deviate from this approach. The federal government should also work with allies to uphold the United States' open, rules-based approach to international standards bodies.

Recommendation: Align national AI strategy with broader U.S. technology strategy.

AI is one of the most transformative technologies in decades and promises to be a core part of our national interests for years to come. To ensure a government-wide approach to AI development, AI should be considered a part of the national science and technology strategy and similar federal strategies.

The National AI Initiative Act formalizes interagency coordination and strategic planning on AI initiatives. The CHIPS and Science Act directs OSTP to work with the National Science and Technology Council (NSTC) to develop a comprehensive national science and technology strategy every four years to ensure research and development meets our strategic directives.[85] AI should be an explicit part of this national strategy.

[85] Supra 57.

Recommendation: Explore how to accelerate scientific discovery across disciplines with AI.

AI has the potential to accelerate research in all fields of science. Federal science agencies are investing substantially in fundamental AI R&D and its application to various STEM fields. With this experience and expertise, agencies can understand the resource barriers that impede AI R&D and can offer potential solutions that would broaden AI research in the U.S. For example, agencies could study and recommend additional infrastructure investments to better harness AI for scientific discovery.

Interdisciplinary research that combines AI with fields like disease prevention, environmental sciences, and manufacturing can deliver tools to address our most complex challenges in new ways. Federal agencies have historically partnered to solve these problems and achieve common goals. NSF, which funds university research across all non-biomedical disciplines and numerous STEM education programs, should be key in promoting interdisciplinary AI research.

Recommendation: Support AI R&D by small businesses.

Small businesses are the backbone of the U.S. economy. Accordingly, Congress should continue to support small business' capacity to conduct and advance AI R&D. Programs like Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) are serviced by 10+ federal agencies, with extramural R&D budgets exceeding $100 million and $1 billion, respectively.[86] Both programs are designed to allow small businesses to commercialize advanced technological solutions and ultimately promote technology transfer between industry, academia, and the government. Agencies like NSF and DOD should continue contributing significantly to AI-focused SBIR/STTR awards; these programs can support foundational and applied AI research to enhance national security and our R&D infrastructure.[87]

[86] U.S. Small Business Administration. "About the Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) Programs." SBIR.gov, https://www.sbir.gov/about.
[87] National Science Foundation. "Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) Programs." National Science Foundation, https://new.nsf.gov/funding/opportunities/nsf-small-business-innovation-research-small-0.

Recommendation: Encourage international collaboration with likeminded allies and partners on R&D

Global issues, like AI governance, spur international dialogue. While governments like the PRC deploy AI technologies to surveil and control their populations, the United States has an opportunity to lead the world in the responsible and ethical design, development, and deployment of AI technologies. This will benefit from international cooperation on research and standardization. In 2019, the Organization for Economic Cooperation and Development (OECD) Recommendation on Artificial Intelligence included investing in AI R&D as the first recommendation for national policies and international cooperation. The United States has signed onto this commitment, and it remains a major focal point for international cooperation on AI R&D.

The United States should continue to engage and lead in these international fora to demonstrate an interest in R&D cooperation and foster a global culture of responsible AI development that respects applicable international law, individual privacy, and human rights. Congress should support federal AI R&D efforts that build public confidence in AI technology and foster shared values and social priorities with like-minded nations. Related activities could include the U.S. Science Envoy Program and other federal science programs that support international cooperation, facilitate shared infrastructure, and ensure U.S. access to international talent.

CIVIL RIGHTS & CIVIL LIBERTIES

Background

AI systems, especially those used to automate complex or intensive processes, have brought undeniable benefits to our daily lives. However, if used without the proper design, understanding, and safeguards, AI systems can also cause harm.

An AI model, and software systems more generally, can produce misleading or inaccurate outputs. Acting or making decisions based on flawed outputs can violate laws, exacerbate harms, or create new ones.

Some harms can potentially have wide-ranging effects across large segments of the population; they can even constitute crimes or violations of constitutional rights, civil rights, or civil liberties. Particularly in situations that implicate fair and equal access to government services and benefits, it is essential to ensure that the federal government lawfully protects the public's civil rights and liberties.

Adverse effects from flawed or misused technologies are not new developments but are consequential considerations in designing and using AI systems. Businesses, government, and law enforcement agencies have used technologies with inaccuracies or flawed designs that have affected American's civil rights. In one example, faulty facial recognition technology used by law enforcement has led to wrongful arrests.[1]

[1] Thaddeus Johnson et al, "Facial recognition systems in policing and racial disparities in arrests," Government Information Quarterly, Volume 39, Issue 4, 2022, 101753, https://www.sciencedirect.com/science/article/abs/pii/S0740624X22000892?via%3Dihub.; see also: Tate Ryan-Mosely, "The new lawsuit that shows facial recognition is officially a civil rights issue", MIT Technology Review, 14 April 2021, https://www.technologyreview.com/2021/04/14/1022676/robert-williams-facial-recognition-lawsuit-aclu-detroit-police/.

One state's software system, designed to detect unemployment insurance fraud, produced numerous improper fraud charges and financial penalties.[2] Finally, a retailer's use of a flawed facial recognition system improperly identified people as shoplifters.[3]

Concerns about these harms were being considered in 2020 when Executive Order. 13960, "Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government," was signed by President Trump.

That order stated:

> *"Agencies are encouraged to continue to use AI, when appropriate, to benefit the American people. The ongoing adoption and acceptance of AI will depend significantly on public trust. Agencies must therefore design, develop, acquire, and use AI in a manner that fosters public trust and confidence while protecting privacy, civil rights, civil liberties, and American values, consistent with applicable law and the goals of [my earlier] Executive Order, 13859."*[4]

In 2022, the Biden Administration's Office of Science and Technology Policy (OSTP) released the *Blueprint for an AI Bill of Rights*, a white paper positioned as non-binding and not constituting U.S. government policy.[5] This document lays out a number of policies that could support and protect civil rights and civil liberties, as well as promote U.S. values in the design, deployment, and governance of AI systems.

Deprivation of Americans' constitutional rights, civil rights, and civil liberties through automation and automated decisions is a serious concern that deserves in-depth examination. While the House AI Taskforce has endeavored to examine AI policy issues through multiple lenses, further exploration of the potential pitfalls associated with AI adoption is warranted in many cases, including civil rights and civil liberties. Committees with jurisdiction over these issues should consider areas implicating Americans' rights, including law enforcement adoption of AI-enabled technology. When AI adoption can impact these rights and liberties, time and resources should be invested in exploring the scope of the problem, considering potential pitfalls, and identifying solutions. For more information on elections and First Amendment implications from synthetic content, please see the **Content Authenticity** chapter.

[2] Alejandro De La Garza, "States' Automated Systems Are Trapping Citizens in Bureaucratic Nightmares With Their Lives on the Line," Time, 28 May 2020, Tihttps://time.com/5840609/algorithm-unemployment/me,

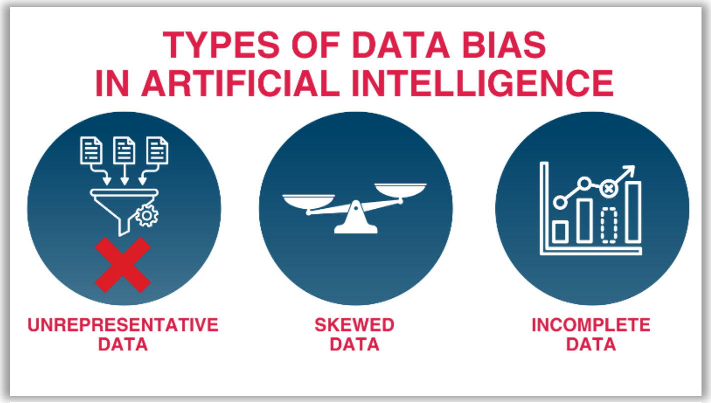
[3] Federal Trade Commission. "Rite Aid Banned from Using AI Facial Recognition after FTC Says Retailer Deployed Technology without Reasonable Safeguards." Federal Trade Commission, Dec. 2023, www.ftc.gov/news-events/news/press-releases/2023/12/rite-aid-banned-using-ai-facial-recognition-after-ftc-says-retailer-deployed-technology-without.

[4] Executive Office of the President. "Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government." Federal Register, 8 Dec. 2020, www.federalregister.gov/documents/2020/12/08/2020-27065/promoting-the-use-of-trustworthy-artificial-intelligence-in-the-federal-government. This E.O. builds on Executive Order 13859, issued in 2019 and available at https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence.

[5] AI Bill of Rights. The White House, https://www.whitehouse.gov/ostp/ai-bill-of-rights/.

AI Can Lead to Flawed Actions and Decision Making

Because AI models are data-driven, problems with training data can cause models to perform differently than designed or expected. Specifically, the training data used can be imbalanced, incomplete, or otherwise limited, sometimes in subtle and difficult-to-detect ways. Training data can include historical patterns that have not been subjected to sufficiently rigorous scrutiny or analysis by AI model trainers. These flaws in training data can fail to account for important nuances or misrepresent particular groups or types of decisions.



Source: Proxima - Addressing Bias in AI for Medical Applications: A Comprehensive Guide

It is essential that training data is representative and is otherwise high quality for its intended purpose. If not, a model trained on data with these shortcomings can produce inaccurate, misleading, or otherwise flawed outputs. AI models can produce outputs with flaws that are skewed or uneven in that they disproportionately affect one or more groups of people. Using these kinds of flawed outputs in decision-making could risk engaging in bias and discrimination against groups, including protected classes such as race, sex, or veteran status.

One major set of risks caused by improper design and use of AI systems is harmful bias, which can occur when an algorithm produces results that are systemically skewed. Bias can be introduced into an AI system when it is created, or it can emerge because of the manner in which the AI system is used.

The chance of producing such a skewed model is amplified when the training data of an AI model reflects historical bias,[6] does not adequately represent certain groups or otherwise exhibits a statistical distribution that does not comport with the desired use of the model.

[6] Leonardo Nicoletti and Dina Bass, "Humans Are Biased. Generative AI Is Even Worse." Bloomberg, 9 June 2023, www.bloomberg.com/graphics/2023-generative-ai-bias/.

The National Institute of Standards and Technology (NIST) has described three categories of bias that can influence AI systems and their output.[7] First, systemic biases result from the procedures and practices of particular institutions, which may not be consciously discriminatory but may have disadvantaged certain social groups. These biases can then be reflected in datasets used to train AI systems and left unaddressed by the norms and practices of AI development and deployment.

Second, statistical and computational biases result from errors that occur when the data the AI system is trained on is not representative of relevant populations. These biases arise when algorithms are trained on one type of data and cannot accurately extrapolate beyond that.

Finally, human biases can result from common cognitive phenomena such as anchoring bias, availability heuristic, or framing effects that arise from adaptive mental shortcuts but can lead to cognitive bias. These errors are often implicit and affect how an individual or group perceives and acts on information.[8]

> ***“When discussing bias in AI it is important to keep in mind that not all bias is harmful and not all AI bias is due to human bias.***

Biases in AI systems can contribute to harmful actions or negative consequences and produce unwarranted, undesirable, or illegal decisions. Examples include decisions disadvantaging people based on one or more protected characteristics of that person, such as a person’s race, sex, or veteran status.

While discrimination based on protected classes is illegal, even when bias does not directly violate a law, it can still be harmful if that system is used to make consequential decisions like whether to hire or fire people, how to diagnose a disease, or whether to grant an individual credit. The harm resulting from bias and any improper use of or reliance upon AI systems is unjustifiable, whether or not existing laws prohibit it.

When discussing bias in AI, it is important to keep in mind that not all bias is harmful, and not all AI bias is due to human bias.

First, not all bias is inherently harmful. Statistical and computational biases that arise in an analysis are a normal and expected part of data science, machine learning, and some of the most popular contemporary AI technologies.

[7] Reva Schwartz et al., “Towards a Standard for Identifying and Managing Bias in Artificial Intelligence,” National Institute of Standards and Technology, March 2022, https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1270.pdf.

[8] Bettina J. Casad, et al., "Confirmation Bias." Britannica, 2022, www.britannica.com/science/confirmation-bias.

For example, a model trained to detect a rare disease may be trained with limited data available, and that data may over- or under-represent segments of the true population. That bias could lead to misdiagnoses for some segments of the population. However, if users understand these limitations, they can account for the bias and use the tool appropriately to mitigate the risk of harm.

Second, biased systems are not necessarily the result of human bias in the design or use of the system. As the example above demonstrates, biased systems can arise from a variety of factors, including limited input data. A model may also perpetuate bias in ways other than by learning directly from data that includes protected characteristics. For example, using a variable closely correlated with a protected class in model training can serve as an unintended proxy for that protected class.[9] These kinds of correlations between training data and protected classes can lead to or perpetuate discrimination.

Enforcement

Several agencies are committed to enforcing the federal laws that protect civil rights, non-discrimination, fair competition, consumer protection, and equal opportunity. In April 2023, the Justice Department's Civil Rights Division, the Consumer Financial Protection Bureau, the Equal Employment Opportunity Commission, and the Federal Trade Commission jointly pledged to uphold America's commitment to core principles of fairness, equality, and justice and to enforce existing civil rights laws in automated systems including systems involving AI.[10] They were joined in this pledge in April 2024 by the Department of Education, Department of Health and Human Services, Department of Homeland Security, Department of Housing and Urban Development, and Department of Labor.[11]

These agencies enforce federal laws such as the Civil Rights Act, Fair Housing Act, Americans with Disabilities Act, and Fair Credit Reporting Act. This pledge states, "existing legal authorities apply to the use of automated systems and innovative new technologies just as they apply to other practices." However, this may be challenged given recent Supreme Court rulings that curtail agencies' ability to interpret ambiguous statutes.[12]

Existing agency authorities may not always be clearly applicable to the type of discrimination at issue. Particularly in the case of emerging technologies, regulators may lack the resources or expertise to adequately enforce existing laws in new contexts engendered by emerging technology.

[9] For example, zip code, which is not a protected class, can be correlated with race, a protected class.

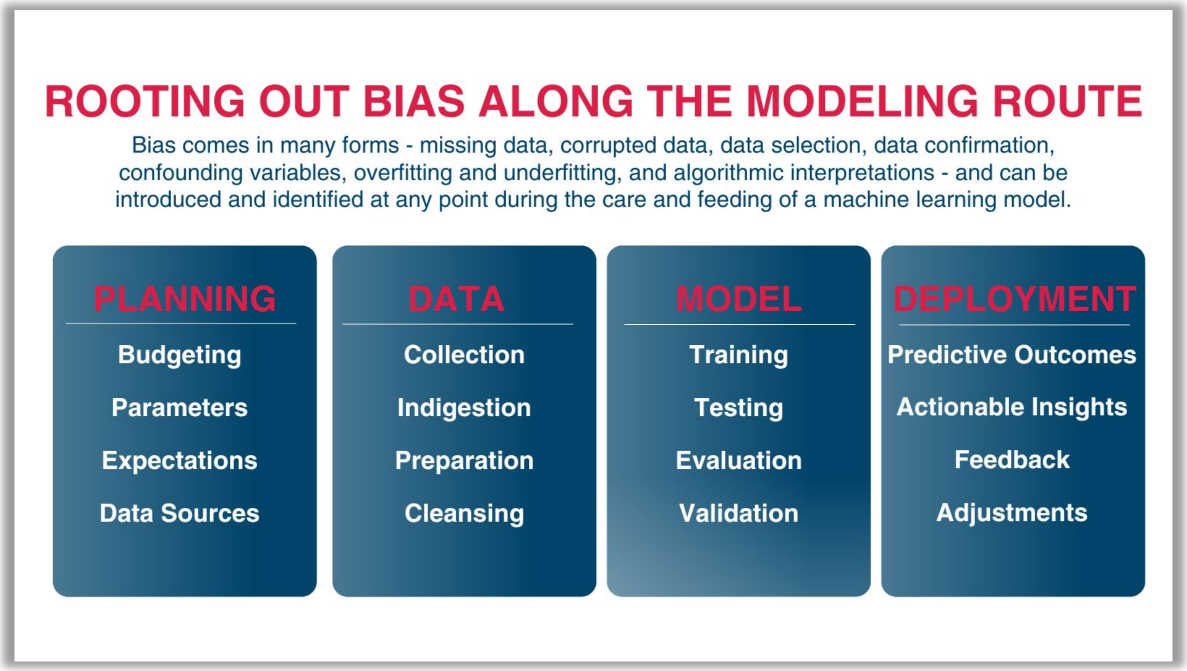
[10] U.S. Department of Justice, "Joint Statement on Enforcement Efforts Against Discrimination and Bias in Automated Systems," U.S. Department of Justice, www.justice.gov/crt/page/file/1581491/dl?inline.

[11] U.S. Department of Justice, "Joint Statement on Enforcement of Civil Rights, Fair Competition, Consumer Protection, and Equal Opportunity Laws in Automated Systems." U.S. Department of Justice, 4 April 2022, www.justice.gov/crt/media/1346821/dl?inline.

[12] Loper Bright Enterprises v. Raimondo, (No. 22-451), 603 U.S. ___ (2024), 144 S. Ct. 2244, (6/28/2024); Relentless, Inc. v. Department of Commerce, (No. 22-1219), 62 F.4th 621 (1st Cir. 2023), slip. op. at 1., 603 U.S. ___ (2024), (7/30/2024). https://www.supremecourt.gov/opinions/23pdf/22-451_7m58.pdf .

Further, many AI systems are highly complex, and the entity that creates or deploys an AI system may have much more information about these systems than those affected. Consequently, regulators and the public often cannot determine the specific factors that algorithmic systems rely upon.

Without sufficient transparency into specifically how AI systems generate their outputs, one must evaluate the AI system as it is deployed to determine whether it has the potential to produce discriminatory decisions. It might not always be apparent how AI systems produce their outputs, what roles these outputs play in human decision-making, or how to correct these flaws.



ROOTING OUT BIAS ALONG THE MODELING ROUTE

Bias comes in many forms - missing data, corrupted data, data selection, data confirmation, confounding variables, overfitting and underfitting, and algorithmic interpretations - and can be introduced and identified at any point during the care and feeding of a machine learning model.

PLANNING	DATA	MODEL	DEPLOYMENT
Budgeting	Collection	Training	Predictive Outcomes
Parameters	Indigestion	Testing	Actionable Insights
Expectations	Preparation	Evaluation	Feedback
Data Sources	Cleansing	Validation	Adjustments

Source: TechTarget - 6 ways to reduce different types of bias in machine learning

Key Findings

Improper use of AI can violate laws and deprive Americans of our most important rights.

AI that is flawed, skewed, biased, or improperly used in decision-making can violate laws or deprive Americans of their constitutional rights, civil rights, and civil liberties. Decisions made with the assistance of AI can potentially cause harms, such as bias or discrimination against protected classes. As Congress considers policies regarding AI design, development, and use across sectors and use cases, a core consideration should be mitigating harmful outcomes impacting American's civil rights and civil liberties.

Understanding the possible flaws and shortcomings of AI models can mitigate potentially harmful uses of AI.

It is not always apparent when AI systems produce flawed outputs or what roles flawed outputs play in human decision-making. Evaluating and documenting the specific characteristics and limitations of AI models, particularly in the context of their intended uses, can reveal risks associated with making decisions based on their outputs and help protect civil rights and civil liberties. Bringing more voices into the design and training process may help improve AI systems and address some of these issues.

Recommendations

Recommendation: Have humans in the loop to actively identify and remedy potential flaws when AI is used in highly consequential decision-making.

This is necessary to safeguard the constitutional rights and civil liberties of Americans affected by AI systems, whether those systems are used by government or private entities.

Recommendation: Agencies must understand and protect against using AI in discriminatory decision-making.

Not only must federal agencies understand and protect against harmful AI use internally, but those agencies tasked with overseeing civil rights and consumer protection must also understand and protect against unlawful AI use by others. This is especially important for decision-making in high-impact areas like banking, healthcare, housing, education, employment, law enforcement, and other public services.

Recommendation: Empower sectoral regulators with the tools and expertise to address AI-related risks in their domains.

Existing sectoral regulators are empowered to hold both public and private sector actors accountable for different violations of constitutional rights, civil rights, civil liberties violations, and consumer protection. To address how AI intersects with their jurisdictions, regulators need tools and technical expertise to understand and properly address AI-related risks. Agencies should adopt tools that allow them to evaluate AI-enabled decision-making and identify and quantify potential shortcomings in AI-enabled decision-making. AI knowledge is especially helpful to regulators in domains such as financial services, healthcare, housing, education, employment, and consumer protection, where many novel types of AI decisions are emerging. One possible approach is for agencies with AI expertise to assist regulators in developing specific research programs to understand and mitigate different risks of AI systems across different contexts and use cases.

Recommendation: Explore transparency for users affected by decisions made using AI.

AI technologies play a larger role in decision-making in areas such as banking, healthcare, housing, education, employment, consumer protection, and law enforcement. Within their domain, sectoral authorities should explore the questions surrounding whether and under which circumstances it would be appropriate to inform individuals when consequential decisions are made by a process that substantially uses AI. A similar concern is whether and to what extent individuals should be informed about the characteristics of AI systems used by government or private entities in making consequential decisions. Improved transparency could also help ensure effective oversight of AI by the public and private sectors, especially AI used in election advertisements and to inform law enforcement and judicial decision-making.

Recommendation: Support standards and technical evaluations to mitigate flawed decision-making involving AI systems.

Improved private sector engagement in and the development of industry-led technical standards could help provide a rigorous technical basis to guide the proper use of AI systems in decision-making. AI standards are much less mature than other types of information technology standards, such as cybersecurity. Such standards for AI will be critical to maximizing transparency, control, evaluation, and accountability of AI systems to foster public trust in AI and lower the chances of producing harmful results. These standards would be most helpful if directed to particular use cases of AI systems.

EDUCATION & WORKFORCE

Background

Worldwide demand for science, technology, engineering, and mathematics (STEM) capable workers, especially for AI-related jobs and careers, continues to grow. But while STEM competencies across U.S. sectors are becoming more essential, K–12 mathematics and science scores in the United States are well below those of many other nations and have stagnated.[1] Despite federal and state efforts, the United States has a significant gap in the appropriate talent needed to research, develop, and deploy AI applications—and this gap is growing.[2] Training and educating American learners on AI topics will be critical to continuing U.S. leadership in AI technology and to the country's economic and national security.

Fostering AI and STEM Talent

Training the future generations of talent in AI-related fields starts with AI and STEM education. While AI literacy is generally useful to all students in our increasingly digital world, additional foundational knowledge and skills are needed to prepare students to work in more technical roles in creating and advancing AI technologies, such as AI researchers, AI/Machine Learning (ML) engineers, and data scientists.[3]

[1] National Science Board. “Talent in U.S. and Global STEM Education and Labor Force.” National Center for Science and Engineering Statistics, 2024, https://ncses.nsf.gov/pubs/nsb20243/talent-u-s-and-global-stem-education-and-labor-force#elementary-and-secondary-mathematics-and-science.

[2] Diana Gehlhaus, et al. “The U.S. AI Workforce: An Analysis of the U.S. Artificial Intelligence Labor Market”. Center for Security and Emerging Technologies, 2021, https://cset.georgetown.edu/publication/u-s-ai-workforce/.

[3] U.S. Department of Defense. “2020 DoD Artificial Intelligence Training and Education Strategy.” National Defense Foundation, 2020, https://nwcfoundation.org/wp-content/uploads/2021/02/2020_DoD_AI_Training_and_Education_Strategy_and_Infographic_10_27_20.pdf.

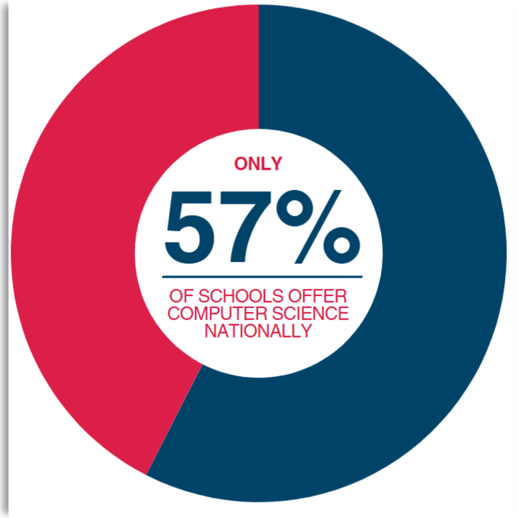
For example, becoming an AI/ML specialist requires students to have core knowledge of operational analysis and mathematics, including trigonometry, linear algebra, calculus, and statistics.[4]

Addressing the future needs of the domestic AI industry must start with fostering AI talent. However, one of the major challenges to fostering domestic AI talent is a widespread lack of basic literacy in STEM concepts.[5]

A 2019 international assessment of mathematics and science showed that United States eighth graders ranked in the middle of education systems in countries with advanced economies, with Singapore, Japan, and others far outpacing the U.S. Since then, recent national assessment scores indicated that mathematics achievement regressed approximately 20 years during the COVID-19 pandemic.[6]

For students seeking other AI-relevant courses, such as computer science, only 57% of U.S. public high schools offer them as opposed to countries where they are widely taught, like China and South Korea.[7] While students earn more mathematics credits in high school and complete more advanced courses than in previous years, their scores on a national mathematics assessment have not improved. Regardless of the potential benefits, efforts to improve science-based training will be challenging if teachers continue to lack the necessary resources.

AI learning should also be nurtured starting from the K–12 level. AI learning has traditionally required advanced programming knowledge that is typically beyond the scope of K–12 settings. However, the emergence of more age-appropriate tools and curricula has enabled educators today to improve the learning process for younger students.[8] Several studies have found the potential to "gamify" the learning experience for younger learners, enabling them to study AI systems in their STEM courses.[9]



Source: NSB Science and Engineering Indicators 2020

[4] U.S. Department of Defense Cyber Workforce. AI/ML Specialist Work Role. U.S. Department of Defense, 2023, https://public.cyber.mil/dcwf-work-role/ai-ml-specialist/.

[5] National Science Board. "Science and Engineering Indicators 2020." National Science Foundation, 2020, https://www.nsf.gov/nsb/publications/2020/nsb202015.pdf.

[6] National Center for Science and Engineering Statistics. Elementary and Secondary STEM Education. National Science Foundation, 2023, https://ncses.nsf.gov/pubs/nsb202331/.

[7] Code.org. "2023 State of Computer Science Education: A Survey of U.S. K–12 Schools." Code.org, 2023, https://code.org/assets/advocacy/stateofcs/2023_state_of_cs.pdf.

[8] Liyao Zou, et al. "A Preliminary Study on the Application of Artificial Intelligence Technology in Meteorological Education and Training." ACM Transactions on Computing Education, https://dl.acm.org/doi/abs/10.1145/3358695.3360939.

[9] Martin, et al. "Systematic review of research on artificial intelligence in K-12 education (2017–2022)." Computers and Education: Artificial Intelligence, ScienceDirect. June 2024, https://www.sciencedirect.com/science/article/pii/S2666920X23000747.

Studies regarding K–12 AI curriculum in the Asia-Pacific region have not only shown a positive influence on students' learning outcomes with various AI concepts such as machine learning, neural networks, and deep learning but have also shown improvement in students' interest in AI courses.[10]

Currently, the most common pathway into the AI workforce is through a four-year degree.[11] Interest in AI-related degrees like computer science is surging—enrollment in computer science degrees grew 249 percent between 2011 and 2020.[12] However, there is some concern that the teaching capacity at universities is not growing proportionally.[13]

Another key challenge is the significant gap in science and engineering degree awards to certain underrepresented groups, including minorities and women. The National Science Board (NSB) describes this issue as "the missing millions"—those who have yet to be engaged in the STEM workforce so that it reflects the region's racial, ethnic, and gender makeup of the general population.[14]

It is critical the United States cultivates STEM talent in every zip code of the country. While progress has been made in recent years,[15] significantly more work is needed to overcome this gap.

Several studies on innovation have clearly demonstrated the benefits of diverse perspectives on innovation capacity and competitiveness.[16] Broadening participation would help provide the talent needed to develop an AI-capable workforce and help ensure the technology is deployed safely and ethically.

As with many technologies, there is a risk that the benefits will extend primarily to communities that are part of its development, raising the possibility that AI could reinforce existing structural, economic, social, and demographic disparities.

[10] Jiahong Su, et al. "A meta-review of literature on educational approaches for teaching AI at the K-12 levels in the Asia-Pacific region." Computers and Education: Artificial Intelligence, vol. 3, 2022, p. 100065, https://doi.org/10.1016/j.caeai.2022.100065.

[11] Rathinam, Sonali, et al. "The U.S. AI Workforce: Analyzing Current Supply and Growth." Center for Security and Emerging Technologies, 2024, https://cset.georgetown.edu/publication/the-u-s-ai-workforce-analyzing-current-supply-and-growth/.

[12] Remco Zwetsloot and Jack Corrigan, "AI Faculty Shortages: Are U.S. Universities Meeting the Growing Demand for AI Skills?" Center for Security and Emerging Technology, July 2022. https://doi.org/10.51593/20190049.

[13] Id.

[14] Supra 5.

[15] National Science Foundation. "NSF Releases Latest Science and Engineering Indicators Report." National Science Foundation News, 2020, https://www.nsf.gov/nsb/news/news_summ.jsp?cntn_id=308617.

[16] Forbes Insights. "Diversity Confirmed to Boost Innovation and Financial Results." Forbes, 15 Jan. 2020, https://www.forbes.com/sites/forbesinsights/2020/01/15/diversity-confirmed-to-boost-innovation-and-financial-results/?sh=45c6ec0c4a6a.

Resources for AI Learning

One key challenge facing the development of domestic AI talent is the lack of access to AI resources, particularly computational power and data, at institutions of higher education.

Even large research institutions do not have the resources to train AI systems of complexity comparable to ChatGPT. Smaller institutions of higher education, minority-serving institutions, community colleges, secondary schools, startups, and small businesses may face deeper challenges in purchasing or otherwise accessing the needed computing resources.[17]

New Workforce Pathways for AI Practitioners

Continued U.S. leadership in AI will require growing the pool of trained AI practitioners, including people with skills in researching, developing, and incorporating AI techniques. This will likely require expanding workforce pathways beyond the traditional educational routes.

AI leadership would also be strengthened by utilizing more of the skilled technical workforce, defined by the National Center for Science and Engineering Statistics as the workforce that is highly skilled in science and engineering fields but does not possess a bachelor's degree or above.[18] Although "AI" was generally not a keyword in job descriptions before 2022, AI-related skills were present in jobs like information technology, data science, and computer engineering.[19]

Now, many workforce pathways, even some not considered technical, are rapidly and continually evolving their training to include explicit AI skills. Interest in AI-related certificates and degrees is growing swiftly,[20] including new industry partnerships with community colleges to train new talent. For example, one Fortune 500 company's AI workforce program expanded to over 85 community colleges in 35 states between 2020 and 2023.[21] However, many companies do not recognize certificates or even associate

[17] Alan Blatecky, et al. Missing Millions: Democratizing Computation and Data to Bridge Digital Divides and Increase Access to Science for Underrepresented Communities. National Science Foundation 2021, https://www.rti.org/publication/missing-millions/fulltext.pdf.

[18] National Science Foundation. Skilled Technical Workforce: Overview and Working Group. National Science Foundation, 2021, https://www.nsf.gov/statistics/stw/docs/skilled-technical-workforce-overview-and-working-group.pdf.

[19] Diana Gehlhaus and Ines Pancorbo, "U.S. Demand for AI Certifications" Center for Security and Emerging Technology, June 2021. https://doi.org/10.51593/20210001.

[20] Jackie Snow. "Students of All Ages Returning to School for AI." GovTech, 2024, https://www.govtech.com/education/higher-ed/students-of-all-ages-returning-to-school-for-ai.

[21] Jobs for the Future. AI-Ready Workforce Report. The Center for Artificial Intelligence & the Future of Work, 2023, https://info.jff.org/hubfs/JFF-AI-Ready%20Workforce%20Report_103123-vF.pdf?utm_medium=email&_hsmi=280455825&utm_content=280455825&utm_source=hs_automation.

degrees as suitable credentials,[22] and the adoption of non-traditional hiring pathways, like skills-based recruitment, remains low but is rising.[23]

The federal government and states are also developing workforce programs to support pathways into AI-related jobs. The Department of Labor (DOL), through its authorities under the National Apprenticeship Act of 1937[24] and the Workforce Innovation and Opportunity Act (WIOA),[25] is the nation's primary federal supporter of industrial workforce development, including apprenticeships. DOL has been working with industry partners to facilitate access to AI-related apprenticeship programs.[26] Several states also support AI-related apprenticeships, such as through tax incentives for employers that offer apprenticeships.[27]

However, a persistent challenge is the lack of a workforce framework for AI that identifies and formalizes standard AI roles, and the skills and competencies needed for those roles. Without a standard framework for reference, there can be significant variation between AI curricula and training programs, exacerbating challenges in validating the skills and competencies of AI job seekers.

Understanding the AI Workforce

Little is currently understood about who makes up the "AI workforce," including its demographic makeup, changes in the workforce over time, employment gaps, and the penetration of AI-related jobs across sectors.[28] Further, AI-related work roles, job categories, tasks, skill sets, and competencies are underdeveloped and often undefined. For example, there are no standard criteria for the tasks, skills, or knowledge that workers need to be able to do testing, evaluation, and analysis of AI systems. Without good data on the AI workforce, it will be difficult to understand the abilities, gaps, and needs of this important workforce segment. Without standardized roles, tasks, and the knowledge and skills to perform those tasks, workforce pathways and professional certifications for AI-related jobs will remain immature and variable.

[22] Supra 19.

[23] Agovino, "Skills-Based Hiring Is Gaining Ground." Society for Human Resource Management, 2024, https://www.shrm.org/topics-tools/news/all-things-work/skills-based-hiring-new-workplace-trend.

[24] United States, Congress. National Apprenticeship Act of 1937. Public Law 75-308, 16 Aug. 1937. United States Statutes at Large, vol. 50, pp. 664–666. GovInfo, https://www.govinfo.gov/content/pkg/STATUTE-50/pdf/STATUTE-50-Pg664.pdf.

[25] Office of the Federal Register, National Archives and Records Administration. Public Law 113 - 128 - Workforce Innovation and Opportunity Act. U.S. Government Publishing Office, 21 July 2014, https://www.govinfo.gov/app/details/PLAW-113publ128.

[26] Exiger. "Exiger and IAA Partner to Launch DOL Registered Apprenticeship Program." Exiger, 2024, https://www.exiger.com/perspectives/exiger-and-iaa-partner-to-launch-dol-registered-apprenticeship-program/.

[27] South Carolina Department of Revenue. TC45: Employee's Withholding Allowance Certificate. South Carolina Department of Revenue, 2023, https://dor.sc.gov/forms-site/Forms/TC45.pdf.

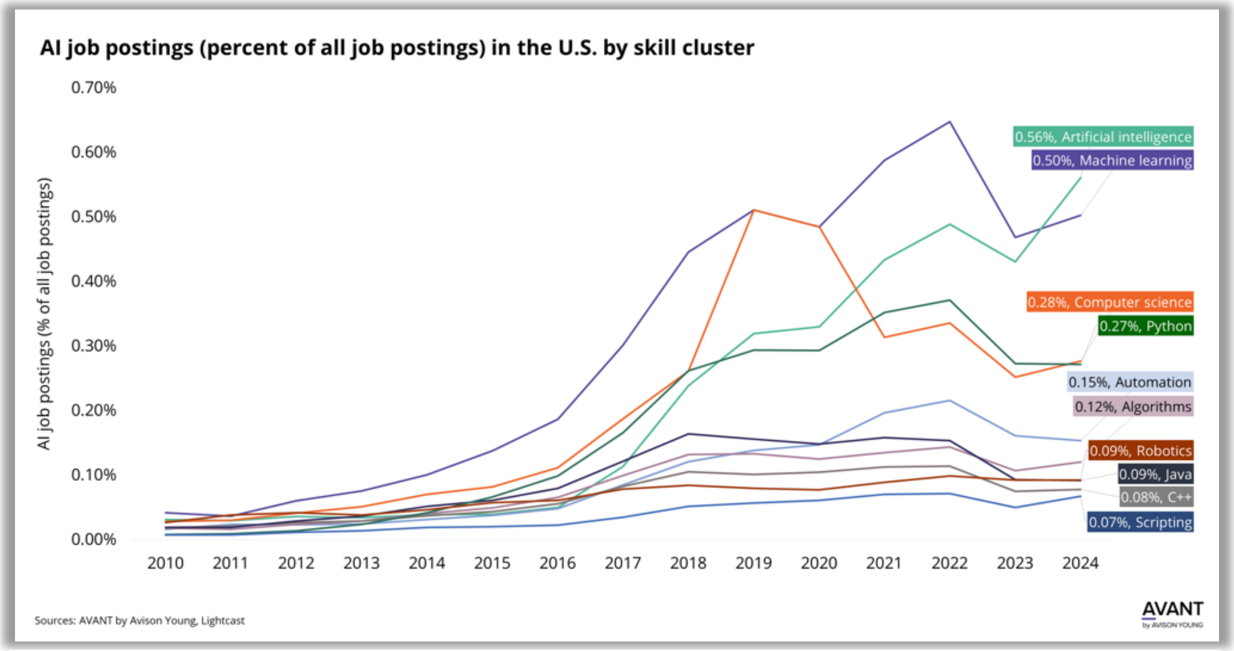
[28] The White House. National Artificial Intelligence Research and Development Strategic Plan: 2023 Update. The White House, May 2023, https://www.whitehouse.gov/wp-content/uploads/2023/05/National-Artificial-Intelligence-Research-and-Development-Strategic-Plan-2023-Update.pdf.

Understanding the AI workforce, including defining roles and skills, is critical for creating educational and talent pipelines for AI.

Updates to workforce training programs and nontraditional hiring pathways are needed to fulfill the growing need for AI practitioners. These programs can help AI job seekers successfully navigate from career education and upskilling programs into the professional AI workforce.

Skills-based hiring will expand talent pools by making it easier for applicants without a bachelor's degree to demonstrate their skills and will help remove barriers to employment for historically underrepresented groups.[29] The United States can strengthen the AI workforce by enabling skills-based hiring, developing an AI workforce framework, and investigating barriers to workforce participation. Standardizing work roles, job categories, tasks, skill sets, and competencies for AI-related jobs will also help enable skills-based hiring. This includes within the federal workforce, as was directed in the *AI in Government Act of 2020,*[30] which is still ongoing.[31]

Further success will also require extensive, well-ordered data analytics to monitor the status of the AI workforce. Supporting better data can help decision-makers focus efforts and investments to address gaps and disparities, including demographic disparities, in the AI workforce. Some of this work is currently being done by the Bureau



Source: Avison Young - Labor demand for AI skills continues to grow across the U.S.

[29] Office of Personnel Management. OPM Releases Skills-Based Hiring Guidance. Chief Human Capital Officers Council, 2024, https://www.chcoc.gov/content/artificial-intelligence-classification-policy-and-talent-acquisition-guidance-ai-government.

[30] Office of the Federal Register, National Archives and Records Administration. Public Law 116 - 260 - Consolidated Appropriations Act, 2021. U.S. Government Publishing Office, 26 Dec. 2020, https://www.govinfo.gov/app/details/PLAW-116publ260.

[31] Supra 29.

of Labor Statistics and the National Center for Science and Engineering Statistics.[32] Congress and the federal government should continue to examine how to improve and streamline such data collection and analysis.

An Evolving AI Workforce

While it is crucial to establish a workforce framework for AI roles, it is also important to stress that the composition of the AI workforce is not static. With AI technology advancing rapidly and being adopted across many sectors, skills for AI practitioners and roles within the AI ecosystem are also quickly evolving. A framework and taxonomy of skills will enable better research and monitoring of emerging skills gaps, improving U.S. leadership in AI.

There are also entirely new roles that have emerged in the AI value chain, and their scope will be heavily influenced by advancing technology and policy. For example, there is an emerging sector of AI auditors who assess and ensure that AI systems are following legal and ethical protocols.[33] However, research into AI measurement, assessment, and assurance best practices is ongoing. Existing roles supporting AI development and integration—such as data center workers and cybersecurity professionals—may also grow or change alongside the technology's evolution.

AI models rely on large and well-curated datasets, which has led to a robust sector of data labeling tasks that require significant human labor. In some cases, this work is outsourced to the Global South, where conditions are rife for worker exploitation.[34] Additionally, some of the most advanced models use "reinforcement learning from human feedback" (RLHF) to align model outputs with human preferences. This technique requires humans to provide feedback that directly "rewards" a model for preferred outputs.

In some cases, this means having low-wage gig workers manually flag undesirable or toxic content to make the model fit for public consumption.[35] Because parts of the internet contain large amounts of disinformation, bias, and hate speech, without intervention, these systems are prone to mimic toxic content. This kind of human feedback requires workers to have appropriate skills and training to interpret the content correctly. As companies seek to improve AI systems on complex topics, there will need to be a commensurate increase in the skills required for humans to provide relevant feedback in RLHF or similar AI training paradigms.

[32] Hill, National Center for Science and Engineering Statistics. Artificial Intelligence & Technology: A National Perspective. National Science Foundation, 2020, https://ncses.nsf.gov/pubs/nsb20205/artificial-intelligence-technology.

[33] Mark Dangelo. "Auditing AI: The emerging battlefield of transparency and assessment." Thomson Reuters, 2023, https://www.thomsonreuters.com/en-us/posts/technology/auditing-ai-transparency/.

[34] Chinasa T. Okolo and Marie Tano. "Moving Toward Truly Responsible AI Development in the Global AI Market." Brookings, 2023, https://www.brookings.edu/articles/moving-toward-truly-responsible-ai-development-in-the-global-ai-market/.

[35] Billy Perrigo. "OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic." Time, 2023, https://time.com/6247678/openai-chatgpt-kenya-workers/.

Facilitating Public-Private Partnerships to Bolster the AI Workforce

American leadership in many critical and emerging technologies has historically been rooted in the U.S. government-university-industry R&D ecosystem and workforce pipelines.

Successful collaborations between educational institutions, government, and industries should effectively align education and workforce development with market needs and emerging technologies. The National Science Foundation's (NSF) Directorate for Technology, Innovation and Partnerships (authorized in the CHIPS and Science Act of 2022[36]) is a key facilitator for these activities.

To foster the AI workforce needed for continued U.S. leadership in AI, partnerships must occur at every level, including with key workforce stakeholders. Greater collaboration between federal, state, and local entities with nonprofits, industry, and education stakeholders can provide useful opportunities for adopting AI tools and resources in pre-K–12 classrooms.

Partnerships can also create effective workforce pipelines from community or technical colleges to data centers and factory floors. This has been proven through programs such as P-TECH (Pathways in Technology Early College High School), a collaboration between IBM, public schools, and two-year institutions, to offer students a high school diploma, an associate degree in STEM, and immediate professional development work.[37] Further, partnerships can provide unique research experiences and support job pathways for students graduating with undergraduate or graduate degrees. Industry can also offer buy-in to training programs through direct support or by recognizing credentials.

There are also opportunities to develop and cultivate regional expertise through government-university-industry partnerships. Different regions will bring their own industries with individual needs, and employer-led workforce development activities may differ across geographies. Supporting such regional partnerships could also facilitate the equitable and broad dispersion of AI training and the economic opportunities it brings.

AI Literacy and Empowering Educators

As society has become more knowledge-based, citizens have had to learn basic digital competencies to compete equally in the workplace.[38] Digital literacy has extended to

[36] Office of the Federal Register, National Archives and Records Administration. Public Law 117 - 167 - An act making appropriations for Legislative Branch for the fiscal year ending September 30, 2022, and for other purposes. U.S. Government Publishing Office, 8 Aug. 2022, https://www.govinfo.gov/app/details/PLAW-117publ167.
[37] P-TECH: A Global Model for Education and Workforce Development. P-TECH, 2023, https://www.ptech.org/.
[38] Lankshear, Colin, et al. Digital Literacies: Concepts, Policies, and Practices. 2010, https://pages.ucsd.edu/~bgoldfarb/comt109w10/reading/Lankshear-Knobel_et_al-DigitalLiteracies.pdf.

new literacies, such as media, computer, data, and now AI literacy.[39] AI literacy involves understanding how AI functions, using it responsibly, and applying it effectively and ethically in various fields. It spans all disciplines, helping people become informed consumers and responsible creators while approaching digital content critically.

For example, a lack of understanding of AI could lead the public to avoid AI products, missing out on productivity-enhancing or quality-of-life-improving uses of the technology. Meanwhile, rampant or reckless use of AI without sufficient understanding of the technology may lead to AI-related mistakes or harms.[40]

Comprehensive AI literacy is fundamental to ensuring that AI is used in ways that maximize benefits and mitigate harms.[41] Age-appropriate learning curricula and teacher education across disciplines can lead to AI-enabled learners through improved AI conceptual understandings in both K–12 and undergraduate students.[42]

AI LITERACY ENCOMPASSES

- Understanding how AI works
- Using AI responsibly
- Recognizing its social and ethical impacts
- Understanding AI's potential benefits and risks and how to mitigate the risks

AI literacy is crucial not only for developing a skilled workforce and positioning our nation as a leader in this critical field but also for mitigating ethical and other challenges associated with AI. As our adversaries gain access to AI technologies for surveillance, weaponization, and economic competition, it is imperative we promote American leadership through an AI-literate public. From early childhood, AI literacy education will positively affect our everyday lives, work productivity, and social circles. Proficient, widespread understanding of AI will mitigate some negative impacts of this technology, such as recognizing harms in the AI-generated content space.

[39] Siu-Cheung Kong, et al. "Evaluation of an artificial intelligence literacy course for university students with diverse study backgrounds." Computers and Education: Artificial Intelligence, vol. 2, no. 1, 2021, p. 100020, https://www.sciencedirect.com/science/article/pii/S2666920X21000205.

[40] Talagala, Nisha. "The Rise of the AI-Enabled Practitioner." Forbes, 2 Feb. 2022, https://www.forbes.com/sites/nishatalagala/2022/02/02/the-rise-of-the-ai-enabled-practitioner/.

[41] TeachAI Policy. TeachAI, 2023, https://www.teachai.org/policy.

[42] Davy Tsz Kit Ng, et al. "Conceptualizing AI literacy: An exploratory review." Computers and Education: Artificial Intelligence, vol. 2, 2021, p. 100041, https://doi.org/10.1016/j.caeai.2021.100041.

Curricula Development

Facilitating the development and implementation of AI-related and AI-enabled curricula will be critical to improving AI literacy. As curricula is developed by the National Science Foundation (NSF), private sector, and civil society, considerations should cover diverse educational formats, including online learning, vocational training, and continuous professional development to meet evolving workforce needs. AI literacy education should also incorporate proper use and technology ethics at every stage, from K–12 to lifelong learners, to promote the responsible use of AI.

Teacher Needs and Guidance

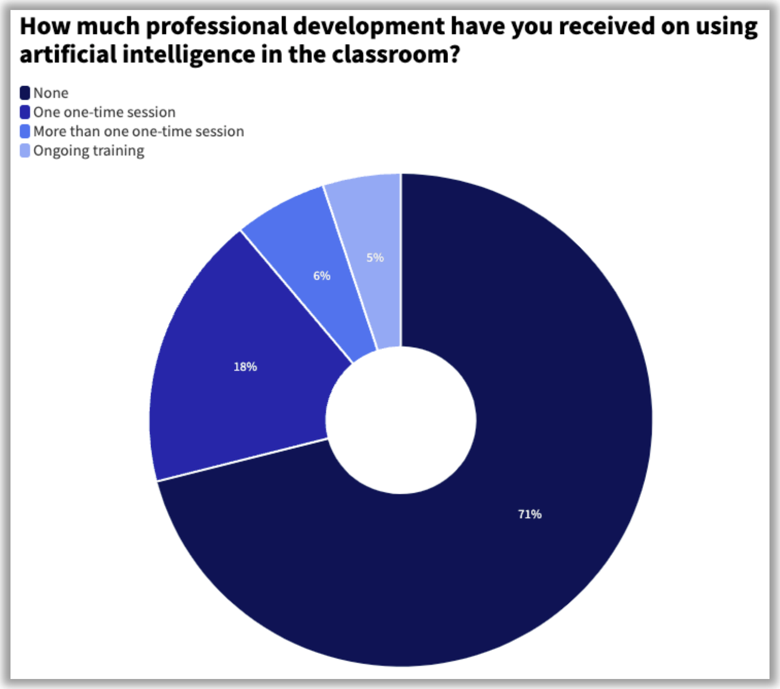
Teachers need knowledge of AI technology to achieve AI literacy and education for students. However, less than a third of teachers have received AI training—one 2024 survey found that 71 percent of K–12 teachers had received no professional learning about using artificial intelligence in the classroom.[43] Educators and administrators simultaneously must learn to operate specific AI products and understand how a broader range of AI systems integrate into their daily activities. Given the individual needs of teachers on a district or state-wide basis, this integration can be highly variable and subject to different jurisdictions and place-based needs.

As AI systems are increasingly integrated into the classroom, educators and administrators will need tools and resources to identify adverse outcomes. Additionally, the distribution of these resources is highly variable across the country, and certain populations, including those with a range of disabilities, may not have access to the positive benefits that AI systems can enable.

[43] Lauraine Langreo, “Teachers Desperately Need AI Training. How Many Are Getting It?,” Education Week, 25 March 2024, https://www.edweek.org/technology/teachers-desperately-need-ai-training-how-many-are-getting-it/2024/03.

Academic integrity is a primary concern among teachers, educators, and administrators at all levels. Some students have used generative AI tools to cheat or plagiarize on their assignments. School districts and universities struggle to write policies to characterize misconduct and govern academic integrity. One of the major challenges in this space is the lack of effective AI detection software. Some detection systems have proven to be inaccurate,[44] while others produce biased outcomes.

For example, some detection systems have inadvertently accused multilingual students because the systems are trained using writing samples from native speakers.[45]



Source: Education Week - Teachers Desperately Need AI Training. How Many Are Getting It?

States are starting to grapple with this technology in their school districts. As of June 2024, only 12 states—Arizona, California, Indiana, Michigan, Mississippi, North Carolina, Ohio, Oklahoma, Oregon, Virginia, Washington, and West Virginia—had issued AI-related guidance from the state department of education or another organization.[46] Districts and higher education institutions have also varied significantly in approaching AI tools. Some systems have banned AI systems (or are reconsidering their bans),[47] some have developed policies for appropriate use,[48] while still others have built their own AI-enabled education tools.[49]

[44] Elkhatat, A.M., Elsaid, K. & Almeer, S. Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. 2023. https://doi.org/10.1007/s40979-023-00140-5.

[45] Weixin Liang et al., "GPT detectors are biased against non-native English writers," Patterns 4, no. 7 (2023), https://doi.org/10.1016/j.patter.2023.100779.

[46] Becky Pringle. Report of the NEA Task Force on Artificial Intelligence in Education. National Education Association, June 2024, https://www.nea.org/sites/default/files/2024-06/report_of_the_nea_task_force_on_artificial_intelligence_in_education_ra_2024.pdf.

[47] Searcey, Dionne. "Despite Cheating Fears, Schools Repeal ChatGPT Bans." The New York Times, 2023, https://www.nytimes.com/2023/08/24/business/schools-chatgpt-chatbot-bans.html.

[48] Shaw, C., Yuan, L., Brennan, D., Martin, S., Janson, N., Fox, K., & Bryant, G., 23 Oct. 2023. Tyton Partners. https://tytonpartners.com/app/uploads/2023/10/GenAI-IN-HIGHER-EDUCATION-FALL-2023-UPDATE-TIME-FOR-CLASS-STUDY.pdf.

[49] "Ed Powered by Individual Acceleration Plan," Los Angeles Unified School District, 2024, http://www.lausd.org/site/default.aspx?PageID=19406.

Because educators and administrators often lack the training and resources to teach AI literacy and to incorporate AI tools into their classrooms, federal, state, and local governments will need to work alongside administrators to enable educators to do so.

Similarly, faculty and administrators will need to adopt new strategies to manage AI use in higher education settings. Institutions can support their faculty by offering clear guidelines and training to enable them to identify benefits and limitations in their courses and determine whether and how to integrate AI into their classrooms and assignments. Federal, state, and local governments should support institutions as they seek to provide these guidelines and training to their faculty.

AI Impact on Workforce

The automation of human jobs has been occurring for centuries amid both promise and concern. The latest technological progression has included machine learning and traditional AI techniques, such as computer vision and automated decision-making processes, which have become commonplace over the past two decades. Using AI to automate tasks across various industries has driven some productivity gains.[50]

However, this automation has also led to the displacement of some jobs that involve repetitive or predictable tasks. As AI has progressed, businesses are increasingly incorporating more advanced AI, including generative AI, into tasks beyond these simpler tasks, including creative work, computer programming, and legal work. While AI may displace some jobs, it will augment existing jobs and create new ones. Some jobs created may require more advanced skills, such as AI system design, maintenance, and oversight; others may require less advanced skills, such as data entry or data labeling. Congress must consider how the U.S. workforce and economy can best incorporate and utilize AI while protecting against potential detriments to existing jobs, workers, and communities.

Understanding Worker Augmentation and Displacement Due to AI

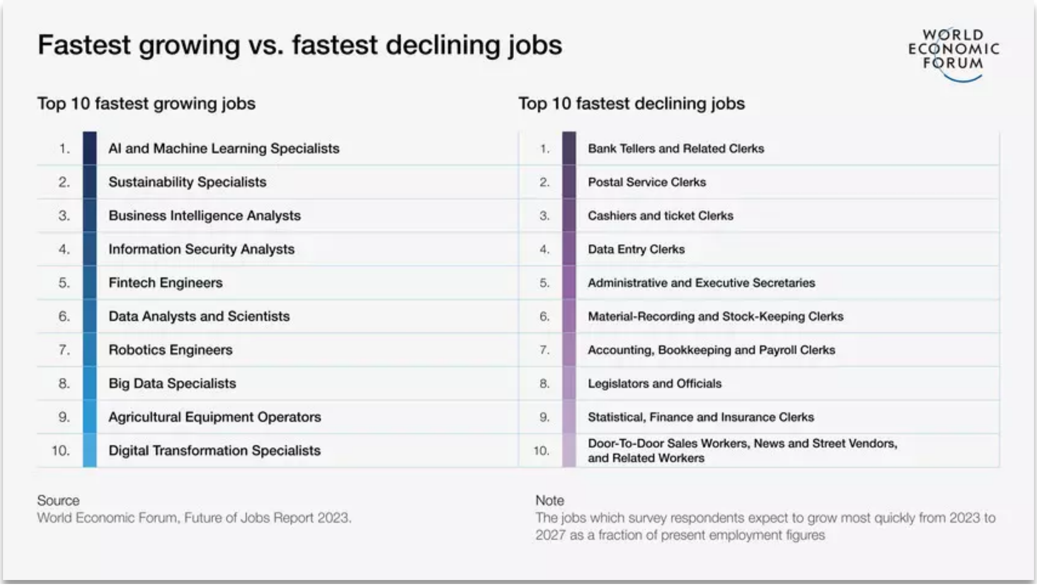
Shifting demographic, social, and economic trends have led to persistent labor shortages, widening skill gaps, and an aging population. In response to talent shortages, businesses across all economic sectors increasingly look to AI-enabled automation and digital transformation to supplement their workforce and increase productivity. It is important to examine how AI will affect the composition and distribution of the workforce.

Generally, AI is expected to augment workers' knowledge and enable them to become more productive. These tools can allow workers to spend more time completing complex problems that draw upon their expertise rather than on mundane tasks. This

[50] OECD/APO, Identifying the Main Drivers of Productivity Growth: A Literature Review, Organisation for Economic Co-operation and Development Publishing, Paris, 2022 https://doi.org/10.1787/00435b80-en.

will require improved AI literacy across the workforce and may require specific up-skilling to perform AI-augmented tasks.

Different sectors will adopt and adapt to AI differently, and there will be sector-specific considerations for AI's ethical and responsible use. For example, a 2024 survey of journalists found that over 70 percent of respondents had started using generative AI in some capacity, but they continue to have concerns about lack of human supervision, inaccurate information, bias, and more.[51]



Fastest growing vs. fastest declining jobs

WORLD ECONOMIC FORUM

	Top 10 fastest growing jobs		Top 10 fastest declining jobs
1.	AI and Machine Learning Specialists	1.	Bank Tellers and Related Clerks
2.	Sustainability Specialists	2.	Postal Service Clerks
3.	Business Intelligence Analysts	3.	Cashiers and ticket Clerks
4.	Information Security Analysts	4.	Data Entry Clerks
5.	Fintech Engineers	5.	Administrative and Executive Secretaries
6.	Data Analysts and Scientists	6.	Material-Recording and Stock-Keeping Clerks
7.	Robotics Engineers	7.	Accounting, Bookkeeping and Payroll Clerks
8.	Big Data Specialists	8.	Legislators and Officials
9.	Agricultural Equipment Operators	9.	Statistical, Finance and Insurance Clerks
10.	Digital Transformation Specialists	10.	Door-To-Door Sales Workers, News and Street Vendors, and Related Workers

Source
World Economic Forum, Future of Jobs Report 2023.

Note
The jobs which survey respondents expect to grow most quickly from 2023 to 2027 as a fraction of present employment figures

Source: World Economic Forum - Fastest growing vs. fastest declining jobs

Some tasks are more likely to be automated by AI, leading to a reallocation of human labor that may displace some workers and industries. The impact of this automation is often felt disproportionately on the lower-wage and low-skilled workers who can least afford to adjust. However, AI tools have also started to displace skilled roles, such as those in some creative and technical industries, traditionally insulated from automation. For example, generative AI tools have been demonstrated to be used in creative professions like graphic design and technical tasks like coding. A 2023 survey of the animation entertainment industry found that 75% of respondents believed generative AI tools have already supported eliminating, reducing, or consolidating jobs in the industry.[52]

[51] Diakopoulos, Nicholas & Cools, Hannes & Li, Charlotte & Helberger, Natali & Kung, Ernest & Rinehart, Aimee & Gibbs, Lisa. "Generative AI in Journalism: The Evolution of Newswork and Ethics in a Generative Information Ecosystem." Research Gate, 2024, https://www.researchgate.net/publication/379668724_Generative_AI_in_Journalism_The_Evolution_of_Newswork_and_Ethics_in_a_Generative_Information_Ecosystem.

[52] CVL Economics. "FUTURE UNSCRIPTED: The Impact of Generative Artificial Intelligence on Entertainment Industry Jobs." Animation Guild, 2024, https://animationguild.org/ai-and-animation/.

Furthermore, not all AI adoption into workforce settings augments or enables skills. Some fields and workplaces have adopted AI in ways that automate complex tasks and leave human workers to perform lower-skilled tasks, a process called "deskilling."

For example, some hospitals have started to explore AI technologies in contexts in which nurses have traditionally played a role, such as treatment decisions.[53] Without careful development and deployment, automating the complex tasks that previously relied on nuanced human decision-making could result in poorer outcomes and a lower-skilled workforce.

It is critical to understand and monitor how skills, jobs, and roles change as AI technologies improve and are increasingly incorporated into the economy. These dynamics will affect long-term educational and vocational pathways into those roles, societal perceptions toward work,[54] and macroeconomic factors like tax revenue.[55]

Training an AI-Enabled Workforce

Harnessing the benefits of AI systems will require a workforce capable of integrating these systems into their daily jobs. The World Economic Forum projects that over the next ten years, 1.1 billion jobs will likely be radically transformed by technologies such as AI.[56] A Microsoft survey notes that more than 82% of business leaders say increased AI use will require new skills from their workers.[57]

According to an IBM survey, executives estimate that 40% of their workforce will need reskilling in response to AI and automation in the next three years, and 87% of executives expect job roles to be "augmented, rather than replaced, by generative AI."[58]

To adapt to AI-augmented roles, incumbent workers will likely need to upskill or improve their performance in their current role to gain AI competency. These shifts in the labor market will also require reskilling or transitioning workers to new roles requiring skills outside their area of focus. To avoid long-term unemployment, displaced workers may benefit from assistance and wage replacement while retraining for an extended period.

[53] National Nurses United. "A.I.'s impact on nursing and healthcare." 2023, https://www.nationalnursesunited.org/artificial-intelligence.

[54] Mirbabaie, M., Brünker, F., Möllmann Frick, N.R.J. et al. The rise of artificial intelligence – understanding the AI identity threat at the workplace. Electron Markets 32, 73–99 (2022). https://doi.org/10.1007/s12525-021-00496-x.

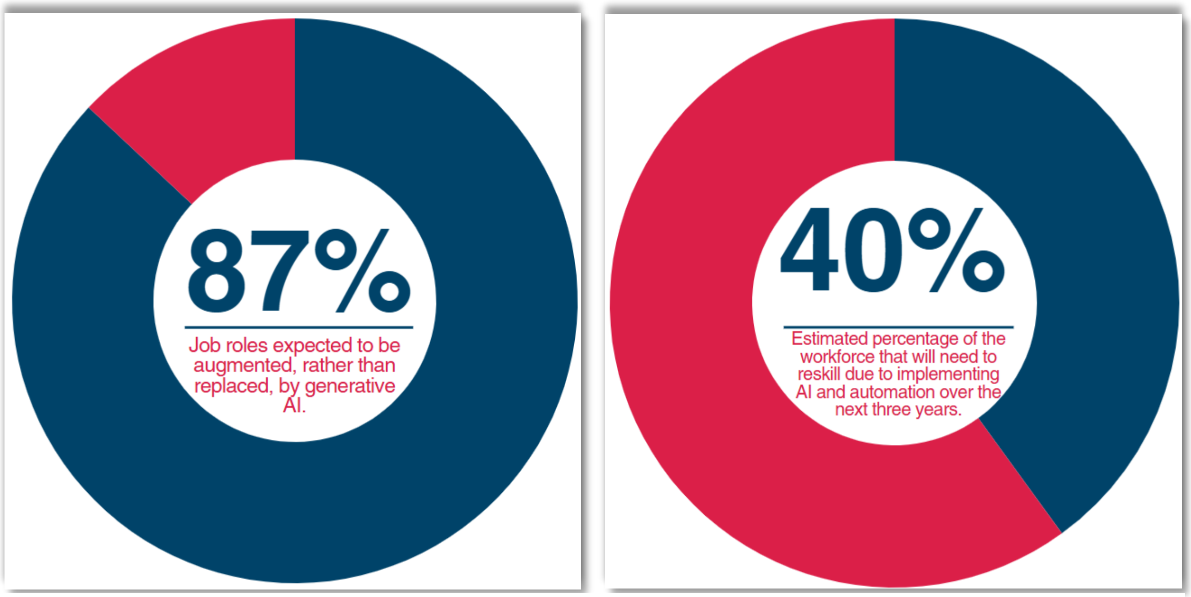
[55] Daron Acemoglu, et al. Taxes, Automation, and the Future of Labor. MIT Sloan School of Management, 2023, https://mitsloan.mit.edu/shared/ods/documents?PublicationDocumentID=7929.

[56] Preston Fore. "To address AI's growing skills gap, IBM is promising to train 2 million people over the next 3 years, Fortune Recommends" Fortune, 29 Sep. 2023, https://fortune.com/education/articles/ibm-commitment-two-million-workers-in-ai-by-2026/.

[57] Id.

[58] IBM. "Research Insights - IBM Institute for Business Value, Augmented work for an automated, AI-driven world." IBM, August 2023, https://www.ibm.com/downloads/cas/NGAWMXAK.

Adopting AI and automation technologies can require significant human infrastructure, including new skills and work routines, changes to physical infrastructure, and changes to cultural norms.



Source: IBM - Augmented work for an automated, AI-driven world

Demand for AI skills has broadly increased since 2010.[59] However, evidence suggests there is currently a large shortage of AI workers and potential workers who can transition into AI roles.[60] Furthermore, this demand is growing across different industry sectors, with higher growth in the information, professional, scientific, and technical services, as well as the finance and insurance sectors.[61]

As AI changes the skills required to do certain jobs, organizations are exploring how to reskill and upskill their existing workforces to remain competitive. For example, Amazon has trained thousands of employees in machine learning through its Machine Learning University program.[62] However, these reskilling and upskilling efforts are challenging and do not guarantee a positive outcome.[63] A 2021 report from the OECD found that

[59] Stanford University. "AI Index Report 2024." Stanford HAI, May 2024, https://aiindex.stanford.edu/wp-content/uploads/2024/05/HAI_AI-Index-Report-2024.pdf.
[60] Diana Gehlhaus and Ilya Rahkovsky. "U.S. AI Workforce, Labor Market Dynamics." Center for Security and Emerging Technologies, 2021, https://cset.georgetown.edu/wp-content/uploads/CSET-U.S.-AI-Workforce-Labor-Market-Dynamics.pdf.
[61] Id.
[62] Amazon Web Services. "Machine Learning and AI Solutions." AWS, 2024, https://aws.amazon.com/ai/machine-learning/.
[63] Jorge Tamayo, et al. "Reskilling in the Age of AI." Harvard Business Review, September 2023, https://hbr.org/2023/09/reskilling-in-the-age-of-ai.

only a fraction of workers participate in these training programs, and those who participate often do not need them.[64]

To achieve the best results across the U.S., such programs may require significant support from federal, state, and local governments and other key partner organizations, such as community and technical colleges, local universities, labor unions, education nonprofits, and workforce development organizations.

AI Use by Employers

Employers are also heavily integrating AI tools into workplace operations. They use AI to manage workers, including assigning tasks, scheduling workers, evaluating performance, hiring, and firing.[65] A 2022 survey of human resources leaders found that 98% planned to rely on software and algorithms to reduce labor costs.[66] However, integrating these technologies into established fields and disciplines can change the conditions and quality of labor in those jobs, including changes to workplace culture.[67]

One challenge with integrating AI into workplace settings is the possibility of errors such as harmful bias, which can occur when an algorithm produces systematically prejudiced results due to erroneous assumptions in the machine learning process.

Bias in automated decision-making, even using non-AI tools, is a recognized problem: in 2015, Amazon acknowledged that its internal AI-driven recruitment technology systematically ranked female applicants' resumes lower than male applicants' because the AI model was trained primarily on male applicants.[68]

The 2022 survey of human resource leaders mentioned previously found only 50% were completely confident that their tech would make unbiased decisions.[69] Though AI has improved substantially over time, a 2023 study found that algorithmic bias in gender, race, and ability of job applicants is still pervasive and stems from limited raw data sets and biased algorithm designers.[70]

[64] OECD (2021), Training in Enterprises: New Evidence from 100 Case Studies, Getting Skills Right, OECD Publishing, Paris, https://doi.org/10.1787/7d63d210-en.

[65] Alexandra Mateescu. "Explainer: Challenging Worker Datafication". Data & Society, November 2023, https://datasociety.net/wp-content/uploads/2023/11/DS_Explainer-Challenging-Worker-Datafication.pdf.

[66] Brian Westfall. "Algorithms Will Make Critical Talent Decisions in the Next Recession—Here's How To Ensure They're the Right Ones." Capterra, 2023, https://www.capterra.com/resources/recession-planning-for-businesses/.

[67] Alexandra Mateescu and Madeleine Clare Elish. "AI in Context: The Labor of Integrating New Technologies." Data & Society, 2019, https://datasociety.net/wp-content/uploads/2019/01/DataandSociety_AIinContext.pdf.

[68] Jeffrey Dastin. "Insight - Amazon scraps secret AI recruiting tool that showed bias against women." Reuters, 2018, https://www.reuters.com/article/usamazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G.

[69] Supra 66.

[70] Chen, Z. Ethics and discrimination in artificial intelligence-enabled recruitment practices. Humanit Soc Sci Commun 10, 567. Nature, 2023. https://doi.org/10.1057/s41599-023-02079-x.

AI technologies are also increasingly being used in workplace surveillance.[71] This includes monitoring and recording video or audio, GPS location, computer activity, and biometric data. This has also included monitoring employees working from home, potentially trespassing into home life.[72] Businesses use this data for productivity tracking and algorithmic discipline, including firings.[73]

Businesses can also use the data they collect to train other AI models, such as improving operational efficiencies and workplace safety.[74]

However, there are concerns that the “datafication” of workers could negatively impact workers and make their workplaces less safe.[75] A 2024 survey of U.S. workers found that workplace monitoring technologies used to discipline workers were strongly related to negative worker health and well-being outcomes.[76] While data-driven analytics may contribute to improved business practices and productivity, these tools may also harm workers’ civil rights or systematically lock certain workers out of employment opportunities altogether.

Understanding Labor Shifts

As job seekers, employers, and education providers seek to keep pace with the evolving skill demands of the economy, timely and granular data on skills and occupations is critical. This crucial work is done by the Department of Labor’s Employment and Training Administration (ETA)—including the O*NET system of occupational characteristics and the Workforce Data Quality Initiative for statewide data—and the Bureau of Labor Statistics (BLS)—including the Employment Projections program and the Work Stoppages program.

Monitoring, understanding, and projecting workforce trends, including wages, is critical for informing educational and social policies and projecting tax revenue. Congress should continue to work with the agencies to monitor shifts in demand across industries, occupations, and skillsets to understand the impact of AI on the workforce, including non-AI jobs.

[71] Wendi Lazar and Cody Yorke. "Watched While Working: Use of Monitoring AI in the Workplace Increases." Reuters, 25 April 2023, https://www.reuters.com/legal/legalindustry/watched-while-working-use-monitoring-ai-workplace-increases-2023-04-25/.

[72] Supra 69.

[73] Colin Lecher. "How Amazon automatically tracks and fires warehouse workers for ‘productivity’." The Verge, 25 April 2019, https://www.theverge.com/2019/4/25/18516004/amazon-warehouse-fulfillment-centers-productivity-firing-terminations.

[74] Brian Warrick, "The Role of Artificial Intelligence in Occupational Safety and Health Practices (OSH)." USF Public Health News, 2024, https://www.usf.edu/health/public-health/news/2024/ai-in-osh-practices.aspx.

[75] Supra 71.

[76] Alexander Hertel-Fernandez. “Estimating the Prevalence of Automated Management and Surveillance Technologies at Work and Their Impact on Workers’ Well-Being.” Washinton Center for Equitable Growth, 2023, https://equitablegrowth.org/research-paper/estimating-the-prevalence-of-automated-management-and-surveillance-technologies-at-work-and-their-impact-on-workers-well-being/.

Existing Programs for Worker Dislocation, Reskilling, and Upskilling

There are several existing programs for workforce development which could be updated to address changes brought forth by AI. These programs include:

- *Programs authorized under the Workforce Innovation and Opportunity Act.*[77] WIOA requires states to strategically align their workforce development programs to coordinate the needs of both job seekers and employers. Funding for WIOA has never reached authorized levels. WIOA's Dislocated Worker Program was funded at $1.095 billion in FY 2023 and only served 212,018 people in program year 2022 (the most recent year with available data).[78] Further, the WIOA Dislocated Worker Program emphasizes short-term training, which is an effective tool for rapid reemployment, but many new jobs created from automation may require higher skill attainment and, thus, longer training. Some of the most common credentials earned through WIOA are for low-wage occupations, such as nursing assistants. The median annual earning of a training program on the Eligible Training Provider list is $29,388.[79]

- *Trade Adjustment Assistance (TAA).* DOL's TAA ended in 2022, but previously, it helped workers dislocated by foreign trade become reemployed in an in-demand industry by providing access to training, employment and case-management services, allowances for job search and relocation, and wage insurance. It cost $582.1 million in FY 2019 and served 28,751 participants.

- *Unemployment Insurance (UI).* DOL oversees the UI program, which provides income support to workers who lose a job through no fault of their own while they seek reemployment.

- *Education and Work-Based Learning Programs.* Although not specifically targeted toward dislocated workers, the DOL's Employment and Training Administration provides training programs and other services, including Skills Training Grants, the Adult Literacy and Education Initiative, and work-based learning programs that provide education, training, and learning opportunities to prepare individuals for future employment.

Additional efforts, including partnerships with state, local, or non-governmental entities, should be considered to support workers and workforce skilling and re-skilling.

[77] Supra 25.

[78] U.S. Department of Labor, Employment and Training Administration. PY 2022 WIOA National Performance Summary. U.S. Department of Labor, 2022, https://www.dol.gov/sites/dolgov/files/ETA/Performance/pdfs/PY2022/PY%202022%20WIOA%20National%20Performance%20Summary.pdf.

[79] David Deming, et al. "Navigating Public Job Training: Harvard Project on Workforce." Harvard Kennedy School, 2023, https://www.pw.hks.harvard.edu/post/publicjobtraining.

Labor Laws for Modern Needs

Employers deploying AI in the workforce must adhere to anti-discrimination and labor laws. Existing law protects employees and job applicants from employment discrimination based on race, color, religion, sex, and national origin, as well as disability discrimination.

The Equal Employment Opportunity Commission (EEOC) has released guidance on employers' use of AI.[80] [81] The EEOC needs continued support to ensure that nondiscrimination protections are fully enforced. Further, Congress should investigate transparency legislation to ensure that AI technologies deployed in the workplace are not violating workers' rights. Please see the chapter on **Civil Rights** for more discussion of advancing civil rights protections in the age of AI.

AI used in the workplace must respect existing workers' and employers' rights under the National Labor Relations Act. U.S. labor and employment laws have governed workplace organizing and bargaining activities for decades. As an example of the broad implications of AI, the National Labor Relations Board (NLRB) has completed work on the interactions of emerging technologies and protected activities or unfair practices under those laws.[82]

The Occupational Safety and Health Act (P.L.91-596) requires employers to provide workers with a workplace free of recognized hazards.[83] The Occupational Safety and Health Administration (OSHA) should monitor its standards or recommended practices and require updating as AI technologies are increasingly incorporated into the workplace. Congress could also consider legislation that supports research through the National Institute of Occupational Safety and Health to examine how AI technologies can enhance or reduce workplace safety and well-being.[84]

Workers will continue to be an important input in the development of AI systems. Congress should ensure that policies overseeing those workers are fair and ethical.

[80] U.S. Equal Employment Opportunity Commission. Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence. U.S. Equal Employment Opportunity Commission, 2023, https://www.eeoc.gov/laws/guidance/americans-disabilities-act-and-use-software-algorithms-and-artificial-intelligence.

[81] U.S. Equal Employment Opportunity Commission. "EEOC Releases New Resource on Artificial Intelligence and Title VII." U.S. Equal Employment Opportunity Commission, 2023, https://www.eeoc.gov/newsroom/eeoc-releases-new-resource-artificial-intelligence-and-title-vii.

[82] National Labor Relations Board. "NLRB General Counsel Issues Memo on Unlawful Electronic Surveillance and Monitoring." National Labor Relations Board, 2023, https://www.nlrb.gov/news-outreach/news-story/nlrb-general-counsel-issues-memo-on-unlawful-electronic-surveillance-and.

[83] United States, Congress. Occupational Safety and Health Act of 1970. Public Law 91-596, 29 Dec. 1970. United States Statutes at Large, vol. 84, pp. 1590–1620. GovInfo, www.govinfo.gov/content/pkg/STATUTE-84/pdf/STATUTE-84-Pg1590.pdf.

[84] Jay Vietas. "The Role of Artificial Intelligence in the Future of Work." Centers for Disease Control and Prevention. NIOSH Science Blog, 24 May 2021, https://blogs.cdc.gov/niosh-science-blog/2021/05/24/ai-future-of-work/.

Congress could explore legislation and oversight activities in their jurisdictions to promote sound job principles.

For example, the federal government could ensure appropriate standards for data labeling or reinforcement learning work done for federal programs and federally funded research.

Key Findings

AI is increasingly used in the workplace by both employers and employees.

It is likely that workers will increasingly work with or alongside AI systems, which will require pathways to upskill an AI-enabled workforce.

Fostering domestic AI talent and continued U.S. leadership will require significant improvements in basic STEM education and training.

Other nations are ahead of the U.S. in K–12 mathematics and science education. Addressing the future needs of the American AI industry will require that AI skills be bolstered by both workforce training and K–12 education.

U.S. AI leadership would be strengthened by utilizing a more skilled technical workforce.

The highly skilled workforce in science and engineering fields, without a bachelor's degree or above, can add to the ranks of U.S. AI talent. This workforce can be trained through AI-related certificate programs and industry training programs.

AI adoption in America requires AI literacy.

A lack of understanding of AI could lead the public to avoid AI products, missing out on productivity-enhancing or quality-of-life-improving uses of the technology.

K–12 educators need resources to promote AI literacy.

To achieve AI literacy and education for students, teachers need knowledge of AI technology, including AI training on using AI in the classroom.

Recommendations

Recommendation: Invest in K–12 STEM and AI Education and broaden participation.

U.S. K–12 STEM education is lagging, especially in comparison to other nations. Continued U.S. leadership in AI will require a targeted focus on improving K–12 STEM and AI education. The federal government has several initiatives to improve STEM education in rural and underserved communities. Congress should continue to support those efforts, especially in areas related to AI.

Recommendation: Bolster U.S. AI skills by providing needed AI resources.

One key challenge facing the advancement of the United States' AI talent pipeline is a lack of access to AI resources, particularly computational power and data, at institutions of higher education. One potential solution is the NAIRR Pilot (discussed in the **Research, Development, &Standards** chapter), which connects U.S. researchers and educators with computational, data, software, model, and training resources.

Recommendation: Develop a full understanding of the AI workforce in the United States.

Understanding the AI workforce, including defining roles and skills, is critical for creating educational and talent pipelines for AI. There currently exists little understanding of who makes up the "AI workforce"—including demographic makeup, changes in the workforce over time, employment gaps, and the penetration of AI-related jobs across sectors. Further, AI-related work roles, job categories, tasks, skill sets, and competencies are underdeveloped and often undefined. Without good data on the AI workforce, it will be difficult to understand the abilities, gaps, and needs of this important workforce segment.

Congress could ensure federal agencies support extensive, well-ordered data analytics to monitor the status of the AI workforce. Supporting better data can help policymakers focus efforts and investments to address gaps and disparities, including demographic disparities, in the AI workforce. Some of this work is currently being done by the Bureau of Labor Statistics and the National Center for Science and Engineering Statistics.[85]

[85] Hill, National Center for Science and Engineering Statistics. Artificial Intelligence & Technology: A National Perspective. National Science Foundation, 2020, https://ncses.nsf.gov/pubs/nsb20205/artificial-intelligence-technology.

Recommendation: Facilitate public-private partnerships to bolster the AI workforce.

The United States can improve its AI-related education and workforce development activities by facilitating public-private partnerships in different regions across the country. Collaborations between educational institutions, federal and state governments, and industries effectively align education and workforce development with market needs and emerging technologies. To foster the AI workforce needed for continued U.S. leadership in AI, partnerships must occur at every level, including with key workforce stakeholders such as employers, training providers, community-based organizations, labor unions, career and technical education organizations, economic development organizations, and other public sector organizations.

Recommendation: Develop regional expertise when supporting government-university-industry partnerships.

Congress should support public-private partnerships in education and workforce development to bolster workforce pathways into AI jobs. Different regions will have their own industries with individual needs, and employer-led workforce development activities may differ across geographies. Supporting regional expertise would facilitate the equitable and broad dispersion of AI training, as well as the economic opportunities it brings. To support these outcomes, Congress should support regional innovation programs at the National Science Foundation, the Department of Energy, and the Economic Development Administration.

Recommendation: Broaden pathways to the AI workforce for all Americans.

The United States can strengthen the AI workforce by enabling skills-based hiring, developing an AI Workforce Framework, and investigating how to eliminate barriers to workforce participation.

Congress should explore how to update workforce training programs and nontraditional hiring pathways to fulfill the rising need for AI practitioners and to ensure AI job seekers can successfully navigate from career education and upskilling programs into the professional AI workforce. Skills-based hiring will expand talent pools by making it easier for applicants without a bachelor’s degree to demonstrate their skills and will help remove barriers to employment for historically under-represented groups.[86]

Congress should develop improved guidance and strategies for the public and private sectors to implement skills-based hiring practices and standards for professional certifications and credentials in AI-related fields and disciplines.

[86] Office of Personnel Management. OPM Releases Skills-Based Hiring Guidance. Chief Human Capital Officers Council, 2024, https://www.chcoc.gov/content/artificial-intelligence-classification-policy-and-talent-acquisition-guidance-ai-government.

Recommendation: Support the standardization of work roles, job categories, tasks, skill sets, and competencies for AI-related jobs.

Legislation should be explored to support a common understanding and lexicon around AI-related tasks and competencies to address these issues. Federal agencies should support extensive, well-ordered data analytics to monitor the status of the AI workforce.

Supporting better data can help policymakers focus efforts and investments to address gaps and disparities, including demographic disparities, in the AI workforce. Legislation could improve and streamline such data collection.

Recommendation: Evaluate existing workforce development programs.

Rapid AI adoption may displace some job roles and lead to job losses. This chapter discusses several existing U.S. federal workforce development programs to address worker dislocation, support retraining, and mitigate other adverse impacts of automation. Congress should evaluate these programs. They may need to be expanded or updated to adequately address new challenges brought forth by AI.

Congress should also consider whether other programs, including partnerships with state, local, or non-governmental entities, should be used to implement workers and workforce skilling and re-skilling.

Recommendation: Promote AI literacy across the U.S.

AI literacy is critical in empowering the American public to responsibly use and respond to AI technologies. The United States will need to promote a basic understanding of AI technologies and their societal impact, akin to basic digital literacy. Continued federal support for existing programs and methods to equitably scale AI to classroom settings will be critical to incorporating AI into our country's school systems.

Federal agencies should provide support for professional development and teacher preparation on using these technologies in the classroom, especially in a manner consistent with the student's learning goals. NSF and ED both have programs focused on enabling educators and administrators to have this training, such as NSF's Teacher Corps, NSF's Computer Science for All, and ED's Teacher Quality Partnership.

Recommendation: Empower U.S. educators with AI training and resources.

Congress should explore leveraging federal funds, such as Every Student Succeeds Act[87] Title II-A and IV-A funds, for professional learning and educator development in AI literacy. Continued federal support for these programs will enable educators to incorporate AI into our country's K–12 school systems.

State education agencies and school districts often go through time- and energy-intensive efforts to choose products for their school systems. Congress and the Administration should support these policymakers in choosing these products. The budget implications of using certain AI-enabled educational tools can also be unclear and challenging to quantify. Programs such as ED's What Works Clearinghouse[88] can be useful models to guide evaluation efforts to help inform these decisions.

Recommendation: Support NSF curricula development.

Congress should support the NSF's efforts to promote curricula development for AI-related fields through competitive awards for institutions of higher education, industry consortiums, and education nonprofits. Congress should also explore other mechanisms to support the adoption of these curricula, such as through programs supported by ED and partnerships between state educational entities, federal entities, industry, education nonprofits, and institutions of higher education.

Recommendation: Monitor the interaction of labor laws and worker protections with AI adoption.

As AI continues to drive increased worker productivity and economic prosperity, the entire economy should share those productivity gains. Employers deploying AI in the workforce must adhere to existing anti-discrimination and labor standards. However, there may be gaps in understanding how existing laws apply to new AI technologies and how to enforce those laws. Congress should monitor these developments to ensure existing laws adequately address worker and employer needs in the modern age.

[87] Office of the Federal Register, National Archives and Records Administration. Public Law 114 - 95 - Every Student Succeeds Act. U.S. Government Publishing Office, 9 Dec. 2015, https://www.govinfo.gov/app/details/PLAW-114publ95.

[88] Institute of Education Sciences, "What Works Clearinghouse (WWC)," IES: Institute of Education Sciences, U.S. Department of Education, https://ies.ed.gov/ncee/wwc/FWW.

INTELLECTUAL PROPERTY

Background

Advances in AI technology have introduced new issues for intellectual property (IP) laws, raising questions on how the ownership, creation, and protection of art, writings, brands, inventions, and other creations should be treated.

Generative AI technology has become more adept at tasks resembling human creativity, raising further questions about how works created using generative AI should be treated under our IP laws.

Four main categories of intellectual property rights exist in the United States: patents, trademarks, copyrights, and trade secrets.[1] Patents and copyrights are generally protected under federal law.[2] In contrast, trademarks and trade secrets are protected under both state and federal laws.[3]

[1] Verigan, Teresa. "Protecting Intellectual Property in the United States: A Guide for Small and Medium-Sized Enterprises in the United Kingdom." U.S. Patent and Trademark Office, 15 March 2018, www.uspto.gov/sites/default/files/documents/UK-SME-IP-Toolkit_FINAL.pdf.

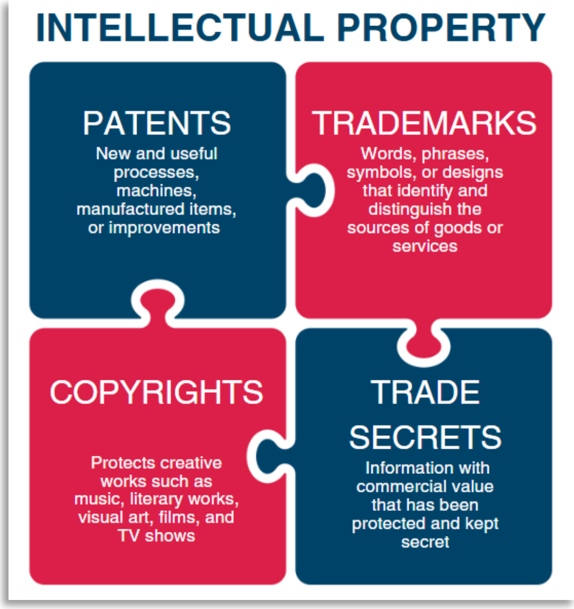
[2] House of Representatives, Congress. 17 U.S.C. 301 - Preemption with respect to other laws. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title17/USCODE-2023-title17-chap3-sec301; see also: Bonito Boats, Inc. v. Thunder Craft Boats, Inc., 489 U.S. 141, 152 (1989) ("Thus our past decisions have made clear that state regulation of intellectual property must yield to the extent that it clashes with the balance struck by Congress in our patent laws.").

[3] Zirpoli, Christopher T. "Artificial Intelligence Prompts Renewed Consideration of a Federal Right of Publicity." Congressional Research Service, 29 Jan. 2024, https://crsreports.congress.gov/product/pdf/LSB/LSB11052. ("Existing federal IP laws provide examples of both approaches, as the Patent Act and Copyright Act largely preempt state laws while the Lanham Act and Defend Trade Secrets Act do not.").

Patents

Utility patents protect new and useful processes, machines, articles of manufacture, compositions of matter, or improvements thereof.[4]

Under Title 35, such an invention must be novel and cannot already exist in preceding technology or knowledge ("prior art") to be patentable.[5] Further, the claimed invention must not be merely an obvious improvement over the prior art.[6] Additionally, the patent document itself must meet certain requirements with respect to disclosures, level of detail, and specificity of language, among other things.[7]


INTELLECTUAL PROPERTY
PATENTS
New and useful processes, machines, manufactured items, or improvements
TRADEMARKS
Words, phrases, symbols, or designs that identify and distinguish the sources of goods or services
COPYRIGHTS
Protects creative works such as music, literary works, visual art, films, and TV shows
TRADE SECRETS
Information with commercial value that has been protected and kept secret

These requirements reflect the fundamental nature of a patent as a bargain in which innovators are incentivized to create truly new inventions by granting them certain legal rights for a limited period of time, and society benefits from the required public disclosure of those inventions to enable other innovators to learn from them and develop further innovations.[8]

U.S. patents are obtained by filing a patent application with the U.S. Patent and Trademark Office (USPTO), which examines the application to determine whether it meets all relevant statutory criteria (such as those described above) before granting the patent.[9] Once issued, patents generally have a term of 20 years from the date the application was originally filed with the USPTO.[10]

[4] House of Representatives, Congress. 35 U.S.C. 101 - Inventions patentable. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title35/USCODE-2023-title35-partII-chap10-sec101; see also: Supra 1.

[5] House of Representatives, Congress. 35 U.S.C. 102 - Conditions for patentability; novelty. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title35/USCODE-2023-title35-partII-chap10-sec102

[6] House of Representatives, Congress. 35 U.S.C. 103 - Conditions for patentability; non-obvious subject matter. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title35/USCODE-2023-title35-partII-chap10-sec103.

[7] House of Representatives, Congress. 35 U.S.C. 112 - Specification. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title35/USCODE-2023-title35-partII-chap11-sec112.

[8] Amgen, et al. v. Sanofi, et al., 143 S.Ct. 1243, 1251 (2023) ("In exchange for bringing 'new designs and technologies into the public domain through disclosure,' so they may benefit all, an inventor receives a limited term of 'protection from competitive exploitation.'").

[9] House of Representatives, Congress. 35 U.S.C. 2 - Powers and duties. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title35/USCODE-2023-title35-partI-chap1-sec2; see also: House of Representatives, Congress. 35 U.S.C. 131 - Examination of application. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title35/USCODE-2023-title35-partII-chap12-sec131.

[10] House of Representatives, Congress. 35 U.S.C. 154 - Contents and term of patent; provisional rights. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title35/USCODE-2023-title35-partII-chap14-sec154.

Copyrights

Copyrights protect original works of authorship during the author's lifetime and for a limited time afterward.[11] Copyright is a form of IP that protects creative works such as music, literary works, visual art, and audiovisual works like films and TV shows.[12] As soon as a work is created and fixed in a tangible medium (e.g., written down or recorded), the author/artist automatically has a copyright on that work.[13]

Importantly, the scope of copyright protection is limited to the expressive portions of a work and does not extend to "any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work."[14] Copyright protection is also limited unless the copyright is registered with the U.S. Copyright Office (USCO).[15]

Copyright registration enables the copyright owner to file an infringement lawsuit and to access the full scope of available remedies.[16]

A copyright confers on its owner a "bundle" of rights enumerated under the Copyright Act.[17] Among those rights are the right to "reproduce the copyrighted work in copies or phonorecords" and to "prepare derivative works based upon the copyrighted work."[18] To reproduce a work or prepare derivative works (or other protected acts), permission (i.e., a license) must first be obtained from the copyright owner.[19] Whether the work is publicly accessible or freely available for certain uses (e.g., for viewing) has no legal effect on whether a license is required for other uses (e.g., reproduction/copying, derivative works).[20]

[11] Supra 1.
[12] House of Representatives, Congress. 17 U.S.C. 102 - Subject matter of copyright: In general. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title17/USCODE-2023-title17-chap1-sec102.
[13] Id.
[14] Id.
[15] House of Representatives, Congress. 17 U.S.C. 401 - Notice of copyright: Visually perceptible copies. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title17/USCODE-2023-title17-chap4-sec401; see also: House of Representatives, Congress. 17 U.S.C. 502 - Remedies for infringement: Injunctions. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title17/USCODE-2023-title17-chap5-sec502; see also: House of Representatives, Congress. 17 U.S.C. 503 - Remedies for infringement: Impounding and disposition of infringing articles. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title17/USCODE-2023-title17-chap5-sec503; see also: House of Representatives, Congress. 17 U.S.C. 504 - Remedies for infringement: Damages and profits. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title17/USCODE-2023-title17-chap5-sec504; see also: House of Representatives, Congress. 17 U.S.C. 505 - Remedies for infringement: Costs and attorney's fees. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title17/USCODE-2023-title17-chap5-sec505
[16] Id.
[17] House of Representatives, Congress. 17 U.S.C. 106 - Exclusive rights in copyrighted works. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title17/USCODE-2023-title17-chap1-sec106.
[18] Id.
[19] Id. ("the owner of copyright under this title has exclusive rights to do and to authorize . . .")
[20] Id.; see also: House of Representatives, Congress. 17 U.S.C. 201 - Ownership of copyright. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title17/USCODE-2023-title17-chap2-sec201.

In addition to the “idea” limitation noted above, another important limitation on the rights of a copyright owner relevant to AI is “fair use,” namely, unlicensed use of a copyrighted work defined as non-infringing in light of an analysis of certain statutory factors.[21]

The four statutory factors in the fair use analysis are (1) the nature of the copyrighted work; (2) the amount and substantiality of the portion used in relation to the work as a whole (i.e., how much is used and how important it is to the work); (3) the effect of the use on the value of the work; and (4) the purpose and character of the use (e.g., commercial purpose, educational purpose, research purpose, criticism/commentary/parody, etc.).[22] Among other things, the fair use exception prevents copyrights from being used to stifle First Amendment-protected free speech.[23]

Trade Secrets

Trade secrets include any information maintained as a secret, having commercial value, and for which reasonable steps have been taken to protect and keep them secret.[24] Unlike patents and copyrights, a trade secret does not expire as long as the required conditions are met.[25]

Unlike patents, for which infringement is based on strict liability (i.e., there is no requirement for a “bad” act or intent), a trade secret claim only arises when the trade secret is acquired through “improper” means, such as theft, misrepresentation, breach of duty to maintain secrecy, or espionage.[26] For example, no claim arises when a trade secret is acquired through independent discovery or reverse engineering.[27]

Trademarks

Trademarks (for goods) and service marks (for services) are words, phrases, symbols, or designs, or a combination thereof, that identify and distinguish the sources of goods or services.[28] Similar to copyrights, full protection under federal trademark law is only available if the mark is registered with the USPTO, which requires that the mark must be used, or intended to be used, in commerce.[29] Unlike patents and copyrights, marks do not expire as long as required renewal actions are taken.[30]

[21] House of Representatives, Congress. 17 U.S.C. 107 - Limitations on exclusive rights: Fair use. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title17/USCODE-2023-title17-chap1-sec107.
[22] Id.
[23] Eldred v. Ashcroft, 537 U.S. 186, 219 (2003) (“In addition to spurring the creation and publication of new expression, copyright law contains built-in First Amendment accommodations . . . the ‘fair use’ defense allows the public to use not only facts and ideas contained in a copyrighted work but also expression itself in certain circumstances.”) (citations omitted).
[24] Supra 1.
[25] “Trade Secret Intellectual Property Toolkit - USPTO.” U.S. Patent and Trademark Office, 17 Oct. 2023, www.uspto.gov/sites/default/files/documents/tradesecretsiptoolkit.pdf.
[26] Id.
[27] Id.
[28] Supra 1.
[29] “Trademark FAQs.” United States Patent and Trademark Office, 5 Aug. 2014, www.uspto.gov/learning-and-resources/trademark-faqs#type-browse-faqs_1223.
[30] Id.

Other IP-Like Rights

Beyond the four traditional categories of IP rights, certain "personhood" rights are recognized within the United States and are sometimes treated as IP rights. First, under federal law, Section 43(a) of the Trademark Act, also known as the Lanham Act, provides a limited cause of action for commercial use of an individual's likeness in a manner that misleads consumers as to the affiliation, connection, or endorsement of a product, service, or a business/organization, by that individual.[31] However, such a claim requires a minimum level of fame or recognizability, without which a claim cannot be advanced.[32]

Additionally, under various state laws, an individual's name, image, likeness, and voice can be protected under right of privacy laws; right of publicity laws, including variations on name, image, and likeness rights; and anti-revenge pornography laws, among others. Most states have enacted some form of legislation addressing these areas.[33] States, such as Tennessee and Texas, have developed statutory and common law protections for name, image, and likeness, while others, such as New York, rely solely on statutory provisions. Each state's laws are different and have their own nuances and limitations.[34]

IP Issues Raised by AI

The rapid development of generative AI, in particular, has raised a number of IP-related issues, primarily in four fundamental areas: (1) the ingestion of IP-protected works for training AI systems; (2) the implication of IP rights and the availability of IP protection with respect to the output of AI systems; (3) transparency in the training, functionality, and outputs of AI systems; and (4) the protection of individuals "personhood" rights from abuses of AI systems.

Training of AI Models and Ingestion of IP-protected Works

Generally, generative AI models are trained through machine learning using an extremely large amount of data.[35] The large body of this training data, also known as a corpus, will be tailored to the purpose of the model. For example, a large language model (LLM) is trained on a massive amount of different forms of language (e.g., articles, webpages, and all text documents found on a portion of the internet).

[31] House of Representatives, Congress. 15 U.S.C. 1125 - False designations of origin, false descriptions, and dilution forbidden. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title15/USCODE-2023-title15-chap22-subchapIII-sec1125.

[32] Electra v. 59 Murray Enters., 987 F.3d 233, 258 (2nd Cir. 2021).

[33] "NIL Legislation Tracker." Saul Ewing LLP, 1 Jan. 2023, www.saul.com/nil-legislation-tracker; see also: Rothman, Jennifer E. "Right of Publicity State-by-State." Rothman's Roadmap to the Right of Publicity, 5 Aug. 2015, https://rightofpublicityroadmap.com/.

[34] Id.

[35] Martineau, Kim. "What Is Generative Ai?" IBM Research, IBM, 20 April 2023, https://research.ibm.com/blog/what-is-generative-AI.

Generative AI models are typically trained mostly or entirely on real-world data. The models translate that data into an overall mathematical representation of the training corpus that can be used to produce new content based on patterns and probabilities derived from the training corpus.[36] For many generative AI models, their training data included numerous copyright-protected works scraped from the internet.[37] Currently, this is often done without seeking consent or a license or providing any credit or compensation to the copyright owners.[38]

Some commentators, including many content creators and other copyright owners, assert that using copyrighted works for AI training purposes constitutes infringement and is not fair use.[39] They argue that training uses entire works, the works are reproduced (e.g., downloaded) even if only temporarily, the model's mathematical representations of training data should be considered reproduction or a derivative work under copyright law, and that the relevant purpose is the overall purpose of the AI model, such as when the model is itself commercialized or used in commercial products or services.[40] Thus, they contend that developers of AI systems should obtain consent, give credit, and/or pay compensation for using copyrighted works in training most models.[41]

In contrast, other commentators, including many AI developers, assert that using a copyrighted work to train an AI model is fair use and not a reproduction. Thus, they argue that training is not copyright infringement.[42] They point to the transitory nature of any copied data, noting that the data is fed into an artificial neural network (ANN) but typically not retained.[43] Further, they argue that using a copyright-protected work during the training of an AI model has no effect on the value of that work.[44] They also argue that any copies used in training are not distributed or made available to the public but are used only for training.[45]

[36] Id.
[37] Appel, Gil, et al. "Generative AI Has an Intellectual Property Problem." Harvard Business Review, 7 April 2023, https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem.
[38] Id.
[39] Id.
[40] Id.; see also Uriostegui, Hassan. "AI-Copyright Weights: A New Frontier in Intellectual Property Law." Medium, Waken AI, 9 June 2023, medium.com/twinchat/ai-copyright-weights-a-new-frontier-in-intellectual-property-law-d8ee1b6c55ee; see also: Brittain, Blake. "Getty Images Lawsuit Says Stability AI Misused Photos to Train AI." Reuters, 6 Feb. 2023, www.reuters.com/legal/getty-images-lawsuit-says-stability-ai-misused-photos-train-ai-2023-02-06/; see also: Grynbaum, Michael M., and Ryan Mac. "The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work." The New York Times, 27 Dec. 2023, www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html; see also: Brittain, Blake. "Music Publishers Fire Back at Anthropic in AI Copyright Lawsuit." Reuters, 15 Feb. 2024, www.reuters.com/legal/litigation/music-publishers-fire-back-anthropic-ai-copyright-lawsuit-2024-02-15/.
[41] Id.
[42] Wolfson, Stephen. "Fair Use: Training Generative AI." Creative Commons, 17 Feb. 2023, https://creativecommons.org/2023/02/17/fair-use-training-generative-ai/.
[43] Id.
[44] Id.
[45] Zirpoli, Christopher T. "Generative Artificial Intelligence and Copyright Law." Congressional Research Services, 29 Sept. 2023, https://crsreports.congress.gov/product/pdf/LSB/LSB10922.

Courts have not yet settled whether using copyrighted works for AI training infringes the authors' copyright. A legal precedent at the center of the debate regarding AI and fair use is what is known as the *Google Books* case.[46] In that case, the court held that Google had not infringed the copyrights of authors whose books it had digitized, made searchable, and displayed limited portions of online because Google's use was a "transformative" fair use, despite that the displayed text was identical to text from the copyrighted works.[47]

AI advocates argue that the issues presented by AI training are similar to *Google Books* because copyrighted works in AI training data are transformed and incorporated into a mathematical model.[48] However, some copyright holders distinguish *Google Books* from AI-related cases, arguing that many AI models can be induced to produce virtually identical copies of copyrighted works in their training data.[49]

Whether the training of an AI model using copyrighted works constitutes copyright infringement or fair use is currently the subject of ongoing litigation.[50] While these cases work their way through the courts, a number of AI developers have moved forward and penned licensing agreements with some rights holders. For example, some media companies have already signed individual licensing agreements with AI companies, including *News Corp*, *Time*, the *Financial Times*, and *Le Monde* of France.[51]

These licenses have a broad range of terms, including joint product development agreements and traffic referrals.[52] Others are using existing corporate copyright clearinghouses, with some copyright licensing organizations including AI rights in their broad corporate licenses.[53] In addition, a number of startups have been formed to aggregate content (and the associated IP rights) into data collections that will be offered under blanket licenses to AI companies.[54] It is unclear what effect court decisions on fair use will have on these licensing mechanisms.

[46] Authors Guild v. Google, Inc., 804 F.3d 202 (2nd Cir. 2015).
[47] Id.
[48] Supra 42.
[49] Supra 45 ("the Getty Images lawsuit alleges that 'Stable Diffusion at times produces images that are highly similar to and derivative of the Getty Images.' One study has found 'a significant amount of copying' in less than 2% of the images created by Stable Diffusion, but the authors claimed that their methodology 'likely underestimates the true rate' of copying."); see also: Supra 40.
[50] Supra 40. Copyright owners have filed several lawsuits against AI companies, including Getty Images against Stability AI (photographs), the New York Times against OpenAI (news articles), and music publishers against Anthropic (song lyrics).
[51] Rosenblatt, Bill. "The Media Industry's Race To License Content For AI." Forbes, 18 July 2024, www.forbes.com/sites/billrosenblatt/2024/07/18/the-media-industrys-race-to-license-content-for-ai/.
[52] Id.
[53] Id.
[54] Id.

> **“Encouraging innovation and creativity by providing adequate IP protection for AI-assisted innovations and creative works is a key developing issue.**

The United States is not alone in attempting to work through issues of fair use connected to the training of AI. Earlier this year, preliminary draft legislation addressing the training of AI systems circulated in the PRC.[55] It provides, in pertinent part, that subject to certain limitations, an AI developer's use of copyrighted data of others is “a reasonable use of data” that does not require payment to the rights holder.[56]

At the same time, the European Union passed a landmark law, the EU AI Act, earlier this year, which requires general-purpose AI developers to establish a policy to respect the Copyright Directive and publish a sufficiently detailed summary of the content used for training their models. This transparency requirement, which goes into effect in August 2025, may allow creators to better understand if their copyrighted works were used in the training of AI models.

IP Rights, AI Outputs, and AI-Assisted Innovations and Creative Works

The ability to pursue claims for the infringement of IP rights against the outputs of AI systems is generally undisputed and requires little more than the application of existing law. However, encouraging innovation and creativity by providing adequate IP protection for AI-assisted innovations and creative works is a key developing issue.

The USPTO and the USCO have been proactive in issuing guidance on the implications of AI on the availability of patents and copyright registrations and how those agencies handle questions of inventorship and authorship.

Infringement of IP Rights by Outputs of Generative AI

Considering that generative AI models are often trained on copyrighted materials, concerns have been raised about using such models to generate infringing works.[57] One issue in deciding whether infringement occurred is whether these outputs are new, original, or derivative (i.e., infringing) works.[58]

Generative AI has been known to provide outputs that are near exact replicas of existing works under certain specific circumstances.[59] Users may also use a generative AI model to create content that is different from existing work but may closely mirror the style or characteristics of the works of a human author or artist.[60]

[55] Linghan, Zhang, et al. “Artificial Intelligence Law of the People’s Republic of China (Draft for Suggestions from Scholars).” Translated by Ben Murphy, Center for Security and Emerging Technology, 2 May 2024, https://cset.georgetown.edu/publication/china-ai-law-draft/.
[56] Id.
[57] Supra 37.; Supra 45.
[58] Id.
[59] Id.
[60] Id.

The federal courts are currently considering cases that address claims for the infringement of IP rights by outputs of AI systems. As in *Google Books*, where the court applied existing law to determine whether a new technological use case constituted IP infringement, courts are equipped to apply existing IP law that addresses infringement by non-AI-generated works to determine whether AI-generated works are infringing.[61] Current law already provides a framework to determine, for example, whether an allegedly infringing work (regardless of its source) is sufficiently similar to an existing work to constitute a derivative work.[62]

The fair use doctrine is well-developed and can be readily applied by the courts to cases involving generative AI outputs. This is because the fair use factors focus mainly on the use of the work, not its source.[63] For instance, using an AI-generated song in a social media post that closely resembles an existing copyright-protected song presents substantially the same fair use issues as using a non-AI-generated song in the same manner, and unlike AI training, such uses are well-understood under current law.

AI, Authorship, and Copyright Protection

In response to recent developments in generative AI technology and its use by individuals and businesses, the USCO issued new guidance on March 16, 2023, entitled “Registration Guidance: Works Containing Material Generated by Artificial Intelligence.”[64] It also launched an “AI Initiative” to examine copyright law and policy issues raised by AI,[65] including efforts to gather feedback from a broad array of stakeholders[66] and provide information to the public.[67]

[61] Authors Guild v. Google, Inc., 804 F.3d 202 (2nd Cir. 2015); see also: Gil Appel et al., Generative AI Has an Intellectual Property Problem, Harvard Business Review, 7 April 2023, https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem.

[62] Id.; see also: Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith, 598 U.S. 508 (2023).

[63] Id.; Supra 21.

[64] Office of the Federal Register, National Archives and Records Administration. 88 FR 16190 - Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence. Office of the Federal Register, National Archives and Records Administration, https://www.govinfo.gov/app/details/FR-2023-03-16/2023-05321.

[65] U.S. Copyright Office. “Copyright Office Launches New Artificial Intelligence Initiative.” NewsNet Issue 1004 | U.S. Copyright Office, 16 March 2023, www.copyright.gov/newsnet/2023/1004.html.

[66] U.S. Copyright Office. “Spring 2023 AI Listening Sessions.” Spring 2023 AI Listening Sessions | U.S. Copyright Office, 2023, www.copyright.gov/ai/listening-sessions.html?loclr=eanco. Attendees included participants from Microsoft (literary works), Jasper AI (visual arts), the Motion Picture Association (audiovisual works), Spotify (music), and other stakeholders.

[67] Scheland, Nora. “Our Summer of Artificial Intelligence: Copyright Office Hosts Two Webinars on Copyright and AI.” The Library of Congress, 23 Aug. 2023, https://blogs.loc.gov/copyright/2023/08/our-summer-of-artificial-intelligence-copyright-office-hosts-two-webinars-on-copyright-and-ai/.; see also: U.S. Copyright Off., Webinar: Registration Guidance for Works Containing AI-generated Content, https://copyright.gov/events/ai-application-process/?loclr=eanco.; see also: U.S. Copyright Off., International Copyright Issues and Artificial Intelligence, 26 Aug. 2024, https://www.copyright.gov/events/international-ai-copyright-webinar/?loclr=eanco.

The USCO's guidance indicates that copyright registration is only available for human-authored materials.[68] The term "Author," used in both the Constitution and the Copyright Act, has been found by the courts to exclude non-humans.[69] The guidance did not, however, seek to ban all uses of AI in registered works. Instead, it pointed to *Burrow-Giles Lithographic Co. v. Sarony,* a Supreme Court case about the use of camera technology in which the Court ruled that photographs can be copyrighted as long as they represent the "original intellectual conceptions of the [human] author."[70]

When seeking to register a work generated by or created with the assistance of AI, the USCO guidance explained that those using AI in their creations can claim copyright protection only for their human contributions.[71] For example, if "traditional elements of authorship" are determined and executed by an AI in response to solely a prompt—such as a simple prompt requesting a poem in the style of William Shakespeare—the generated material is not the product of human authorship according to the USCO.[72]

Conversely, "a human may select or arrange AI-generated material in a sufficiently creative way that 'the resulting work as a whole constitutes an original work of authorship.'"[73] Under the guidance, applicants must clearly identify human-authored portions and explicitly exclude AI-generated content that is more than *de minimis*.[74]

If applicants have previously submitted or have pending applications that inadequately disclosed AI involvement, they must correct those omissions to ensure the registration's validity.[75] Failure to accurately represent AI contributions can lead to registration cancellation or legal challenges in infringement cases.[76]

It is not yet clear how the USCO will administer the rules and standards put forth in its guidance. What information, and how much of it, is required to establish human authorship of protectible portions of an AI-assisted work is yet to be determined.

The USCO generally does not investigate the truth of authorship claims in copyright registrations.[77] Questions also remain as to how to exclude AI-generated portions of a work when both human and AI-generated output are blended to create it.

[68] Supra 64.
[69] Id.
[70] Id. Further, the USCO noted additional court cases holding that works "authored by non-human spiritual beings" or animals, like a monkey taking a photograph, are not eligible for copyright protection.
[71] Id.
[72] Id.
[73] Id.
[74] Id.
[75] Id.
[76] Id.
[77] Compendium of U.S. Copyright Office Practices, § 503.2 (3d ed. 2021). https://www.copyright.gov/comp3/chap500/ch500-identifying-works.pdf.

In addition to its guidance on registering AI-assisted works, the USCO also issued a notice of inquiry (NOI) on August 30, 2023, as part of its AI Initiative.[78] The NOI requested comments from stakeholders about a range of copyright issues implicated by AI, including the legal status of AI-generated outputs.[79] As of December 2023, the USCO had received over 10,000 comments.[80] Based in part on the feedback it received, the USCO is issuing a series of reports on these issues and is expected to issue a report addressing authorship and copyright registration of AI-assisted works.[81]

In the meantime, the USCO has already issued a handful of decisions rejecting registration of works created or modified with AI tools. On September 5, 2023, the USCO Copyright Review Board (CRB) issued a decision denying registration to a work that was created with AI tools on the basis that the creator refused to disclaim the more than "*de minimis*" portions of the work that were "generated by artificial intelligence."[82]

On December 11, 2023, the CRB affirmed the USCO's decision to refuse to register an artistic work featuring a photograph that had been processed using AI to appear in the style of Vincent Van Gogh's *The Starry Night*.[83] The applicant utilized RAGHAV—an AI tool—that enables users to start by selecting a base image and a style image, identify the level to which the style image should be applied to the base image, and then have the AI generate a final output image.[84] The CRB held that the applicant could not register the output image.[85] It explained that "selecting a single number for a style filter is the kind of *de minimis* authorship not protected by copyright."[86]

Authorship issues are also being raised in the courts. On August 18, 2023, the U.S. District Court for the District of Columbia affirmed a decision by the USCO to reject the registration of art created by a computer system owned by Stephen Thaler.[87]

[78] U.S. Copyright Off., Copyright Office Issues Notice of Inquiry on Copyright and Artificial Intelligence, 30 Aug. 2023, https://www.copyright.gov/newsnet/2023/1017.html?utm_campaign=subscriptioncenter&utm_%20content=&utm_medium=email&utm_name=&utm_source=govdelivery&utm_term=. "The Office is undertaking a study of the copyright law and policy issues raised by generative AI and is assessing whether legislative or regulatory steps are warranted."
[79] Id.
[80] U.S. Copyright Off., Artificial Intelligence Study, https://www.copyright.gov/policy/artificial-intelligence/.
[81] Id. The first report was recently issued, which addresses digital replicas.; see also: U.S. Copyright Off., Copyright and Artificial Intelligence: Part 1: Digital Replicas, July 2024, https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-1-Digital-Replicas-Report.pdf.
[82] In re Theatre D'opera Spatial (U.S. Copyright Off. Bd. of Appeals Sep. 5, 2023), https://www.copyright.gov/rulings-filings/review-board/docs/Theatre-Dopera-Spatial.pdf.
[83] In re Suryast (U.S. Copyright Off. Bd. of Appeals Dec. 11, 2023), https://www.copyright.gov/rulings-filings/review-board/docs/SURYAST.pdf.
[84] Id.
[85] Id.
[86] Id.
[87] Thaler v. Perlmutter, Case No. 1:22-cv-01564, slip op. (D.D.C. 18 Aug. 2023), https://www.copyright.gov/ai/docs/district-court-decision-affirming-refusal-of-registration.pdf. (Mr. Thaler sought to register the work, listing the computer system as the author and himself as merely the owner. The parties moved for summary judgment on the question of "whether a work generated entirely by an artificial system absent human involvement should be eligible for copyright." The court found that "human authorship is an essential part of a valid copyright claim," and affirmed the decision that the work could not be registered, explaining that "[c]opyright has never stretched so far . . . as to protect works generated by new forms of technology operating absent any guiding human hand . . . . Human authorship is a bedrock requirement of copyright.")

The court found that "human authorship is an essential part of a valid copyright claim" and affirmed the decision that the work could not be registered, explaining that "[c]opyright has never stretched so far . . . as to protect works generated by new forms of technology operating absent any guiding human hand . . . human authorship is a bedrock requirement of copyright."[88] The court reasoned that "[n]on-human actors need no incentivization with the promise of exclusive rights under U.S. law, and copyright was therefore not designed to reach them."[89] The decision is currently pending appeal.

These issues raised by the use of AI in producing creative content are already multiplying and can be expected to become increasingly difficult to resolve. For example, country music artist Randy Travis recently released a new song, "Where That Came From," his first release in over a decade.[90]

In 2013, Mr. Travis was hospitalized with a viral infection that eventually led to a stroke and aphasia, rendering him incapable of singing.[91] A specialized AI model was trained on samples of his voice and used to generate a vocal performance of the new song that was tweaked and edited to sound like Mr. Travis, all with his permission and involvement.[92] It is unclear, however, whether that recording can obtain a copyright registration at the USCO under its new policies, given that the vocal performance was entirely generated by an AI model, albeit with human guidance.

AI, Inventorship, and Patentability

Biomedical, pharmaceutical, and software companies are investing heavily in developing AI tools and using them to drive innovation in their respective fields.[93] As noted above, pursuant to the Patent Act, an invention must be useful, novel, non-obvious, and directed to statutory subject matter to be patentable.[94] AI-assisted inventions raise questions related to inventorship, patent subject matter eligibility, as well as novelty and non-obviousness.

[88] Id. at 1-2,8.

[89] Id. at 10.

[90] Maria Sherman, With help from AI, Randy Travis got his voice back. Here's how his first song post-stroke came to be, AP News, 6 May 2024, https://apnews.com/article/randy-travis-artificial-intelligence-song-voice-589a8c142f70ed8ccf53af6d32c662dc.

[91] Id.

[92] Id.

[93] Artificial Intelligence and Intellectual Property: Part III – IP Protection for AI-Assisted Inventions and Creative Works: Hearing before the Subcomm. On Courts, Intellectual, and the Internet of the H. Comm. on the Judiciary, 118th Cong. (2024) (statement of Clair Laporte, Fellow, Ginkgo Bioworks), https://judiciary.house.gov/sites/evo-subsites/republicans-judiciary.house.gov/files/evo-media-document/Laporte%20Testimony.pdf; see also: Intellectual Property and Strategic Competition with China: Part 3 - IP Theft, Cybersecurity, and AI, 118th Cong. (2023) (statement of Robert Sheldon, Sr. Director, Public Policy & Strategy, CrowdStrike), https://judiciary.house.gov/sites/evo-subsites/republicans-judiciary.house.gov/files/evo-media-document/sheldon-testimony_0.pdf.

[94] Supra 5; see also: House of Representatives, Congress. 17 U.S.C. 103 - Subject matter of copyright: Compilations and derivative works. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title17/USCODE-2023-title17-chap1-sec103; see also: House of Representatives, Congress. 17 U.S.C. 112 - Limitations on exclusive rights: Ephemeral recordings. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title17/USCODE-2023-title17-chap1-sec112.

The courts have spoken on a number of issues, and the USPTO has issued some guidance on AI-assisted inventions, but questions remain that create a great deal of uncertainty for investors in fields where AI is being utilized to assist in the innovation process.

Inventorship

Under current U.S. patent law, a patent may only be issued if all inventors are properly disclosed, and only humans may be considered inventors.[95] This means that AI systems also cannot be listed as inventors on a U.S. patent application.[96] The USPTO recently issued guidance on legal and examination-related issues surrounding AI-assisted inventions. The guidance provides that while AI cannot be named as an inventor, the USPTO will consider a natural person who used AI as an inventor if their contribution to the claimed invention is "significant."[97] The guidance provides a "non-exhaustive" list of five "guiding principles" to assist in determining what constitutes a significant contribution.[98]

Alongside the guidance, the USPTO provided examples of how examiners should apply the guiding principles to AI-assisted inventions.[99] Among these include a human prompting an AI system to generate a transaxle design for a toy car.[100] The human could be named as the inventor if the human's contribution to design and testing is deemed significant, such as through substantial experiment-driven modification of the AI-generated design.[101] Another example involves an AI system generating a vast number of potential molecular structures for a new drug and a human chemist synthesizing, analyzing, and selecting promising candidates among them for further testing.[102] If the human plays a significant role in selecting, evaluating, and modifying the candidate compounds, the human could be named the inventor.[103]

[95] Thaler v. Vidal, 43 F.4th 1207 (Fed. Cir. 2022); Mark Masutani & Jacob W. S. Schneider, Making the Case for AI Inventorship; Thaler v. Vidal, Case No. 21-2347 (Fed. Cir.), 7 June 2022, https://www.hklaw.com/en/insights/publications/2022/06/making-the-case-for-ai-inventorship.; see also: Ji Mao, Revisiting AI Inventorship in Thaler v. Vidal, 4 Oct. 2022, https://www.hklaw.com/en/insights/publications/2022/10/revisiting-ai-inteventorship-in-thaler-v-vidal. In Thaler v. Vidal, an applicant (the same Thaler discussed above who sought a copyright registration for an AI-generated work) filed two patent applications for two inventions attributed to his AI system, "Device for the Autonomous Bootstrapping of Unified Science" (DABUS). The applicant argued that DABUS, not a human being, was the true inventor of both inventions. He contended that the Patent Act's inventor requirement did not exclude AI. The USPTO disagreed, denying the applications because the Patent Act, in its view, only recognizes natural persons as inventors. The Federal Circuit focused on the requirement in the Patent Act that an inventor be an "individual" or "individuals," concluding that "individual" refers only to human beings. The court referenced earlier cases that held that corporations could own patents but are not considered inventors themselves. The applicant then petitioned the Supreme Court to hear the case, but certiorari was denied.

[96] Id.

[97] Inventorship Guidance for AI-Assisted Inventions, 89 Fed. Reg. 10043, 13 Feb. 2024, https://www.federalregister.gov/documents/2024/02/13/2024-02623/inventorship-guidance-for-ai-assisted-inventions.

[98] Id.

[99] Id.

[100] Id.

[101] Id.

[102] Id.

[103] Id.

In general, the legal principles underlying the USPTO guidance are rooted in long-standing rules governing joint inventorship, i.e., when a particular person must be added as an inventor on a patent. All (human) inventors must be included and failure to do so may result in the patent being invalid.[104]

The USPTO guidance makes clear that any invention to which no humans contributed significantly (because it was entirely generated by an AI system) is unable to be patented at all due to the lack of a recognized inventor.[105] But in the case of human inventors assisted by an AI system, the guidance applies joint inventorship law by requiring all human inventors who contributed significantly to be listed but allowing the omission of any AI contribution because an AI cannot be an "inventor" under current law.[106]

The guidance contemplates a situation where a patent may be issued to one or more (human) inventors where none of them, individually *or even collectively*, contributed to *all* aspects of the invention because an AI "contributed" to some aspects but does not legally qualify as an "inventor" and, thus, need not be listed. The guidance notes the possibility that humans who created or operated the AI system may qualify as inventors in certain circumstances,[107] but it does not discuss in depth the relevant factors for determining when that would be the case.

The USPTO has yet to issue further guidance regarding what process it will use to determine about human and AI inventorship. For example, prior to the guidance, the USPTO did not typically question or examine whether a patent application correctly listed the inventors, which meant that "applicants rarely need to submit information regarding inventorship."[108]

However, the guidance discusses evaluating facts and evidence to determine whether the listed human inventors made a significant contribution to the invention and emphasizes that examiners "have the ability to require the submission of information that may be reasonably necessary to properly examine" a patent application.[109]

It is unclear what evidence an applicant must now provide about inventorship, what evidence examiners will consider, what standards or criteria examiners will apply, and what the ramifications will be for the examination process and the resulting patent if issued.

[104] Pannu v. Iolab Corp., 155 F.3d 1344, 1349-50 (Fed. Cir. 1998); In re VerHoef, 888 F.3d 1362, 1366-67 (Fed. Cir. 2018); Pannu, 155 F.3d at 1351. (Explaining that any person who contributed to the conception of the claimed invention is a joint inventor and must be listed. To be a joint inventor, the person must have contributed significantly to the invention, and contributions that are insignificant in quality within the context of the full invention do not confer inventorship. But a significant contribution even to only one element of the invention may be sufficient. In the joint inventorship context, a patent may be granted to the joint inventors collectively for their collective conception, even if none of the individual inventors conceived of the entire invention.)
[105] USPTO Guidance, Supra 97.
[106] Id.
[107] Id.
[108] Id.
[109] Id.

The guidance also leaves unclear what sorts of AI use and which AI systems trigger these requirements. It does not address how the agency will adapt to the additional resources needed and burdens placed on its examiners.

These open questions leave unclear how the guidelines for patenting AI-assisted inventions will influence the effectiveness and predictability of patent protection for those inventions, how investors will evaluate potential investments in startups and R&D in both the AI industry and other industries in which AI is used to assist innovation, and the overall impact on innovation generally. Finally, how the USPTO adjusts to processing applications for AI-assisted inventions remains to be seen, particularly if the availability of AI technology creates a flood of new patent applications (which are potentially more challenging to evaluate under the new guidelines).

It is also unclear whether courts will interpret inventorship law in the context of AI in the same way as the USPTO guidance and how courts will address these questions in general. As the guidance itself indicates, it does not have the force of law and merely constitutes a statement of USPTO policy that will only directly affect the USPTO's actions.[110] Therefore, courts may apply a different rule altogether absent legislation.

Subject Matter Eligibility, Novelty, and Non-Obviousness

Section 101 of Title 35 provides that "[w]hoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefore, subject to the conditions and requirements of this title."[111] This is known as the subject matter eligibility requirement.

The issue of what qualifies as statutory subject matter has received significant attention in recent years due to shifts in section 101 jurisprudence, including with respect to AI. Section 101 has been interpreted as limiting the ability of an inventor to receive a patent directed to certain categories of inventions, including abstract ideas, laws of nature, and natural phenomena.[112]

As the Supreme Court explained in *Alice Corp. v. CLS Bank International*, "in applying the §101 exception, we must distinguish between patents that claim the building blocks of human ingenuity and those that integrate the building blocks into something more."[113]

[110] Id.
[111] Supra 4.
[112] See Bilski v. Kappos, 561 U.S. 593 (2010); Mayo Collaborative Servs. v. Prometheus Labs., 566 U.S. 66 (2012); Association for Molecular Pathology v. Myriad Genetics, 569 U.S. 576 (2013); Alice Corp. Pty. v. CLS Bank Int'l, 134 S.Ct 2347 (2014); see also: Emily Blevins & Kevin Hickey, Congressional Research Service IF12563, Patent-Eligible Subject Matter Reform: An Overview, 3 Jan. 2024, https://crsreports.congress.gov/product/pdf/IF/IF12563/1. (Cases involving subject matter that was found to be ineligible in recent years have included business methods using computers or communications networks for hedging price risk or mitigating settlement risk in financial transactions; methods for calibrating the dosage of a particular drug; and isolated human DNA segments).
[113] Alice Corp. v. CLS Bank International,134 S.Ct. 2347, 2354 (2014) (internal quotation marks omitted).

Some stakeholders have expressed concerns that subject matter ineligibility resulting from this decision may be an issue for patenting AI-related inventions because “they may be characterized as methods of organizing human activities, mental processes, or mathematical concepts.”[114]

AI also raises issues relating to the novelty[115] and non-obviousness[116] requirements for patentability. Both require that the claimed invention be an advance over the “prior art,” i.e., preexisting technology and knowledge. For example, a claimed invention cannot have patent protection if it was previously “patented, described in a printed publication, or in public use, on sale, or otherwise available to the public”—e.g., it cannot be patented if someone else had already patented it, or it was already available to the public—because then it is not actually a new invention.[117] Similarly, it cannot represent merely an obvious step over the prior art because it is not much of an invention if most people in that field (“persons having ordinary skill in the art”) would have found it obvious.[118]

Concerns have been raised that artificial intelligence may raise the bar of non-obviousness to unachievable levels for most human inventors and inventions. Because AI systems can ingest and combine nearly endless numbers of prior art references, it has been posited that, theoretically, only a rare invention could exceed the knowledge and performance of AI.[119] For example, one commentator has noted that “[e]xpanding the scope of prior art to almost an infinite collection greatly raises the non-obviousness and, therefore, the patentability bar.”[120] If one of ordinary skill in the art would ordinarily use AI, then the question of obviousness becomes a question of whether an AI would reach the same result.[121] As AI systems grow more powerful, it will be increasingly difficult to find inventions that would be beyond even an ordinary AI’s capabilities, i.e., “non-obvious” and, thus, patentable.

Transparency of AI Inputs and Outputs

IP intersects with the transparency of AI systems and their use in three main ways. The first is transparency regarding what copyrighted works are ingested and used to train a particular AI model. The second is transparency regarding whether a particular output can be connected to existing IP rights, such as IP-protected works used to train the model. The third is transparency regarding the involvement and role of AI in producing outputs, particularly outputs that may themselves be protectible by IP.

[114] Supra 112.
[115] Supra 12.
[116] Supra 6.
[117] Supra 12.
[118] Supra 6.
[119] Riddhi Setty, AI Use Risks Drop in New Patents as Ideas Are Rendered ‘Obvious’, Bloomberg Law, 12 July 2023, https://news.bloomberglaw.com/ip-law/ai-use-risks-drop-in-new-patents-as-ideas-are-rendered-obvious.
[120] Lexi Heon, Comment, Artificially Obvious But Genuinely New: How Artificial Intelligence Alters the Patent Obviousness Analysis, 53 Seton Hall L. Rev. 359, 378, 2022, https://scholarship.shu.edu/cgi/viewcontent.cgi?article=2912&context=shlr.
[121] Id.

Some commentators have called for increased transparency of training data to train AI models and generate specific outputs. Some have proposed, for example, requiring developers to provide all ingested content used to train their AI models in a public repository, such as with the USCO, or at least a summary or other report of their training data.[122]

For example, this would potentially allow creators to see whether their copyrighted content was included in the training corpus of an AI model to identify infringement. Others have advocated for disclosures of the source of data and input that led to a particular output from an AI system.[123] In addition to potentially allowing creators to see whether their copyrighted works were involved in generating a particular output, it could also allow users of the AI system to assess and understand outputs.[124]

Others, including AI technology companies, have conversely argued that they should not be subject to regulation for the use of copyright materials. These include assertions that such transparency proposals related to the source of inputs and training data could be an expensive and significant technological or logistical challenge, especially for AI developers who may need to document millions (if not billions) of potentially IP-protected works. Additionally, they argue that it could be difficult to comply even with disclosure requirements limited to works with IP protection, as determining whether a particular work may be protected is often not straightforward and could be extremely challenging to do at scale.

Some of these commentators are also concerned about whether mandatory disclosure would forcibly reveal the trade secrets and proprietary information of AI developers. In contrast with large language models (LLMs) that absorb massive amounts of online data by scraping data from the internet, niche or specialized AI models are developed using curated data sets. Special-purpose AI models can be built for biomedical or pharmaceutical research, medical diagnosis, and other highly specialized tasks with little room for error.[125]

[122] Core Principles for Artificial Intelligence Applications, Human Artistry Campaign, https://www.humanartistrycampaign.com/. ("Complete recordkeeping of copyrighted works, performances, and likenesses, including the way in which they were used to develop and train any AI system, is essential. Algorithmic transparency and clear identification of a work's provenance are foundational to AI trustworthiness. Stakeholders should work collaboratively to develop standards for technologies that identify the input used to create AI-generated output."); see also: "H.R.7913 - 118th Congress (2023-2024): Generative AI Copyright Disclosure Act of 2024." Congress.gov, Library of Congress, 9 April 2024, https://www.congress.gov/bill/118th-congress/house-bill/7913.

[123] Recommendation of the Council on Artificial Intelligence, Organization for Economic Co-operation and Development § 1.3 May 2, 2024, https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449.

[124] Id.

[125] See, e.g., Neil Savage, Tapping into the drug discovery potential of AI, Nature, 27 May 2021, https://www.nature.com/articles/d43747-021-00045-7; see also: Pranav Kumar, The Rise of Bespoke AI Models: Tailoring AI to Meet Specific Enterprise Needs, Medium, 16 July 16, 2024, https://medium.com/@k.pranav_22/the-rise-of-bespoke-ai-models-tailoring-ai-to-meet-specific-enterprise-needs-e5202da2a535; see also: Martin Willemink et al, Preparing Medical Imaging Data for Machine Learning, Radiological Society of North America, 18 Feb. 2020, https://pubs.rsna.org/doi/full/10.1148/radiol.2020192224.

Niche or specialized AI models can be trained, for example, on internal company data, such as proprietary chemical and biological information collected through years of research or authoritative sources of information carefully selected to reduce errors.[126] Such specialized AI models attach a great deal of value and expense to the data sets used for training, and significant investments are made in data curation, cleaning, and preparation.[127] Some commentators in the industry argue that requiring the disclosure of training data may discourage investment and effort spent to identify, assemble, clean, curate, and otherwise develop specialized training data if disclosure requirements mean it would potentially be available to competitors without the need for that investment or effort.

Industry groups, the White House, and lawmakers have also advocated for increasing transparency when a particular visual or audio output has been generated or modified by AI,[128] which could have implications for IP rights. A number of technical options have been discussed, including appending provenance information in metadata and mandating visible or invisible watermarking on such outputs.[129]

AI developers disagree on the technical feasibility of applying such watermarking in a way that resists removal[130] and have expressed concerns that mandatory watermarking would run afoul of the First Amendment.[131] Many of these concerns are not specific to IP and are discussed in more detail elsewhere in this report.

Abuse of Identity-Based Rights by AI-Generated Digital Replicas

Perhaps the most high-profile IP issue raised by AI technology concerns the protection of individuals who have found their image, likeness, or voice usurped or abused by others through AI-generated digital replicas, such as “deepfakes.”

[126] Supra 125., (“Others such as Recursion Pharmaceuticals, which recently raised $436 million in its initial public offering, are generating vast amounts of bespoke data on cellular behavior in the hope that these can be mined using AI to reveal biological insights that could inform the discovery of innovative drugs.”).

[127] Id.; see also: Intellectual Property and Strategic Competition with China: Part 3 - IP Theft, Cybersecurity, and AI: Hearing before the Subcomm. On Courts, Intellectual, and the Internet of the H. Comm. on the Judiciary, 118th Cong. 2023, https://judiciary.house.gov/sites/evo-subsites/republicans-judiciary.house.gov/files/evo-media-document/brennan-testimony_0.pdf (“a model’s performance is only as good as the data it is trained on. Scale has worked on nearly every generative AI advancement and LLM released. We have also pioneered many of industry’s best practices today around data fine-tuning, red teaming, and test and evaluation.”).

[128] AI Output Disclosures: Use, Provenance, Adverse Incidents, National Telecommunications and Information Administration 27 March 2024, https://www.ntia.gov/issues/artificial-intelligence/ai-accountability-policy-report/developing-accountability-inputs-a-deeper-dive/information-flow/ai-output-disclosures; see also: "S.2765 - 118th Congress (2023-2024): Advisory for AI-Generated Content Act." Congress.gov, Library of Congress, 12 Sept. 2023, https://www.congress.gov/bill/118th-congress/senate-bill/2765.; see also: Chad Heitzenrater, The Case For and Against AI Watermarking, RAND, 17 Jan. 2024, https://www.rand.org/pubs/commentary/2024/01/the-case-for-and-against-ai-watermarking.html.

[129] Siddarth Srinivasan, Detecting AI fingerprints: A guide to watermarking and beyond, Brookings, 4 Jan. 2024, https://www.brookings.edu/articles/detecting-ai-fingerprints-a-guide-to-watermarking-and-beyond/.

[130] Id.

[131] The Fire, Artificial intelligence, free speech, and the First Amendment, Foundation for Individual Rights and Expression, https://www.thefire.org/research-learn/artificial-intelligence-free-speech-and-first-amendment.

Generative AI has been used to facilitate generating deepfakes, which began as "synthetic media where a person in an image or video is swapped with another person's likeness" but has now expanded to images, videos, and audio of people generated entirely by AI systems.[132] As discussed in the chapter on **Content Authenticity**, alarms have been sounded about the serious risks deepfake videos pose to the ability to trust information included in videos and audio recordings and to the broader dissemination of accurate information.[133]

In the creative space, deepfake audio tracks by well-known artists have been released, garnering millions of streams before they were successfully removed. In early 2023, an audio track for a song entitled "Heart on My Sleeve" was uploaded to a number of streaming services and appeared to be a new collaboration between recording artists Drake and The Weeknd (although on some platforms, it did not include Drake or The Weeknd in the title or track description, or was explicitly labeled as an "AI song").[134] Although many saw the track as a legitimate release by the artists, it was actually an AI-generated deepfake that drove millions of streams before it was ultimately removed.[135]

Given that many recording artists' income derives from their voice, the harm to an artist whose voice is replicated in this way is evident, particularly if the replicated voice is used for content that would, for example, cause the artist reputational harm, offend fans, or cause confusion with respect to contractual obligations. Further, considering that a recording artist's main "products" are the artist's music, the production of AI-generated deepfake songs essentially forces the artist to compete against his or her own digital replica, potentially funneling money away from the artist and confusing consumers.

Deepfakes have also been used to create fake product endorsements. AI-generated appearances of Elon Musk, Tom Cruise, and Leonardo DiCaprio have been used in marketing campaigns without their approval or endorsement.[136] In October 2023, actor Tom Hanks released a statement explaining that "[t]here's a video out there promoting some dental plan with an AI version of me. I have nothing to do with it."[137]

[132] Meredith Somers, Deepfakes, explained, MIT, 21 July 2020, https://mitsloan.mit.edu/ideas-made-to-matter/deepfakes-explained.

[133] Don Fallis, The Epistemic Threat of Deepfakes, Philos Technol. 2021; 34(4): 623-643, 6 Aug. 2020, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7406872/#Sec1title.

[134] Joe Coscarelli, An A.I. Hit of Fake 'Drake' and 'The Weeknd' Rattles the Music World, New York Times, 24 April 2023, https://www.nytimes.com/2023/04/19/arts/music/ai-drake-the-weeknd-fake.html; see also: Chris Willman, AI-Generated Fake 'Drake'/'Weeknd' Collaboration, 'Heart on My Sleeve,' Delights Fans and Sets Off Industry Alarm Bells, Variety, 17 April 2023, https://variety.com/2023/music/news/fake-ai-generated-drake-weeknd-collaboration-heart-on-my-sleeve-1235585451/.

[135] Id.

[136] Patrick Coffee, 'Deepfakes' of Celebrities Have Begun Appearing in Ads, With or Without Their Permission, Wall Street Journal, 25 Oct. 2022, https://www.wsj.com/articles/deepfakes-of-celebrities-have-begun-appearing-in-ads-with-or-without-their-permission-11666692003.

[137] Ronald & Jack Guy, Tom Hanks says dental plan video uses 'AI version of me' without permission, CNN, 2 Oct. 2023, https://www.cnn.com/2023/10/02/entertainment/tom-hanks-ai-dental-plan-video-intl-scli/index.html.

As with recording artists, the harm from the use of replicated likenesses of actors, models, or other individuals whose livelihoods depend on their face or likeness is evident—in addition to potential reputational harm, an actor or model could, in theory, have to compete with his or her own digital replica to secure endorsement deals or even possibly acting work in films, TV, or commercials.

Beyond being used to engage in criminal fraud and create deepfake videos of public figures, music tracks, and product endorsements, generative AI has also been used to make deepfake pornography. In fact, the vast majority of all deepfake videos on the internet are pornographic in nature. Between 90 and 95% of online deepfake videos are non-consensually generated pornography, 90% of which target a female victim.[138]

Deepfake pornography began to appear in late 2017, with a Reddit user named "deepfakes"—a user name that ultimately came to describe the whole category of AI-generated videos more broadly—utilizing open-source AI to make and share videos that inserted the faces of female celebrities into existing pornography.[139] More recently, however, deepfake pornography has gone beyond celebrities.[140]

Deepfakes are now being used to target private individuals for inclusion in fake pornographic materials (often called "revenge porn"), with some studies identifying at least 100,000 victims, mostly underage girls.[141]

The legal framework surrounding the protection of individuals' identifying characteristics such as name, image, voice, and likeness presents a complex landscape characterized by federal and state-level laws, as well as notable gaps and ongoing legislative efforts to address emerging challenges.

At the federal level, existing IP rights, including trademarks and copyrights, lack comprehensive mechanisms to prevent various forms of misuse of personal attributes in the era of AI-generated content.

While federal trademark law, as codified in the Lanham Act, protects words, names, symbols, or devices distinguishing goods or services, it currently offers limited protection for individual likeness rights and is untested with respect to AI.[142]

[138] Karen Hao, Deepfake porn is ruining women's lives. Now the law may finally ban it, MIT, 12 Feb. 2021, https://www.technologyreview.com/2021/02/12/1018222/deepfake-revenge-porn-coming-ban/.
[139] Id.
[140] Id.
[141] Id..; see also: Natasha Singer, Spurred by Teen Girls, States Move to Ban Deepfake Nudes, New York Times, 22 April 2024, https://www.nytimes.com/2024/04/22/technology/deepfake-ai-nudes-high-school-laws.html; see also: Kerry Breen, New Jersey teen sues classmate for allegedly creating, sharing fake AI nudes, CBS News, 9 Feb. 2024, https://www.cbsnews.com/news/new-jersey-teen-sues-classmate-for-allegedly-creating-sharing-fake-ai-nudes/; see also: Kat Tenbarge and Liz Kreutz, A Beverly Hills middle school is investigating students sharing AI-made nude photos of classmates, NBC News, 27 Feb. 2024, https://www.nbcnews.com/tech/misinformation/beverly-vista-hills-middle-school-ai-images-deepfakes-rcna140775.
[142] Supra 31; Christopher T. Zirpoli, "An Introduction to Trademark Law in the United States," Congressional Research Service, IF12456, 24 July 2023, https://crsreports.congress.gov/product/pdf/IF/IF12456.

Federal copyright law protects expressive works but does not broadly cover personal attributes like voice or likeness.[143] In other words, a copyright would protect a particular recording of a person's voice, but no copyright would prevent someone from digitally replicating that person's voice to create new recordings that are not copies or derivatives of an existing recording.

State laws addressing the misuse of an individual's name, image, likeness, and voice fall into three main categories: rights of privacy; rights of publicity, including variations on name, image, and likeness rights; and anti-revenge pornography laws.

Most states have enacted some form of legislation addressing these issues.[144] However, the effectiveness of state laws is limited by jurisdictional constraints and inconsistencies among statutes, creating challenges for individuals seeking redress against misuse of their personal attributes. States like Tennessee and Texas have developed statutory and common law protections for name, image, and likeness, while others like New York rely solely on statutory provisions, and each state's laws are different and have their own nuances and limitations.[145]

These gaps and ambiguities in state and federal protection for identity-based rights, such as name, image, and likeness rights, have led to calls for federal legislation to address digital replicas. The USCO concluded in a recent report on digital replicas, that "new federal legislation is urgently needed" to address the speed and scale of production and dissemination of digital replicas enabled by generative AI technology.[146] In reaching this conclusion, the USCO found that state laws are inconsistent and insufficient to address the problems exposed by AI.[147]

The Office also found existing federal laws to be insufficient. According to the USCO, the Copyright Act protects original works of authorship but does not prevent unauthorized duplications of image or likeness.[148] With respect to the Lanham Act's false endorsement provisions, the USCO noted that many federal courts require general fame or celebrity to access its protections and that "[i]t may be difficult for many individuals, including less famous artists and performers, to prove that the challenged conduct is likely to confuse consumers regarding the plaintiff's association with, or approval of, the defendant's commercial activity."[149]

[143] Supra 12; Karl M. Zielaznicki, et al., The Intersection of Generative AI and Copyright Law, Troutman Pepper, 21 July 2023, https://www.troutman.com/insights/the-intersection-of-generative-ai-and-copyright-law.html.
[144] NIL Legislation Tracker, SAUL EWING, https://www.saul.com/nil-legislation-tracker; see also: Jennifer E. Rothman, Rothman's Roadmap to the Right of Publicity, https://rightofpublicityroadmap.com.
[145] Id.
[146] Supra 81.
[147] Id.
[148] Id.
[149] Id.

According to the USCO, federal legislation should address the definition of digital replica, the persons protected, the term of protection, prohibited acts, secondary liability, licenses and assignments, accommodations of First Amendment concerns, remedies, and interaction with state law.[150]

The USCO report also concluded that federal legislation should “not sweep too broadly” but target replicas that “convincingly appear to be the actual individual being replicated.”[151] The report further recommended that federal protection for identity-based rights should not extend significantly beyond the individual's lifetime, that such rights should be licensable but not permanently transferable, that First Amendment protections be robustly preserved, and that state laws should not be completely preempted.[152]

Legal Clarity

Applicable laws, regulations, and agency guidance on authorship/inventorship and the ability to obtain IP protection for AI-assisted creative works and inventions should be clarified to promote the development and application of powerful new AI tools for content creation and innovation while protecting human creators and innovators.

As discussed above, the USCO and USPTO have already issued new guidance on the authorship and inventorship requirements for obtaining copyright registrations and patents for AI-assisted creative works and inventions.[153] They are preliminary, and both agencies are engaged in extensive consultations with stakeholders to gather information before finalizing their policies.[154] Since both agencies have left some questions unanswered about the implementation and effect of their guidance, additional clarity would be beneficial.

The USCO guidance highlights the difficulty of distinguishing between AI and human contributions to blended works, how the USCO will approach this question and the information and evidence that creators will be required to submit, among other things.

Some other potential effects of the USCO’s guidance also need further consideration. For instance, the justification for denying copyright protection to creators who use AI to overcome disabilities seems murky at best when those creators otherwise exercise full creative control and direction over the work. It is also unclear who, if anyone, would be harmed by allowing copyright protection for such AI-assisted creations or why all AI tools universally must be treated so differently than other tools of creation.

[150] Id.
[151] Id.
[152] Id.
[153] Supra 64.; Supra 97
[154] Supra 66.; Supra 67.

The USPTO's guidance regarding inventorship similarly needs clarification. For example, the guidance concludes that a patent application may list all human inventors (i.e., those who contributed significantly to at least one element of the invention) and omit any contribution deemed to be from an AI system.[155] However, it also bans patents for inventions for which the only contribution deemed significant enough is attributed to an AI system independent of any human.[156]

It is unclear whether courts will agree with the USPTO and likewise treat an AI system as an independent contributor or, instead, as a tool of the human inventors. Questions also remain as to whether courts will diverge from the USPTO and conclude that an issued patent is invalid if one aspect of the invention was solely the contribution of an AI system independent of any human, even if the listed (human) inventors were responsible for all other aspects.

Additionally, the guidance creates a new procedural hurdle for patent applicants by requiring for the first time the submission of evidence to assess inventorship in the case of AI-assisted inventions.[157] This seems to indicate that the contributions of human inventors may be assessed against the contributions deemed to be made by the AI system. It is unclear what effect such an analysis will have on the viability of obtaining patents on inventions assisted by AI and whether proving inventorship will be more difficult if the AI system's "contribution" is considered independently. It is also unclear what evidence will be required, what standards patent examiners will apply, and whether these new inventorship inquiries will significantly affect the burdens on applicants and the USPTO.

Similar to the USCO guidance, the USPTO's guidance does not remove all uncertainty from the issue of patent inventorship. The lack of clarity might render U.S. law less favorable to some applicants than the patent laws of other jurisdictions. Although the guidance appears to be concerned with the specter of inventions being mass-produced simply by pressing a button on an AI system, thus far, there has been little indication that such a phenomenon is close at hand.

Generally, inventions require testing, evaluation, and refinement by humans to satisfy patentability requirements and even further development to be incorporated into products and services. To the extent that current patentability requirements prove insufficient, policymakers should focus on narrowly tailored measures that do not broadly deter the application of AI technology and instead strike a balance that preserves incentives to innovate in AI technology and its use. For inventorship, the focus should remain on the human inventors and the sufficiency of their knowledge, understanding, and inventive contributions rather than the tools they used (whether AI or non-AI).

[155] Supra 97.
[156] Id.
[157] Id.

Anticipatory Regulation

Although a range of IP-related issues have been raised by recent developments in AI technology and its applications, some are not yet ripe for government intervention. It will be vital to avoid overreach and understand the potential costs and benefits as much as possible. Any new IP-related legislation or regulations should target specific known issues or problems; tailor definitions, requirements, and consequences narrowly; reduce uncertainty rather than increase it; and focus on improving the ability of the private sector to innovate and creators to thrive.

Key Findings

It is unclear whether legislative action is necessary in some cases, and a number of IP issues are currently in the courts.

Due to the rapidly shifting landscape and tremendous uncertainty of AI's impacts on IP, Congress should exercise caution when considering legislation related to AI and IP.

Generative AI poses a unique challenge to the creative community.

The rapid development of generative AI raises a number of IP-related issues for content creators. While continued progress in AI is welcome, those advancements must not stifle the continued flourishing of human creativity.

It is often difficult for creators to know if their copyrighted works are being used by AI developers.

While the question of fair use in the context of generative AI remains the subject of ongoing litigation, content creators often do not know if their copyrighted works were used in the training of a model.

The global IP policy landscape presents challenges and opportunities to both developers and creators.

A legislative or regulatory environment that significantly increases operational costs for AI developers may cause companies to offshore operations to more permissive environments. For example, the People's Republic of China is already considering steps to encourage relocation to China through more developer-friendly copyright laws.

However, amid advancements in generative AI, a legislative environment that fails to ensure creators' IP rights are protected may harm America's creative community. Furthermore, jurisdictions such as the European Union that are beginning to include transparency requirements for AI developers may encourage creators to seek out opportunities outside of the United States – and this may be true for other jurisdictions that enact robust IP protections for creators.

While some use cases are legitimate and protected forms of expression, the proliferation of harmful deepfakes and digital replicas is a significant and ongoing challenge.

In some cases, digital replicas can be used in the context of speech protected by the First Amendment and for other legitimate purposes. However, some types of deepfakes and digital replicas of individuals' likenesses, voices, and other identifying characteristics are harmful applications of AI technology that are already in use today.

Recommendations

Recommendation: Clarify IP laws, regulations, and agency activity.

Applicable laws, regulations, and agency guidance on AI authorship/inventorship and the ability to obtain IP protection for AI-assisted creative works and inventions should be clarified. Such clarity should ideally promote the development and application of powerful new AI tools for content creation and innovation while still protecting human creators and innovators.

Recommendation: Appropriately counter the growing harm of AI-created deepfakes.

The proliferation of deepfakes and harmful digital replicas is a real and current problem. Although digital replicas and deepfakes have existed for many years, AI technology has vastly amplified the size of the problem by making high-quality, realistic replicas accessible to nearly anyone with little effort.

Congress could address this problem in several ways. One way would be to empower individuals to protect their identity-based rights and establish nationwide protections while avoiding encroaching on speech that is protected by the First Amendment.

CONTENT AUTHENTICITY

Background

Generative AI systems include AI that can generate text, image, video, and audio/voice content.[1] These systems are trained on a large set of existing written, visual, or audio data. The systems identify statistical patterns in this training data and then create novel content that matches these patterns. As generative AI systems continue to be trained with greater amounts of data and more powerful computing resources, they can produce outputs with increasing quality and realism.

When a video, image, or audio is perceived as reflecting a faithful recording of an event, even though it has been generated or substantively manipulated by AI, it is often referred to as a “deepfake.” Deepfakes represent a subset of the general category of synthetic content, or content that is fully or partially altered or created using any audio, video, image creation, or editing tool. Generative AI technologies can be productively used in many ways to stimulate creativity, productivity, and entrepreneurship. Synthetic content has broad applications in marketing, sales, entertainment, and product development. However, synthetic content can also be malicious, harmful, or misleading and be implicated with issues of fraud and consent.

Risks and Harms from Synthetic Content

Risks and harms created by synthetic content, such as financial scams, exist in open information ecosystems like the internet. However, even greater risks arise from AI systems capable of generating realistic synthetic text, images, videos, and audio.

[1] Stryker, Cole, and Mark Scapicchio. “What Is Generative Ai?” IBM, 22 March 2024, www.ibm.com/topics/generative-ai. Some technologies capable of generating content include transformer-based models, generative adversarial networks, and certain types of autoencoders. Bengesi, Staphord, et al. “Advancements in generative AI: A comprehensive review of gans, GPT, autoencoders, diffusion model, and Transformers.” IEEE Access, vol. 12, 2024, pp. 69812–69837, https://doi.org/10.1109/access.2024.3397775.

Generative AI can produce harmful synthetic content much faster and at a lower cost than previous technologies. The most popular contemporary image generation tools have some guardrails to impede misuse, such as functionality that prevents creating content that portrays public figures.[2] Nevertheless, it is relatively easy for a layperson to easily remove such guardrails or acquire generative AI tools that lack sufficient guardrails. Improvements in the capabilities of generative AI systems make it more difficult to detect synthetic content. As a result, while casual observers could easily identify a fake image created with previous technologies, generative AI can produce content that challenges even the most discerning viewer.

> **“While casual observers could easily identify a fake image created with previous technologies, generative AI can produce content that challenges even the most discerning viewer.**

Harms from synthetic content can have concentrated or widespread effects on an individual, organization, community, or target population.[3] For example, synthetic content can be used to malign an individual or perpetrate fraud. In contrast, synthetic content used to distribute misleading or inaccurate information on a social media platform can be spread widely and to targeted populations.

As with traditional content, synthetic content is broadly protected as free speech under the First Amendment. This fundamental constitutional principle underscores the importance of fostering communication and creativity in America. Even if the content is, or might be, synthetic or faked, that alone does not justify attempts to prohibit its creation or distribution.

Americans’ freedom of expression is protected by the First Amendment, even if that expression is conveyed through synthetic content. However, some types of content have been excluded from these protections. For example, abusive material produced involving real children is illegal.[4] And although there is no federal law restricting the use of AI tools to generate nonconsensual intimate imagery, several states have considered laws to curtail the practice.[5]

[2] Usage Policies.” OpenAI, 10 Jan. 2024,https://openai.com/policies/usage-policies/.
[3] Reducing Risks Posed by Synthetic Content, NIST, 30 April 2024, https://airc.nist.gov/docs/NIST.AI.100-4.SyntheticContent.ipd.pdf.
[4] House of Representatives, Congress. 18 U.S.C. 2256 - Definitions for chapter. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2011-title18/USCODE-2011-title18-partI-chap110-sec2256; see also: House of Representatives, Congress. 18 U.S.C. 2252A - Certain activities relating to material constituting or containing child pornography. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title18/USCODE-2023-title18-partI-chap110-sec2252A; see also: House of Representatives, Congress. 18 U.S.C. 1466A - Obscene visual representations of the sexual abuse of children. U.S. Government Publishing Office, https://www.govinfo.gov/app/details/USCODE-2023-title18/USCODE-2023-title18-partI-chap71-sec1466A
[5] Associates, MultiState. “Most States Have Enacted Sexual Deepfake Laws.” Multistate.Ai, 28 June 2024, https://www.multistate.ai/updates/vol-32.

Other laws focus on restricting the nonconsensual dissemination of intimate images of another person, irrespective of the technology used to produce those images.[6]

Synthetic content protected by the First Amendment can still have real or perceived negative effects. For instance, it can normalize activities such as gender-based violence and violence against children. Synthetic content can also discourage political participation or intimidate and silence those targeted. In the case of child sexual abuse material (CSAM), the proliferation of fake images can impede efforts by law enforcement to identify and help victims.[7] When the government attempts to address these harms, it should be careful to uphold constitutional protections. Technical solutions to identify and address the consequences of synthetic content are discussed further below.

Nonconsensual Intimate Images

One of the most pervasive harms from synthetic content generated by contemporary AI systems is the creation and distribution of nonconsensual intimate images (NCII), particularly when accompanied by threats. Contemporary AI systems can replicate an individual's likeness in many forms, including voice-cloning, text-to-speech synthesis, face-swapping, face-morphing, full-body puppetry, and lip-syncing.[8] While many uses of these technologies are relatively benign, they can also be used to cause severe reputational, emotional, and other devastating harms when intimate imagery of a person is widely shared online without their consent.

A 2023 study found that 98% of deepfake videos online are pornographic, with 99% of those being of women.[9] One in three deepfake tools allows users to create nonconsensual pornography.[10] It takes about 30 minutes to create such an image or video at no cost, starting from only one clear image of a face.[11] While celebrities are most commonly the subject of deepfakes, there is also a growing epidemic of teenagers using these deepfake "nudify" apps to create NCII of classmates or teachers in order to bully and harass.[12]

[6] Devendorf, John. "An Overview of Revenge Porn Law." LawInfo, 29 May 2023, https://www.lawinfo.com/resources/sex-crime/revenge-porn.html.

[7] "Generative AI CSAM Is CSAM." National Center for Missing & Exploited Children, 11 March 2024, https://www.missingkids.org/blog/2024/generative-ai-csam-is-csam.

[8] Kietzmann, Jan, et al. "Deepfakes: Trick or Treat?" Business Horizons, vol. 63, no. 2, March 2020, pp. 135–46, https://doi.org/10.1016/j.bushor.2019.11.006.

[9] "2023 State of Deepfakes: Realities, Threats, and Impact." Security Hero, 26 Sept. 2023, https://www.homesecurityheroes.com/state-of-deepfakes/.

[10] Id.

[11] Id.

[12] Cochran, Lexi Lonas. "From Deepfake Nudes to Incriminating Audio, School Bullying Is Going Ai." The Hill, The Hill, 6 June 2024, https://thehill.com/homenews/education/4703396-deepfake-nudes-school-bullying-ai-cyberbullying/.

> **“One in three deepfake tools allows users to create nonconsensual pornography. It takes about 30 minutes to create such an image or video at no cost, starting from only one clear image of a face.**

NCII can be used for harms such as emotional and reputational harm, extortion, silencing political participation, and other malicious or criminal activities.

Harmful uses of this technology are not limited to adults. In one month in 2023, the Internet Watch Foundation found that 3,000 AI-generated images of illegal CSAM were posted to a dark web forum, with the vast majority of these being realistic pseudo-photographs.[13] More than 99% of these images were of girls, primarily aged 7-13.

Further, a 2023 report found hundreds of instances of exploitative images of children in a public dataset used to train AI text-to-image generation models.[14] Other evidence suggests CSAM-trading communities have been able to re-train AI models, possibly with photos of existing victims.[15]

Fraud and Financial Scams

As with instant messaging in the 1990s and digital currencies in the 2010s, scammers are often the earliest adopters of new technologies. As generative AI systems have grown in popularity and availability, scammers have rushed to adopt the technology to enable fraud and social engineering tactics, thereby imposing financial costs on individuals, businesses, and the economy.

One prominent fraud employing synthetic content is using voice cloning technology to impersonate family members or colleagues and attempt to extract money from unknowing victims. For example, scammers impersonated the CEO of an energy company to convince an employee to transfer €220,000 into a foreign bank account.[16]

[13] How AI Is Being Abused to Create Child Sexual Abuse Imagery, Internet Watch Foundation, 1 Oct. 2023, www.iwf.org.uk/media/q4zll2ya/iwf-ai-csam-report_public-oct23v1.pdf.

[14] Thiel, David. “Identifying and Eliminating CSAM in Generative ML Training Data and Models.” Stanford Digital Repository, 20 Dec. 2023, https://purl.stanford.edu/kh752sm9123.

[15] Pfefferkorn, Riana. ADDRESSING COMPUTER-GENERATED CHILD SEX ABUSE IMAGERY: LEGAL FRAMEWORK AND POLICY IMPLICATIONS, February 2024, https://s3.documentcloud.org/documents/24403088/adressing-cg-csam-pfefferkorn-1.pdf.

[16] Stupp, Catherine. “Fraudsters Used AI to Mimic CEO’s Voice in Unusual Cybercrime Case.” WSJ Pro Cybersecurity, The Wall Street Journal, 30 Aug. 2019, www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402.

In another horrifying scenario, a woman was called by a distraught voice impersonation of her daughter, claiming to be kidnapped and demanding a ransom. Jennifer DeStefano testified in front of the Senate Judiciary Committee:

> *"It was my daughter's voice. It was her cries, her sobs. It was the way she spoke. I will never be able to shake that voice out of mind. It's every parent's worst nightmare to hear your child pleading with fear and pain, knowing that they are being harmed and you are helpless and desperate."*[17]

Large Language Models (LLM) also make it easier for malicious actors to generate convincing content for "phishing," "vishing," and "spear-phishing" attacks. Fortunately, existing defenses, such as spam filters and detection methods, are generally capable of detecting and removing these attack messages.[18] It is possible that AI-enabled phishing and spear-phishing attacks may not be more effective at the population level, even though their content is more convincing.[19]

Finally, there are concerns about scam robocalls using AI-generated voice or AI-generated spam message content. Robocalls are already unlawful in many situations under Do Not Call restrictions, which are currently opt-in for most wired and wireless telephone subscribers.

The FCC also has the STIR/SHAKEN technology mitigation that operates in the cellular network and telephone system layer, mandated by the FCC as a result of the 2019 Pallone-Thune TRACED Act *(P.L. 116-115)*.[20] This is an important component in reducing spoof and scam calls at the technical layer that operates beneath the media presentation mode where AI-generated content would be introduced.

The FCC released updates on the status of the rollout of the STIR/SHAKEN tech in March 2023. Further, the FCC released a rulemaking AI-generated voices in robocalls illegal.[21]

[17] DeStefano, Jennifer. "United States Senate Written Statement of Jennifer DeStefano Abuses of Artificial Intelligence." U.S. Senate Committee on the Judiciary, 13 June 2023, www.judiciary.senate.gov/imo/media/doc/2023-06-13 PM - Testimony - DeStefano.pdf.

[18] Marlow, Simon. "Fighting Spam with Haskell." Engineering at Meta, Meta, 26 June 2015, https://engineering.fb.com/2015/06/26/security/fighting-spam-with-haskell/.

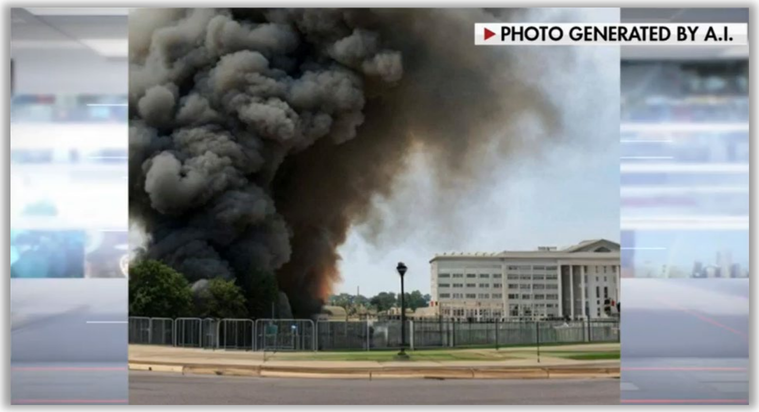
[19] Kapoor, Sayash, and Arvind Narayanan. How to Prepare for the Deluge of Generative AI on Social Media. Knight First Amendment Institute, 16 June 2023, https://knightcolumbia.org/content/how-to-prepare-for-the-deluge-of-generative-ai-on-social-media.

[20] U.S. Federal Communications Commission. Pallone-Thune Telephone Robocall Abuse Criminal Enforcement and Deterrence Act (TRACED Act), 23 Sept. 2021,https://www.federalregister.gov/documents/2021/09/23/2021-14711/pallone-thune-telephone-robocall-abuse-criminal-enforcement-and-deterrence-act-traced-act.

[21] U.S. Federal Communications Commission. "FCC Makes AI-Generated Voices in Robocalls Illegal." 8 Feb. 2024. https://www.fcc.gov/document/fcc-makes-ai-generated-voices-robocalls-illegal.

Integrity of Information

Contemporary AI systems have proven to be effective tools for generating false narratives that both appeal to and influence targeted audiences in various languages.



Source: Fox Business - Fake Pentagon explosion image goes viral on Twitter

These tools can automate the spread of false information, whether it is intended to mislead or not. The examples are numerous: there was a synthetic video of Ukrainian President Zelenskyy appearing to tell his soldiers to surrender[22] and manipulated video claiming a candidate for Chicago mayor promoted police brutality.[23] A synthetic image of an explosion at the Pentagon was widely spread online, in many cases unknowingly,[24] and it caused a 30 point drop in the S&P 500.[25] Recently, AI-generated online news stories have also increased significantly.[26]

Several nation-states have started using synthetic content to undermine people's trust in information ecosystems. For example, a 2023 RAND Corporation report found that the Chinese Communist Party appears to be interested in using AI systems to generate false content aimed at specific populations.[27] Moreover, these researchers showed the viability of using AI systems that generate synthetic text to improve "astroturfing campaigns," which are coordinated campaigns that give the perception of grassroots support for certain issues through deceptive practices.[28]

Some stakeholders have raised the possibility that widely available access to synthetic content generation capabilities will harm the integrity of information on the internet, further eroding trust in media and critical institutions.

[22] Allyn, Bobby. "Deepfake Video of Zelenskyy Could Be 'tip of the Iceberg' in Info War, Experts Warn." NPR, 16 March 2022, https://www.npr.org/2022/03/16/1087062648/deepfake-video-zelenskyy-experts-war-manipulation-ukraine-russia.

[23] Joe Concha, "The impending nightmare that AI poses for media, elections," The Hill, 23 April 2023, https://thehill.com/opinion/technology/3964141-the-impending-nightmare-that-ai-poses-for-media-elections/.

[24] OSINTdefender, X (formerly Twitter), 22 May 2023., https://x.com/sentdefender/status/1660650575569059840.

[25] Ian Krietzberg, "S&P Sheds $500 Billion from Fake Pentagon Explosion," The Street, 22 May 2023, https://www.thestreet.com/technology/s-p-sheds-500-billion-from-fake-pentagon-explosion.

[26] "Tracking AI-Enabled Misinformation: Over 1100 'Unreliable AI-Generated News' Websites (and Counting), Plus the Top False Narratives Generated by Artificial Intelligence Tools." NewsGuard, 19 Aug. 2024, https://www.newsguardtech.com/special-reports/ai-tracking-center/ .

[27] Marcellino, William, et al. The Rise of Generative AI and the Coming Era of Social Media Manipulation 3.0: Next-Generation Chinese Astroturfing and Coping with Ubiquitous AI. RAND Corporation, 2023, https://www.rand.org/pubs/perspectives/PEA2679-1.html.

[28] Id.

As generative AI technology improves and more synthetic content is disseminated via information ecosystems, people may lose confidence in determining whether the content is real or fake.[29]

Further, a mass proliferation of fake content could foment a "liar's dividend," making it easier for people to spread incorrect information to avoid accountability for the truth.[30] While this is a concerning trend, the true impact of synthetic content on information integrity remains unknown.[31]

As synthetic content becomes increasingly common, audiences may shift how they assess the trustworthiness of online content. Institutions designed to address false or misleading content, including fact-checkers and moderators, may adjust their tactics to address the theoretical proliferation of synthetic content meant to deceive.

Unauthorized Use of Likeness for Commercial Purposes

In some cases, an AI-generated likeness of another individual has been used without their consent for commercial gain. For example, the estate of the late George Carlin filed a lawsuit alleging a YouTube creator used an AI voice cloning system to impersonate the comedian and perform jokes based on his previous performances.[32] AI has also been used to generate fake celebrity endorsements for products, which the FTC recently prohibited in a final rule.[33] Several jurisdictions within the United States recognize certain rights related to "personhood" and "right of publicity."[34] Please see the chapter on **Intellectual Property** for more information.

[29] Hao, Karen. "The Biggest Threat of Deepfakes Isn't the Deepfakes Themselves." MIT Technology Review, 10 Oct. 2019, https://www.technologyreview.com/2019/10/10/132667/the-biggest-threat-of-deepfakes-isnt-the-deepfakes-themselves/.

[30] Chesney, Bobby, and Danielle Citron. "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security." California Law Review, December 2019, https://www.californialawreview.org/print/deep-fakes-a-looming-challenge-for-privacy-democracy-and-national-security.

[31] Helmus, Todd C., and Bilva Chandra. Generative Artificial Intelligence Threats to Information Integrity and Potential Policy Responses, RAND Corporation, 16 April 2024, https://www.rand.org/content/dam/rand/pubs/perspectives/PEA3000/PEA3089-1/RAND_PEA3089-1.pdf.

[32] Kuo, Christopher. "George Carlin's Estate Sues Podcasters Over A.I. Episode." The New York Times, 26 Jan. 2024, https://www.nytimes.com/2024/01/26/arts/carlin-lawsuit-ai-podcast-copyright.html.

[33] "Federal Trade Commission Announces Final Rule Banning Fake Reviews and Testimonials." Federal Trade Commission, 14 Aug. 2024, https://www.ftc.gov/news-events/news/press-releases/2024/08/federal-trade-commission-announces-final-rule-banning-fake-reviews-testimonials.

[34] Roesler, Mark, and Garrett Hutchinson. "What's in a Name, Likeness, and Image? The Case for a Federal Right of Publicity Law." American Bar Association, 16 Sept. 2020, www.americanbar.org/groups/intellectual_property_law/publications/landslide/2020-21/september-october/what-s-in-a-name-likeness-image-case-for-federal-right-of-publicity-law/. Please see the intellectual property chapter of this report for additional information.

Copyright Issues

Copyright law protects the rights of creators by granting exclusive rights over their original works. The creation of synthetic content relies on acquiring and modifying other content, which can infringe on the copyright of third parties. However, the copyright doctrine of fair use permits certain uses of third-party content without their permission. Please see the chapter on Intellectual Property for more information.

Elections

AI-generated images, video, and audio can also affect election information in the United States. More advanced AI techniques will increase the effectiveness and the ease of deploying these technologies for sharing inaccurate election information or voter suppression. There have already been attempts to influence voters. There was an AI robocall mimicking President Biden in New Hampshire, discouraging Democrats from voting in the primary,[35] and faked audio portraying President Trump seeming to disparage Republican voters.[36] The scrutiny faced by national-level elections can help quickly identify and correct inaccurate information, but falsehoods about smaller and local elections may go unchecked for longer.

Universities and companies, including AI developers and social media platforms, are developing technical tools and measures to identify, track, and disclose the manipulation of digital media.[37] However, such tools may vary in effectiveness and might not be readily available to the public. Federal agencies are also attempting to address these challenges. The U.S. Election Assistance Commission (EAC) released an “AI Toolkit for Election Officials.”[38] Additionally, the Cybersecurity and Infrastructure Security Agency (CISA) maintains the “Cybersecurity Toolkit and Resources to Protect Elections.”[39] The federal government can also play a key role in providing local election assistance to support AI readiness, election worker preparedness, and cybersecurity through Help American Vote Act (HAVA) Election Security Grants.

Technical Solutions to Identifying Synthetic Content

Although there is no perfect technical solution to harm from synthetic content, there are several promising approaches to promote trust between content producers, content distributors, and the public. Some techniques may be sufficient to address specific goals related to disclosing when content has been created by generative AI.

[35] Michael Graham, “FCC Issues $6M Fine for Bogus Biden Robocalls,” Government Technology, 27 Sept. 2024, https://www.govtech.com/public-safety/fcc-issues-6m-fine-for-bogus-biden-robocalls.

[36] Liles, Jordan. “CLIP Features AI-Generated Trump Voice Calling Republicans ‘Dumbest Group of Voters.’” Snopes, 13 July 2024, https://www.snopes.com/fact-check/trump-republicans-ai-dumbest-voters/.

[37] Anya Schiffrin, et al., “AI Startups and the Fight Against Mis/Disinformation Online: An Update,” German Marshall Fund, 26 July 2023, https://www.gmfus.org/news/ai-startups-and-fight-against-misdisinformation-online-update.

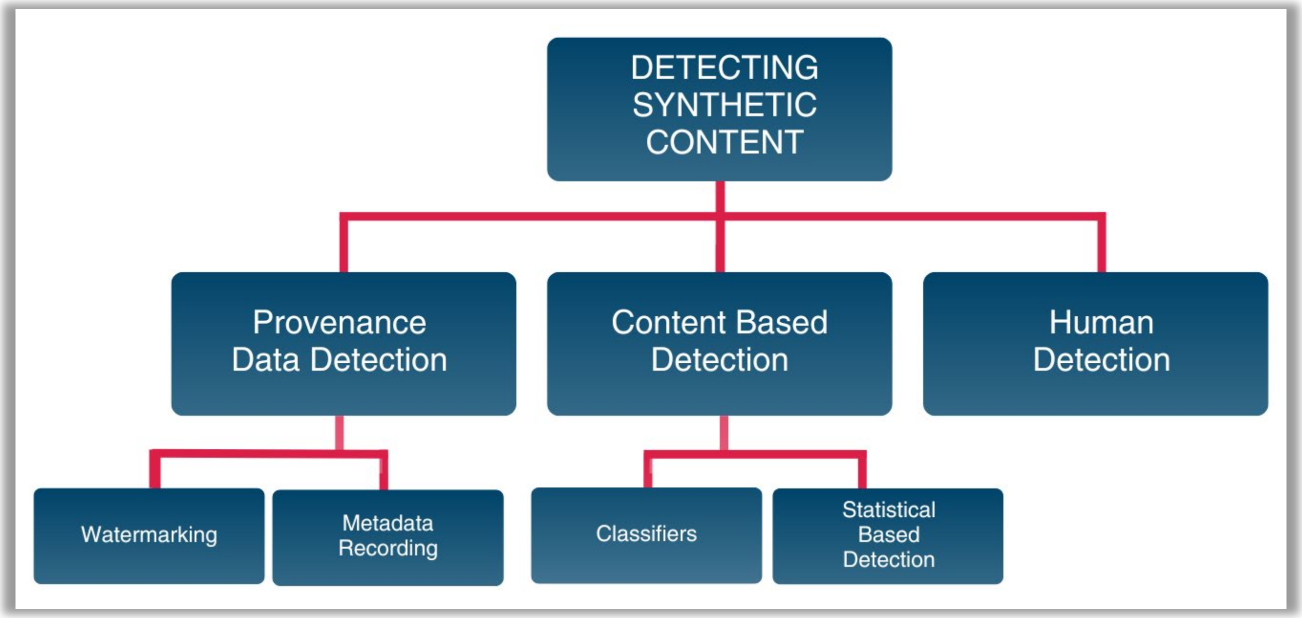
[38] U.S. Election Assistance Commission, “AI Toolkit for Election Officials,” August 2023, https://www.eac.gov/sites/default/files/2023-08/AI_Toolkit_Final_508.pdf.

[39] America’s Cyber Defense Agency, “Election Security,” Cybersecurity & Infrastructure Security Agency, 2024, https://www.cisa.gov/topics/election-security/protect2024.

For example, covert or invisible watermarks are more useful for experts or software than laypeople to determine content authenticity.[40] Sensitive domains, such as national security, may benefit from a multi-pronged approach that utilizes more than one technique to detect synthetic content. For example, critical communications related to election security may require layered technical solutions to prevent overreliance on any one control.[41] Moreover, the effectiveness of particular technical solutions can depend significantly on the type of content being authenticated. Watermarks are readily embedded in images, for instance, but it is more difficult to employ watermarks with text.

Detection

Synthetic content detection refers to the group of techniques and tools used to detect whether content was generated by AI. Detection can broadly fit into three categories: provenance data detection, automated content-based detection, and human authentication.



DETECTING SYNTHETIC CONTENT

Provenance Data Detection

Content Based Detection

Human Detection

Watermarking

Metadata Recording

Classifiers

Statistical Based Detection

Provenance Data Detection

Provenance refers to the origin and history of content. For example, the provenance of an AI-generated image may include information related to its creation (e.g., the model version and user prompt) and any subsequent editing.

[40] Supra 3.
[41] Id.

Provenance detection and tracking techniques use source authentication and integrity verification methods to embed information on digital content to indicate synthetic or authentic origins and identify that embedded information.[42]

Provenance includes digital signatures and cryptographic trust, similar to how internet browsers use encryption and certificates to establish secure connections to banking websites and enable secure logins. An open standard for content provenance and authenticity is being developed by The Coalition for Content Provenance and Authenticity (C2PA), an organization composed of technology companies.[43]

Current methods for provenance data detection employ digital watermarking and metadata recording:

- *Watermarking* involves embedding information in an image to indicate the origins of the content.[44] Watermarks can be overt, meaning they can be perceived directly by humans (e.g., visible changes to an image, audible changes to a sound recording), or covert, meaning they can only be detected by software. Overt watermarks create some level of transparency but are relatively easy to manipulate.

  For example, watermarks in images can be cropped out, altered, or removed entirely unless they cover the majority of the content.[45] On the other hand, the effectiveness of a covert watermark is contingent on whether it can resist being detected and removed from the content—a process that has some probability of failure.[46] Researchers have shown that all invisible watermarks can be removed.[47]

- *Metadata Recording* involves embedding information about digital content to enable content provenance. Metadata is simply information about the content, such as its properties, structure, origin, time and date of creation, author, and more.

  Metadata can be produced whenever content is created, downloaded, or modified, making it easy to manipulate. Metadata can be associated with content through cryptographic processes, such as digital fingerprints and cryptographic signatures.[48]

[42] Id.
[43] Coalition for Content Provenance and Authenticity. C2PA, 2024, https://c2pa.org/.
[44] Srinivasan, Siddarth. "Detecting AI Fingerprints: A Guide to Watermarking and Beyond." Brookings, 4 Jan. 2024, https://www.brookings.edu/articles/detecting-ai-fingerprints-a-guide-to-watermarking-and-beyond/.
[45] "Authenticating AI-Generated Content Exploring Risks, Techniques & Policy Recommendations." Information Technology Industry Council (ITI), January 2024, www.itic.org/policy/ITI_AIContentAuthorizationPolicy_122123.pdf.
[46] Supra 3.
[47] Zhao, X., Zhang, K., Su, Z., Vasan, S., Grishchenko, I., Kruegel, C., ... & Li, L. (2023). Invisible image watermarks are provably removable using generative AI. arXiv preprint arXiv:2306.01953., https://arxiv.org/pdf/2306.01953.
[48] Id.

Provenance data tracking has several limitations. As discussed above, there is no fully robust, reliable watermarking method.[49] Mechanisms to make them more effective can be either resource-intensive or diminish the utility or quality of the content (e.g., some ways of making a watermark more robust can distort the image).[50,51] Likewise, metadata has several limitations, including issues related to privacy and security.[52]

Provenance data tracking can potentially harm information integrity or falsely promote trust in online content. A 2023 study on the provenance of digital content revealed users trusted media more when it had provenance data.[53] However, other research suggests users often fail to differentiate between two similar concepts: the credibility of content's provenance information and the credibility of the content itself.[54] In other words, a verified history of some digital content does not necessarily imply that the content is accurate and unaltered. As a result, provenance data can potentially be exploited to lend credibility to synthetic content by misleading people into believing it is real.[55] Additionally, legitimate information that lacks watermarking or metadata for any number of reasons (e.g., lack of resources, security impediments) may be incorrectly dismissed as fake.

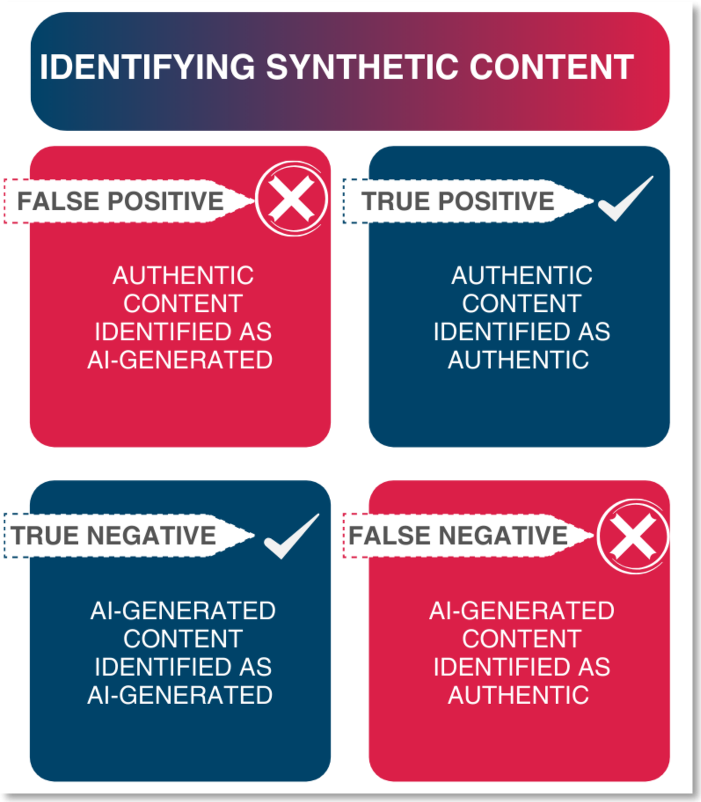
Content-based Detection

Content-based detection consists of post-hoc techniques to identify synthetic content after it has been generated. These techniques primarily involve classifiers, AI tools designed and trained to classify data into categories such as "fake" and "not fake." Classifiers are already used together with many security and risk mitigation technologies, such as in email spam filtering, to identify whether content is generally 'acceptable' or not.[56] By analyzing patterns in the style of images and other content, classifiers attempt to detect if the content is AI-generated.[57]

Some classifier systems are better than others at identifying AI-generated content, but they can make errors in both directions; a classifier can mistake AI-generated content as authentic ("false negative") and can mistake authentic content for synthetic content ("false positive").

[49] Kate Knibbs. "Researchers Tested AI Watermarks—and Broke All of Them." Wired, https://wired.me/culture/researchers-tested-ai-watermarks-and-broke-all-of-them/.
[50] Begum, Mahbuba, and Mohammad Shorif Uddin. "Digital Image Watermarking Techniques: A Review." MDPI, Multidisciplinary Digital Publishing Institute, 17 Feb. 2020, https://www.mdpi.com/2078-2489/11/2/110.
[51] Wan, Wenbo, et al. "A Comprehensive Survey on Robust Image Watermarking." ScienceDirect, Elsevier, 2 March 2022, https://www.sciencedirect.com/science/article/abs/pii/S0925231222002533#b0310.
[52] Supra 3.
[53] Feng, K. J. Kevin, et al. "Examining the Impact of Provenance-Enabled Media on Trust and Accuracy Perceptions." arXiv, 2023, https://arxiv.org/abs/2303.12118.
[54] Feng, Kevin, et al. "Examining the Impact of Provenance-Enabled Media on Trust and Accuracy Perceptions." Proceedings of the ACM on Human-Computer Interaction, vol. 7, no. CSCW2, Association for Computing Machinery, September 2023, pp. 1–42, https://dl.acm.org/doi/pdf/10.1145/3610061.
[55] Supra 47.
[56] Longshot AI, The Science Behind AI Content Detectors: Insights into Their Methods and Accuracy, 15 April 2024, https://www.longshot.ai/blog/working-of-ai-detectors
[57] Id.

Currently, the effectiveness of classifiers ranges from 20% to 98%, depending on the classifier and the AI that generates the content.[58]



IDENTIFYING SYNTHETIC CONTENT

FALSE POSITIVE ✗
AUTHENTIC CONTENT IDENTIFIED AS AI-GENERATED

TRUE POSITIVE ✓
AUTHENTIC CONTENT IDENTIFIED AS AUTHENTIC

TRUE NEGATIVE ✓
AI-GENERATED CONTENT IDENTIFIED AS AI-GENERATED

FALSE NEGATIVE ✗
AI-GENERATED CONTENT IDENTIFIED AS AUTHENTIC

Another type of content-based detection is statistical detection.[59] Statistical detection is a method of identifying statistical anomalies in the distribution of pixels, speech, and frequencies that correlate to artificially generated content.[60] This method is less mature than classifiers.

The National Institute of Standards and Technology (NIST) has identified several issues, including generalizability, reproducibility, interpretability, explainability, and data input, which affect detection methods across different types of synthetic content.[61]

Further, detectors are designed for specific types of media, such as photos, video, text, or audio, and each has its own challenges and limitations. For example, synthetic audio detection developed for one language will not perform well when detecting synthetic audio in other languages or dialects.[62]

Human Authentication

Human authentication refers to human involvement in determining whether an image or content was AI-generated. Human verification of content as potentially AI-generated is one example of human authentication.[63] It leverages the power of human expertise and possibly the perspectives of a wide range of people.

[58] Id.
[59] Id.
[60] Fernando Martin-Rodriguez et al. "Detection of AI-Created Images Using Pixel-Wise Feature Extraction and Convolutional Neural Networks." National Library of Medicine, PubMed Central, 2023, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10674908/.
[61] Supra 3.
[62] Cuccovillo, Luca, et al. "Open Challenges in Synthetic Speech Detection." arXiv, 26 Jan. 2023, https://arxiv.org/pdf/2209.07180.
[63] Id.

Although human authentication is more time-consuming, it can be used judiciously in the most critical parts of a system or output where accurate authentication is most needed.

Research suggests humans can judge whether an image is AI-generated with many contemporary models,[64] but there are significant challenges in determining whether the text is AI-generated.[65] There are several drawbacks to human authentication. Like many manual approaches to technical or automated tasks, this approach requires large numbers of people sifting through content.

Detection methods relying on humans are difficult to scale because they require extensive labor and high costs due to the large amounts of content to be reviewed. Humans' detection accuracy will also differ, possibly greatly, based on their personal experiences and expertise. Humans also may be unable to identify all cases of AI-generated content, especially when such content is designed to fool or deceive humans.[66]

Digital Identity

The increasing prevalence of fraud—whether aided by synthetic content or not—is driving organizations to improve how they authenticate the identities of real individuals. The Federal Trade Commission (FTC) received over 1.1 million identity theft complaints in 2023, an increase of 6.3% from 2022.[67] Online platforms are also seeing significant rises in AI-generated accounts. For example, a Stanford researcher found over 1,000 AI-generated profiles on LinkedIn in 2022.[68]

Digital identity management systems are used to provision identities to users and manage authentication, authorization, and data sharing based on identity. These systems may operate within an organization, across several organizations, or on the internet. In its most basic implementation, a digital ID simply recreates a physical ID, such as a driver's license, in a digital format. In more sophisticated implementations, a fully integrated digital identity system can provide verification processes in both the online and physical world.

Unlike other means for authenticity controls and detection, digital identity management requires verifying and, therefore, knowing the identity of the person accessing the

[64] Groh, Matthew, et al. "Human Detection of Political Speech Deepfakes across Transcripts, Audio, and Video." arXiv, 2022, https://arxiv.org/abs/2202.12883.

[65] Kreps, Sarah, R. Miles McCain, and Miles Brundage. "All the News That's Fit to Fabricate: AI-Generated Text as a Tool of Media Misinformation." Journal of Experimental Political Science 9.1 (2022): 104–117. https://www.cambridge.org/core/journals/journal-of-experimental-political-science/article/abs/all-the-news-thats-fit-to-fabricate-aigenerated-text-as-a-tool-of-media-misinformation/40F27F0661B839FA47375F538C19FA59.

[66] Supra 47.

[67] "As Nationwide Fraud Losses Top $10 Billion in 2023, FTC Steps Up Efforts to Protect the Public." Federal Trade Commission, 9 Feb. 2024, https://www.ftc.gov/news-events/news/press-releases/2024/02/nationwide-fraud-losses-top-10-billion-2023-ftc-steps-efforts-protect-public.

[68] Bond, Shannon. "That Smiling Linkedin Profile Face Might Be a Computer-Generated Fake." NPR, 27 March 2022, https://www.npr.org/2022/03/27/1088140809/fake-linkedin-profiles.

website or service. Once the person's identity is verified, it is easier to establish links between individuals and the content they create or modify.

In addition, online platforms that assist in the spread of AI-generated content could utilize verified identity to reduce fraud in AI-generated content.

There are technical solutions that may help support better digital identity management and content authentication, including cryptographic and biometric solutions for digital identity management. Additionally, novel solutions utilizing blockchain technology would provide a tamper-proof history of digital content, preventing people from deleting all provenance data from content and reposting the content as genuine.[69]

However, there are accessibility, privacy, security, and user control challenges for digital identity management systems. If not implemented responsibly, digital IDs might permit increased location-tracking and user profiling, potentially leading to significant privacy and economic harms.

Similarly, blockchain technology, if used inappropriately, could reveal a public, immutable record of a person's information exchanges, including where, when, and why a digital ID was requested. These systems may not even empower users with privacy controls to edit and delete their own data.

Further, centralized storage of personal and sensitive information, contrary to security best practices, would be an appealing target for hackers. Privacy-by-design and security-by-design requirements may ameliorate some of these drawbacks. Federal coordination and research investments may also help promote interoperable next-generation remote identity proofing and verification systems that squarely address these privacy and security challenges.

The federal government has already adopted some identity management technology into its own systems. However, some of these systems are legacy infrastructure and are susceptible to attacks.[70]

Rather than establishing new federal digital ID requirements, the government can be most effective by facilitating coordination and supporting the development of useful technologies. The government has successfully leveraged public-private partnerships several times over the last few decades to help identify and address challenges with digital identity technology.[71]

[69] Lesavre, Loïc, et al. A Taxonomic Approach to Understanding Emerging Blockchain Identity Management Systems, NIST, 14 Jan. 2020, https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.01142020.pdf.

[70] Better Identity in America: A Blueprint for Policymakers, The Better Identity Coalition, July 2018, https://static1.squarespace.com/static/5a7b7a8490bade8a77c07789/t/5d07cd2eca832a0001656624/1560792371066/Better_Identity_Coalition%2BBlueprint%2B-%2BJuly%2B2018.pdf.

[71] Electronic Authentication Partnership (EAP) "EAP Merges into Liberty Alliance." Network Computing, 10 Sept. 2007, https://www.networkcomputing.com/network-infrastructure/eap-merges-into-liberty-alliance.; see also: Trust Framework Solutions (TFS) program "Identity, Credential, and Access Management (ICAM): CISA." Cybersecurity and Infrastructure Security Agency CISA, 15 March 2018, https://www.cisa.gov/safecom/icam.; see also: National

Other Considerations

Technical solutions to various content authentication challenges are stymied by a lack of implementation by online platforms. Metadata usually gets removed from social media websites when content is uploaded onto the platform. For a metadata or watermark solution to be effective, it would also require an interoperable framework accepted by both content producers and distributors. As a result, open-source AI models pose a significant challenge to implementing several technological solutions, including provenance data tracking.[72]

In February 2024, 25 technology companies signed an agreement in Munich to combat deceptive uses of AI in elections. However, adoption across these platforms seems to be inconsistent.[73] For example, Alphabet and X have different approaches to election-related synthetic content.[74,75] Similarly, while hundreds of companies have started to use the C2PA standard to authenticate their content, only a few social media companies recognize and support these content credentials.

Strategy for Trusted Identities in Cyberspace (NSTIC) "NATIONAL STRATEGY FOR TRUSTED IDENTITIES IN CYBERSPACE: Enhancing Online Choice, Efficiency, Security, and Privacy." NIST, April 2011, https://www.nist.gov/system/files/documents/2016/12/08/nsticstrategy.pdf

[72] Srinivasan, Siddarth. "Detecting AI Fingerprints: A Guide to Watermarking and Beyond." Brookings, 4 Jan. 2024, https://www.brookings.edu/articles/detecting-ai-fingerprints-a-guide-to-watermarking-and-beyond/.

[73] A Tech Accord to Combat Deceptive Use of AI in 2024 Elections, AI Elections Accord, 16 Feb. 2024, https://www.aielectionsaccord.com/uploads/2024/02/A-Tech-Accord-to-Combat-Deceptive-Use-of-AI-in-2024-Elections.FINAL.pdf.

[74] Reuters. "Alphabet to Limit Election Queries Bard and AI-Based Search Can Answer." Reuters, 19 Dec. 2023, https://www.reuters.com/technology/alphabet-limit-election-queries-bard-ai-based-search-can-answer-2023-12-19/.

[75] Woollacott, Emma. "X Lifts Ban On Political Ad." Forbes, Forbes Magazine, 30 Aug. 2023, https://www.forbes.com/sites/emmawoollacott/2023/08/30/x-lifts-ban-on-political-ads/.

Key Findings

Synthetic content has many beneficial uses, but if used improperly it can create harms and undermine confidence in information integrity.

Bad actors can use synthetic content to commit fraud, spread false information, and target individuals. Addressing these harms is important and must also be done within the context of protecting First Amendment rights.

There is currently no single, optimal technical solution to content authentication.

While policymakers have a broad set of tools to address challenges with synthetic content, currently, no single tool or initiative is likely to succeed on its own.[76] Many technical solutions could be manipulated to give fake content the veneer of authenticity, exacerbating the information integrity issues discussed above. For example, watermarks can be removed, faked, or rendered ineffective without undue effort.

Technical literacy would help with the content authenticity challenges but would not be sufficient.

Public education about content authentication challenges would not address all challenges. AI literacy campaigns could educate the public on content authenticity issues, but even sufficiently knowledgeable people could still fall prey to inauthentic content.[77]

Digital identity technology allows a person online to verify who they are and reduces fraud.

If major privacy and security concerns are addressed, such technology may allow a person online to prove their identity to other users and online platforms. Once the person's identity is verified, it is easier to reduce fraud perpetrated through the digital content they create, modify, or disseminate.

[76] Helmus, Todd C., and Bilva Chandra. Generative Artificial Intelligence Threats to Information Integrity and Potential Policy Responses, RAND Corporation, 16 Apr. 2024, https://www.rand.org/content/dam/rand/pubs/perspectives/PEA3000/PEA3089-1/RAND_PEA3089-1.pdf.

[77] Research from the Centre for Research on Extremism and Security Threats has shown media literacy has little to no effect on whether people shared disinformation. Many people still engage with synthetic content knowing it to be AI-generated.
Buchanan, Tom. "Why Do People Share Disinformation On Social Media?" Centre for Research and Evidence on Security Threats, 4 Sept. 2020, https://crestresearch.ac.uk/resources/disinformation-on-social-media/.

Recommendations

Recommendation: Support a risk-based, multipronged approach to content authenticity.

Since there is no single solution to content authenticity challenges, Congress should encourage the development of several solutions. Specifically, a combination of evidence-based education programs, technical solutions, and policy solutions to content authenticity challenges would enable the public to navigate the increasingly problematic sea of digital content.

A future multipronged approach to digital content authenticity may focus on demonstrating the provenance of authentic content. In a future where convincing synthetic content can be created and distributed easily and inexpensively, people may eventually assume that content is synthetic unless proven authentic. Creators of authentic content would be incentivized to demonstrate that their content is authentic and unaltered by enabling the provenance of their content to be revealed. In this case, policymakers may still need to educate the populace about understanding and utilizing provenance information. Public and private sector organizations may also need additional incentives to encourage adoption and promote network effects.

Recommendation: Support technical solutions to content authenticity.

Technical solutions can potentially improve information integrity and transparency in information systems. Congress should support additional research, development, and standardization activities related to technical approaches to detecting synthetic content. This could include legislation to authorize research, development, and demonstration to improve the maturation and commercialization of technical solutions to synthetic content. Congress should work with industry to support a standardized ecosystem for technical solutions to synthetic content, such as the standardization of these technical solutions, whether through pre-standardization research, public-private partnerships, direct engagement in international standard setting, or the development of voluntary standards and guidelines for addressing synthetic content.[78]

Congress should also explore whether to authorize activities to support government adoption and interagency coordination for a particular technical solution to content authentication. For example, legislation could encourage the development and adoption of next-generation digital identity systems appropriately.[79]

[78] Additional information is included in the Government Use and Standards chapter.

[79] Berkman, Daniel. Path to Digital Identity in the United States. Information Technology and Innovation Foundation, 23 Sept. 2024, https://itif.org/publications/2024/09/23/path-to-digital-identity-in-the-united-states/.

Recommendation: Address demonstrable harms, not speculative harms of synthetic content.

Policymakers should avoid hyperbolizing the potential harms of synthetic content on our information ecosystems and instead target concrete solutions to address demonstrable harms. Purely technical solutions will not address the risks and harms on their own. Generative AI has significant benefits, and AI image manipulation is not inherently more problematic than image manipulation with long-standing tools such as Adobe Photoshop. Therefore, Congress should explore legal solutions that are narrowly tailored to prevent specific harmful applications of this technology.

Electoral speech is one area where solutions must be nuanced and designed to protect First Amendment rights. Broad or technology-specific policies may create uncertainty and inadvertently chill protected political expression.[80] For more information on this topic, see the **Civil Rights and Civil Liberties** chapter.

Recommendation: Identify the responsibilities of AI developers, content producers, and continent distributors when it comes to synthetic content.

Congress should examine legislation that helps create or identify the legal responsibilities of AI developers, content producers, and content distributors regarding synthetic content. The federal government could play a role in clarifying legal responsibilities for AI developers, content producers, and content distributors. Some responsibilities may include a requirement that relevant parties disclose when content is synthetic, technical requirements underpinning notice and attribution, and other mechanisms to increase the transparency of information.

Recommendation: Examine existing laws related to harmful synthetic content.

Congress should evaluate laws that govern the specific harm created by synthetic content, such as copyright, privacy, and tort law, to determine whether the existing law sufficiently protects against these harms.

Recommendation: Ensure victims have the necessary tools.

Congress should investigate whether victims have sufficient ability to seek redress for harms from digital content, such as NCII, from those who create or distribute these forgeries, as well as legal barriers preventing such redress. Congress should consider other redress mechanisms for victims, such as civil penalties for cases involving AI fraud and NCII.

[80] United States v. Alvarez, 567 U.S. 709 (2012). See also Supra note 1 at 4–5.

OPEN & CLOSED SYSTEMS

Background

Some general-purpose AI systems, referred to as "foundation models," are trained on such large quantities of data that they can be adapted to a wide range of downstream tasks.[1] Many of these AI models are closed because there is limited or no public access to their inner workings. In contrast, many other companies have released their models as "open" because their components may be inspected and are accessible over the internet.[2]

Open models offer many benefits, such as greater customization, reproducibility, transparency, innovation, and accessibility for a thorough evaluation. Importantly, because they are often available for free, open models promote competition by diversifying and expanding the number of individuals and companies that can participate in AI research and development. On the other hand, the availability of powerful open models increases the risk that malicious actors might use them to cause harm, including perpetrating financial fraud, threatening national security, or large-scale identity theft.

Despite often being characterized as either open or closed, there is in fact a continuum of different forms of AI model availability and transparency. For example, any of a model's numerous components can be made available, such as its underlying architecture, the model weights, the training or fine-tuning data, the source code to create, train, or run the model, and associated model documentation. Therefore, different parts of a model can be made open while others remain closed.

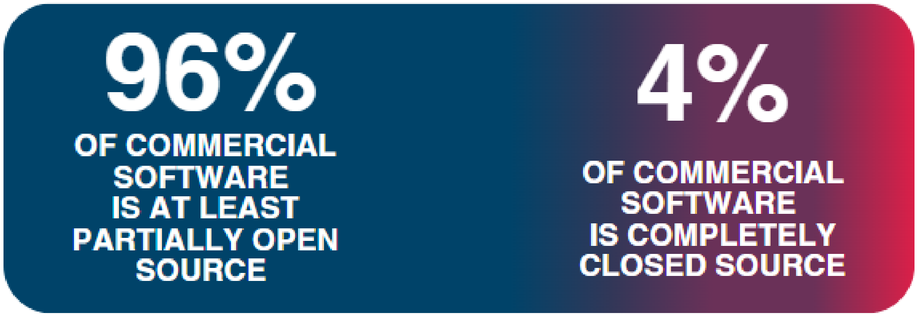
[1] Bommasani, Rishi et. al., "On the Opportunities and Risks of Foundation Models." arXiv, 2022. https://arxiv.org/pdf/2108.07258.

[2] Solaiman, Irene. "The Gradient of Generative AI Release: Methods and Considerations." arXiv, 2023. https://arxiv.org/abs/2302.04844.

Some developers have chosen to limit access to various elements of their models through user-focused web interfaces or application programming interfaces (APIs). Developers can also release their models through structured access mechanisms, sequentially releasing them to select recipients to allow feedback and adjustment before a wider release. Control of model access may also be accomplished with licensing agreements.

Much of the discussion around the risks of open AI models has focused on the release of one major component: model weights. Weights are a series of numerical parameters within a model that are both established by the training process and, once trained, determine the behavior of a model in operation. In most situations, anyone who receives the model weights can run the models.

Therefore, when developers release open model weights, they relinquish exclusive control over the later use of their models. Users who receive only the model weights do not have access to the initial training data or training processes.



96% OF COMMERCIAL SOFTWARE IS AT LEAST PARTIALLY OPEN SOURCE

4% OF COMMERCIAL SOFTWARE IS COMPLETELY CLOSED SOURCE

Source: 1,067 codebases analyzed by the Black Duck Audit Services

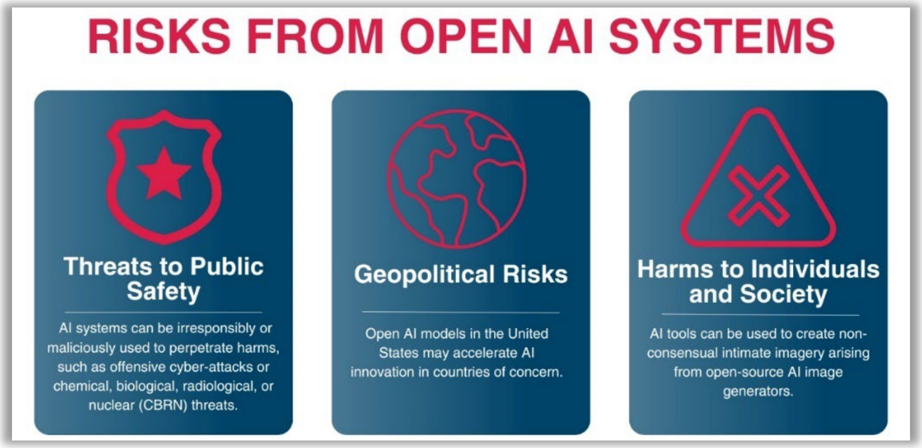
However, they could adjust the model by “fine-tuning” (additionally training) the model, typically to perform better at some specified task. Fine-tuning a model is usually much less computationally intensive than the initial training. Open models, therefore, enable a wider range of parties to fine-tune than the number who have the resources to create their own model.

Downstream modification of a model can thwart the model developer's original intent. For example, researchers have shown that it is possible to purposefully or inadvertently fine-tune a model to remove the guardrails established by the model developer.[3] In contrast, closed model developers can restrict or block access to their models, preventing users from altering them and forestalling malicious activities. However, users with sufficient technical skills may still be able to circumvent the safeguards of closed AI models.

[3] Qi, Xiangyu et. al., “Fine-Tuning Aligned Language Models Compromise Safety, Even When Users Do Not Intend To!” arXiv. 2023, https://arxiv.org/pdf/2310.03693.

Risks from Open AI Systems

There are three primary areas of potential risk associated with open AI systems: threats to public safety from irresponsible or malicious use of open AI systems, accelerated AI innovation in countries of concern, and harm to individuals and society exacerbated by the power of open AI systems.[4]



RISKS FROM OPEN AI SYSTEMS

Threats to Public Safety

AI systems can be irresponsibly or maliciously used to perpetrate harms, such as offensive cyber-attacks or chemical, biological, radiological, or nuclear (CBRN) threats.

Geopolitical Risks

Open AI models in the United States may accelerate AI innovation in countries of concern.

Harms to Individuals and Society

AI tools can be used to create non-consensual intimate imagery arising from open-source AI image generators.

Threats to Public Safety

Powerful, open AI systems may pose risks to public safety because they can be irresponsibly or maliciously used to perpetrate harms, such as offensive cyber-attacks or chemical, biological, radiological, or nuclear (CBRN) threats.[5]

Some threats, such as advanced cyberattacks, remain largely speculative but may be generally available to malicious actors because they would not require specific physical implements.

In contrast, many CBRN risks would be more difficult to perpetrate because they would also require physical products, such as nuclear material or machines used to synthesize biological agents.

While significant safeguards against CBRN attacks already exist, protecting against AI-enabled CBRN threats may require additional hardening of physical defenses. For example, synthetic biology is an area where open AI systems could democratize harmful information.[6] However, even in these situations, malicious parties seeking to use AI systems to cause harm would have to exploit existing vulnerabilities. Therefore, mitigating against misuse of AI-enabled attacks could require improving defenses at sites where biological agents are synthesized.[7]

[4] Dual-Use Foundation Models with Widely Available Model Weights, National Telecommunications and Information Administration, July 2024, https://www.ntia.gov/sites/default/files/publications/ntia-ai-open-model-report.pdf.
[5] Urbina, Fabio et. al., "Dual-use of artificial intelligence powered drug discovery," Nature Machine Intelligence, vol. 3, March 2022, pp. 189-191, https://www.nature.com/articles/s42256-022-00465-9.epdf.
[6] Mouton, Christopher et. al., "The Operational Risks of AI in Large-Scale Biological Attacks," RAND Corporation, 2023, https://www.rand.org/pubs/research_reports/RRA2977-1.html.
[7] Crawford, Forrest et. al., "Securing Commercial Nucleic Acid Synthesis," RAND Corporation, 2024, https://www.rand.org/content/dam/rand/pubs/research_reports/RRA3300/RRA3329-1/RAND_RRA3329-1.pdf.

Geopolitical Risks

A second risk is the geopolitical concern that open AI models in the United States may accelerate AI innovation in countries of concern. Some of the most powerful foreign AI systems are built on open-source American technology;[8] likewise, some top U.S. models are built on foreign open-source systems.[9] While the full nature of these risks remains unknown, restricting access to open AI models could affect U.S. leadership in setting global AI norms and international cooperation on AI systems.

Harms to Individuals and Society

The final risk is that the immense power of open AI tools can be used to harm individuals and society, whether deliberately or negligently. Some of these risks exist already. For example, there has been a significant increase in AI-generated child sexual abuse material (CSAM) and AI-generated non-consensual intimate imagery (NCII) arising from open-source AI image generators.[10]

Cost-Benefit Analysis

Policymakers must weigh the benefits and potential risks of open AI systems. Open AI models enhance competition, innovation, and research. Open development allows developers of all sizes to easily adapt or fine-tune models using propriety data sets, empowering groups that otherwise would have access to advanced AI technology. Open innovation ecosystems also promote competition in downstream markets and reduce market concentration.[11]

Increased openness and transparency along the AI value chain also make it easier to analyze AI systems to ensure compliance with applicable laws.[12]

Similarly, more open ecosystems, including systems with technical safe harbor provisions for public interest research, allow researchers to assess model risks and vulnerabilities more thoroughly, improving overall safety and understanding. Finally, openness and transparency from open-source and open science can lead to better AI governance models.[13]

[8] Mozur, Paul, et al. “China’s Rush to Dominate A.I. Comes with a Twist: It Depends on U.S. Technology.” The New York Times, 21 Feb. 2024, www.nytimes.com/2024/02/21/technology/china-united-states-artificial-intelligence.html.
[9] Tobin, Meaghan and Metz, Cade. “China Is Closing the A.I. Gap with the United States.” The New York Times, 25 July 2024, www.nytimes.com/2024/07/25/technology/china-open-source-ai.html.
[10] Morrish, Lydia. “The Dark Side of Open-Source AI Image Generators.” Wired, Conde Nast, 6 March 2024, www.wired.com/story/dark-side-open-source-ai-image-generators/.
[11] Sayash Kapoor and Rishi Bommasani, et al., “On the Societal Impact of Open Foundation Models.” arXiv, 2024. https://arxiv.org/abs/2403.07918.
[12] Bankston, Kevin and Hodges, Jennifer, et al., “Openness and Transparency in AI Provide Significant Benefits for Society,” Letter to Department of Commerce, Center for Democracy and Technology, Mozilla, and 23 organizations, 25 March 2024, https://cdt.org/wp-content/uploads/2024/03/Civil-Society-Letter-on-Openness-for-NTIA-Process-March-25-2024.pdf.
[13] “Openness and Transparency in AI Provide Significant Benefits for Society,” RStreet, 25 March 2024, https://www.rstreet.org/outreach/coalition-letter-openness-and-transparency-in-ai-provide-significant-benefits-for-society/.

However, these benefits have some limitations. Open ecosystems are unlikely to reduce market concentration in certain parts of the AI supply chain, such as AI hardware.[14] Further, openness alone is unlikely to replace the need for appropriate government agencies to provide meaningful oversight of AI competition.[15]

After significant stakeholder input and weighing benefits and risks, the Department of Commerce produced a report titled "Dual-Use Foundation Models with Widely Available Model Weights" in July 2024.[16] The report focused on the likelihood of threats rather than focusing only on the most dangerous threats. This idea, known as "marginal risk," has policymakers consider the additional risk from open AI models compared to pre-existing technologies and closed AI models.

> **"Open AI models enhance competition, innovation, and research.**

The report finds that "current evidence is not sufficient to definitively determine either that restrictions on such open-weight models are warranted, or that restrictions will never be appropriate in the future." Instead, the report recommends the government actively monitor a portfolio of risks that could arise from these issues and prepare accordingly.

[14] Widder, David et. al. "Open (For Business): Big Tech, Concentrated Power, and the Political Economy of Open AI," Social Science Research Network, 2023. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4543807.
[15] Id.
[16] Supra 4.

Key Findings

Open AI models encourage innovation and competition.

Open-source ecosystems foster significant innovation and competition in AI systems. Many of the most important discoveries in AI were made possible by open-source and open science.[17] The open-source ecosystem makes up roughly 96% of commercial software.[18] The U.S. government, including the Department of Defense, is one of the biggest users and beneficiaries of open-source software.[19]

There is currently limited evidence that open models should be restricted.

The marginal risk approach employed in the Department of Commerce report shows there is currently no reason to impose restrictions on open-weight models. However, future open AI systems may be powerful enough to require a different approach.

[17] Uszkoreit, Jakob. “Transformer: A Novel Neural Network Architecture for Language Understanding.” Google Research, Google, 31 Aug. 2017, https://research.google/blog/transformer-a-novel-neural-network-architecture-for-language-understanding/.

[18] “2024 Open Source Security and Risk Analysis Report.” Synopsys, February 2024, https://www.blackduck.com/resources/analyst-reports/open-source-security-risk-analysis.html#introMenu.

[19] Dahlgren, Masao. “Defense Priorities in the Open-Source AI Debate.” Center for Strategic and International Studies, 19 Aug. 2024, www.csis.org/analysis/defense-priorities-open-source-ai-debate.

Recommendations

Recommendation: Encourage innovation and competition in the development of AI models.

Congress should bolster openness in AI model development and use while continuing to ensure models have appropriate safeguards. Legislation could authorize programs at the National Science Foundation (NSF), National Institute of Standards and Technology (NIST), Department of Energy (DOE), and the Department of Defense (DOD) to improve pathways for open-source ecosystems and improve model cybersecurity, privacy, and governance in these environments. This could include helping to set norms about technical safe harbors for public interest AI researchers, direct incentives to support open-source development, and more. Further, legislation could explore interagency coordination and strategies to support open-source and open-science ecosystems, including through good governance.

Recommendation: Focus on demonstrable harms and physical threats.

Congress should not seek to impose undue burdens on developers in the absence of clear, demonstrable risk. Instead, Congress should work to protect existing vulnerabilities that are exacerbated by highly capable AI systems. Congress and the Administration should focus on the “marginal risk” and “marginal benefit” of AI systems when developing policy that may affect open model development.

Recommendation: Evaluate chemical, biological, radiological, or nuclear (CBRN) threats in light of AI capabilities.

Congress should investigate and take appropriate action to reduce the risk of CBRN attacks augmented with advanced AI systems. Regulations should restrict physical tools needed to carry out attacks rather than restrict access to information already available without AI. Congress should identify gaps in existing regulations and enforcement regarding misuse scenarios and consider legislation to address any gaps discovered.

The Committees on Homeland Security, Judiciary, Energy and Commerce, Foreign Affairs, Armed Services, and Science, Space, and Technology should explore these risks and consider legislation accordingly.

Recommendation: Continue to monitor the risks from open-source models.

The available evidence today gives us no reason to restrict the availability of open-source AI models. Nevertheless, Congress should establish mechanisms that allow it to monitor any risks that might arise from open-source models in the future. With advanced warning of altogether new risks or increased likelihood of known risks, Congress could better respond with appropriate policies in the future.

ENERGY USAGE & DATA CENTERS

Background

After years of minimal growth, demand for electrical power in the United States is projected to increase through 2030[1] at a rate not seen for decades. According to the Department of Energy (DOE), electricity consumption in the United States has grown at a steady 0.5% per year in the last two decades.[2] However, recent estimates suggest an annual growth of at least 0.9% through the end of the decade and an increase in the five-year cumulative growth forecast from 2.6% to 4.7%.[3]

This growth will be fueled by a surge in the number of data centers, expanded uses of artificial intelligence (AI) by data centers, onshoring of manufacturing, and increased electrification.[4] Although these changes can promote local and statewide economic development, they also create new challenges. Accompanying the predictions of soaring demand are warnings that the electric grid cannot reliably meet future needs. To responsibly steward our energy future, we must understand the energy usage and efficiency issues related to AI adoption, development, and deployment.

[1] IEA (2024), Electricity 2024, IEA, Paris https://www.iea.org/reports/electricity-2024; or see: John D. Wilson and Zach Zimmerman. “The Era of Flat Power Demand is Over.” Grid Strategies LLC, December 2023, https://gridstrategiesllc.com/wp-content/uploads/2023/12/National-Load-Growth-Report-2023.pdf; or see: Robert Walton, "U.S. electricity load growth forecast jumps 81% led by data centers, industry: Grid Strategies." Utility Dive, 13 Dec 2023, https://www.utilitydive.com/news/electricity-load-growing-twice-as-fast-as-expected-Grid-Strategies-report/702366/; or see: Scott DiSavino. “U.S. Power Use to Reach Record Highs in 2024-2025 - EIA.” Reuters, 6 Feb. 2024, https://www.reuters.com/world/us/us-power-use-reach-record-highs-2024-2025-eia-2024-02-06/.

[2] U.S. Energy Information Administration. Electricity Annual Data. Energy Information Administration, October 2024, https://www.eia.gov/electricity/annual/.

[3] Supra 1

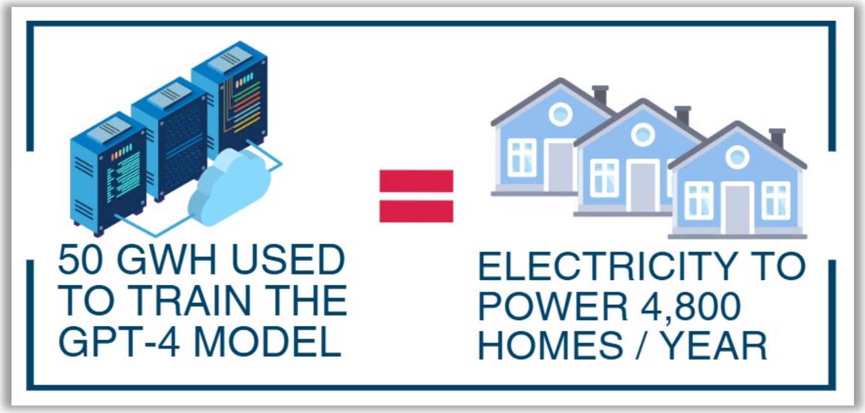
[4] Cy McGeady, "Strategic Perspectives on U.S. Electric Demand Growth." Center for Strategic and International Studies. 2023, https://www.csis.org/analysis/strategic-perspectives-us-electric-demand-growth.

AI Energy Consumption

Significant amounts of power are needed to create and use the most advanced AI models, such as GPT-4o, Llama 3, or DALL·E 3. AI models' energy needs are expected to grow as their capabilities become increasingly sophisticated.

The energy consumption profile of AI models can generally be divided into model creation ("training") and model use ("inference"). Well before an AI model is first used, significant energy is consumed in its training. The energy requirements for model training primarily depend on the model's size (number of parameters), the quantity of the training data, and the speed with which training must be completed.[5]

Today's simpler AI models have millions of parameters, while the most powerful have trillions. In the past several years, the most advanced models have followed empirical "scaling laws," in which a model's capabilities grow in proportion to the amount of training data and the number of model parameters. A natural consequence of scaling laws is that advancements in model performance will require ever-increasing amounts of computing power to train and use the model.


50 GWH USED TO TRAIN THE GPT-4 MODEL
ELECTRICITY TO POWER 4,800 HOMES / YEAR

The most advanced AI models, such as large language models (LLMs), must be trained with imposing amounts of data[6] while utilizing expensive computing hardware known as AI accelerators.[7]

AI accelerators are designed to rapidly perform the types of computations at the heart of modern AI and machine learning. Consequently, the training time of an AI model can be drastically reduced using hundreds or thousands of AI accelerators simultaneously, albeit with considerable investments in both hardware and energy.

The larger the model, the greater the number of accelerators required to train it in a reasonable amount of time. Training state-of-the-art AI models also requires high-speed networks to move large amounts of data among these hardware components. Operating such massive computing systems produces significant heat, which requires even more energy to power robust cooling systems.

[5] Additional determinants of AI model energy usage include the associated infrastructure overhead and data center efficiency.

[6] Meta's Llama 3 model was trained with 15 trillion tokens. https://ai.meta.com/blog/meta-llama-3/. This is approximately 60,000 gigabytes of text, assuming an average of four bytes per token.

[7] Some types of AI accelerators include graphics processing units (GPUs), neural processing unit (NPUs), and Tensor Processing Units (TPUs).

A report from Stanford University concluded that training the AI model GPT-3 likely required approximately 1.3 gigawatt-hours[8] of electricity.[9] In 2023, its successor model, GPT-4, required an estimated 50 GWh to train.[10]

For reference, the average U.S. household consumption is 10.5 megawatt-hours of electricity per year. In other words, the 50 GWh used to train the GPT-4 model is approximately the same amount of electricity as 4,800 homes would consume in a year.[11]

Each new generation of frontier AI model has been larger and trained on much more data than the previous generation. Consequently, each successive advancement in AI models has required significantly more energy to train than the previous.

Once a model is trained, it can be put to use, providing responses to users' requests. As in training, model inference can also require substantial investments in energy and specialized hardware. Energy requirements are even greater if the model is designed to achieve superior performance in any of several different categories.

In one such performance category, the model's outputs exhibit very high quality or accuracy. For example, a model could be designed to generate very realistic images, make accurate weather predictions, or precisely direct the movement of an autonomous vehicle.

Another performance category is the model's latency, i.e., the model's response time. In applications such as chatbots, low latency is valued because users expect to receive responses to their queries in real time, emulating the pace of a conversation with another person.

While each instance of the model inference process is far less energy intensive than training, in the aggregate of hundreds of thousands of simultaneous users and all outputs produced over a long period of time, inference by advanced AI models can far surpass the energy consumption of model training. For example, one estimate puts ChatGPT's operating consumption in 2023 at 564 MWh each day,[12] which means that every three days of use requires more power than the amount used in training GPT-3.

Opportunities to reduce the power consumption of model inference—such as utilizing hardware efficiencies and intelligently selecting models of appropriate size for the user's needs—will also be critical to broader policies related to AI energy usage.

[8] Energy is expressed in units of watt-hours (Wh) and power in watts (W). For example, a 40 W lightbulb run continuously for 24 hours takes 40 x 24 = 960 watt-hours of energy or almost 1 kilowatt-hour (kWh).

[9] Stanford Human-Centered AI Institute. AI Index 2024 Report. 2024, https://aiindex.stanford.edu/report/.

[10] Cohen, Ariel. "AI Is Pushing the World Toward an Energy Crisis." Forbes, 23 May 2024, https://www.forbes.com/sites/arielcohen/2024/05/23/ai-is-pushing-the-world-towards-an-energy-crisis/.

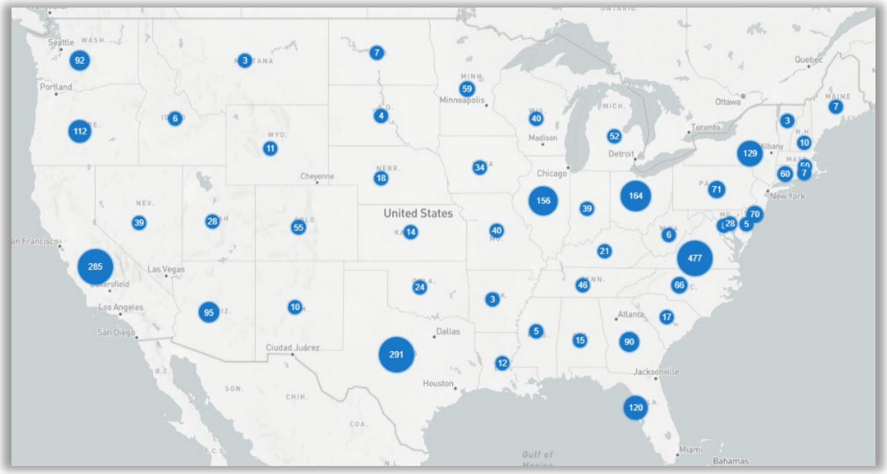
[11] U.S. Energy Information Administration. "Electricity Use in Homes." Energy Information Administration, https://www.eia.gov/energyexplained/use-of-energy/electricity-use-in-homes.php.

[12] De Vries, Alex. "The growing energy footprint of Artificial Intelligence." Joule, vol. 7, no. 10, October 2023, pp. 2191–2194, https://doi.org/10.1016/j.joule.2023.09.004.

An additional issue is that increased AI usage and the associated demand for AI chips heighten concerns about semiconductor supply chains, including critical mineral needs. Accordingly, AI's energy and natural resource requirements must account for the resources consumed in constructing new chips, data centers, power plants, and transmission infrastructure, as well as the downstream effects of AI-enabled efficiencies.

Data Centers

Data centers are physical facilities that house computing hardware, networking, and storage resources. They are composed of information technology (IT) equipment, such as servers and routers, as well as support infrastructure, including ventilation, cooling, and electrical power subsystems. Today, the data center sector is dominated by "hyperscalers" from the largest cloud computing companies. A sizeable data center can house several thousands of servers in a million-square-foot facility.



[Source: Data Center Map – Data Center Concentration](#)

The power demanded by data centers is often used as a proxy for the energy consumed by AI systems, which companies typically do not disclose to the public.

Other factors that contribute to energy consumption in data centers include customer demand, the energy efficiency of the data center, model complexity, the number of resource-intensive queries, and the rate of hardware replacement.

Google's self-reported energy demand for machine learning activities at its data centers has remained at or below 15% despite recent growth in both the use of AI models and the number of users.[13]

This consistency could be explained by increased efficiency in how Google manages its data center operations. Nevertheless, it is possible that eventually increased demand for AI-related services will outpace such efficiencies.

[13] Vida Rozite, et al. "Why AI and Energy Are the New Power Couple." Energy Information Administration, 2 Nov. 2024, https://www.iea.org/commentaries/why-ai-and-energy-are-the-new-power-couple.

It is also difficult to track the overall energy usage of data centers because of a lack of transparency into the industry and the absence of a distinct category for self-reporting via the U.S. Energy Information Administration's Commercial Building Energy Consumption Survey.[14]

Reports like Lawrence Berkeley National Laboratory's *United States Data Center Energy Usage Report* (2016) estimate consumption based on proxies and trends, which magnifies the uncertainty in understanding actual usage.[15]

Unlike traditional electrical loads, many data centers consume power in large quantities and consistently high levels throughout the year. Consequently, any new power generation sources to support the growth of data centers must deliver power in different ways than traditional sources.

In addition to electricity, data centers also require vast amounts of clean water, primarily for cooling hardware. In many cases, this water can be recirculated and returned to its source, but in some cases, it is not. The average data center has a water usage effectiveness of 1.8 L of water consumed (not recirculated) per 1 kWh of IT equipment energy. However, the hyperscale data centers can be an order of magnitude lower.[16]

The energy and water consumption of data centers is not spread uniformly throughout the United States; it is typically clustered around specific regions. For example, fifteen states accounted for an estimated 80% of the national data center load in 2023, with Virginia ranked the highest.[17]

Growth in AI Energy Use

The surge in generative AI and the proliferation of other large predictive AI models has caused a marked increase in estimated data center energy use since 2022. Projections for AI-related energy use vary widely. One report projects U.S. data center energy usage (excluding cryptocurrencies) to comprise between 4.6% to 9.1% of total U.S. energy use by 2030 (up from about 4% in 2024).[18]

[14] U.S. Energy Information Administration. "Commercial Sector Energy Consumption." Energy Information Administration, 2018, https://www.eia.gov/consumption/commercial/.
[15] Arman Shehabi, et al. "United States Data Center Energy Usage." Lawrence Berkeley National Laboratory, June 2016, https://eta.lbl.gov/publications/united-states-data-center-energy.
[16] Mary Zhang. "Data Center Water Usage." DGT Infra, https://dgtlinfra.com/data-center-water-usage/.
[17] EPRI, "Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption", Electric Power Research Institute, 2024, https://www.epri.com/research/products/000000003002028905.
[18] Id.

The nation's largest grid operator, PJM Interconnection, predicts net energy growth of 2.4% per year from 2024 through 2034, primarily due to data centers.[19] Load growth from data centers is expected in other states and regions, including Arizona, Georgia, and Ohio, among many others.[20] For example, American Electric Power is anticipating up to 15 GW of new demand in its footprint by the end of the decade.[21]

According to their environmental reports, the carbon emissions of technology companies are soaring—Google reports their emissions have increased 50% since 2019,[22] Microsoft reports a 30% increase since 2020,[23] and Meta reports a 66% increase since 2021.[24] And this is projected to continue increasing: these hyperscale companies have already announced plans to build several additional gigawatt-scale (up to 5 GW) data centers to meet the AI demand. [25, 26, 27]

This sudden energy demand has put a strain on power grid planners, and it is difficult to rapidly respond to the demand for power with new supply. New AI models are developed on a timescale of six months, and new data centers take one to two years to construct.

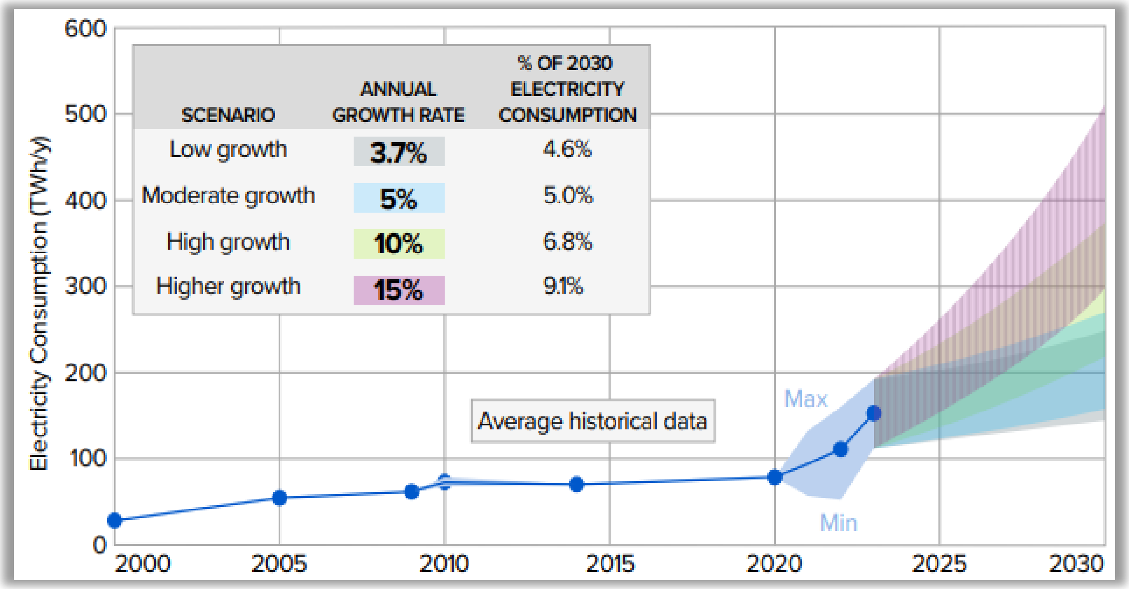
In contrast, new power plants take five to ten years, and new power transmission infrastructure takes 15 to 20 years. Proper planning now for new power generation and transmission is critical to laying the groundwork for AI adoption and innovation, as well as for achieving climate and emissions goals.

Projecting the future of AI-related energy usage is difficult because it depends on accurate predictions of AI development, adoption, and improvements. AI hardware has been continually improving in performance and energy efficiency. NVIDIA claims their current GPUs are 25x more energy efficient for the same performance as the previous generation.[28]

[19] PJM. "PJM Publishes 2024 Long-Term Load Forecast." PJM, 2024, https://insidelines.pjm.com/pjm-publishes-2024-long-term-load-forecast/.
[20] Darren Sweeney. "Rising Data Center Demand Forces Reckoning with U.S. Utility Decarbonization Goals." S&P Global Market Intelligence, 17 March 2024, https://www.spglobal.com/marketintelligence/en/news-insights/latest-news-headlines/rising-datacenter-demand-forces-reckoning-with-us-utility-decarbonization-goals-80889360.
[21] Ethan Howland. Utility Dive. "AEP faces 15 GW of new load, driven by Amazon, Google, other data centers: interim CEO Fowke." Utility Dive, 1 May 2024, https://www.utilitydive.com/news/aep-data-centers-amazon-google-load-growth-epa/714806/.
[22] Google. Google 2024 Environmental Report. 2024, https://sustainability.google/reports/google-2024-environmental-report/.
[23] Microsoft. Microsoft 2024 Sustainability Report. 2024, https://www.microsoft.com/en-us/corporate-responsibility/sustainability/report.
[24] Meta. Meta 2024 Sustainability Report. 2024, https://sustainability.atmeta.com/2024-sustainability-report/.
[25] Gordon, Cindy. "Microsoft and OpenAI Partnering on Stargate: A $100B U.S. Data Center." Forbes, 31 March 2024, https://www.forbes.com/sites/cindygordon/2024/03/31/microsoft-and-openai-partnering-on-stargate-a-100b-us-data-center/.
[26] Miller, Rich. "The Gigawatt Data Center Campus Is Coming." Data Center Frontier, 2024, https://www.datacenterfrontier.com/hyperscale/article/55021675/the-gigawatt-data-center-campus-is-coming.
[27] For reference, the average nuclear reactor generates 1 GW of power.
[28] Harris, Dion. "Blackwell Scientific Computing." NVIDIA Blog, 2024, https://blogs.nvidia.com/blog/blackwell-scientific-computing/.

While some developers pursue ever-larger general-purpose AI models, concurrent progress is also being made in advancing smaller, specialized AI models that require less power to train and use.

Data centers can report their power usage effectiveness (PUE) to describe their energy efficiency—calculated as the ratio of total power use to the power use by IT equipment. However, PUE does not capture the overall efficiencies of the models being run. There is a need for metrics, benchmarks, and vocabulary to describe if AI models use energy efficiently for their tasks. For example, it would be extremely energy-inefficient to use the power of a generative AI model to perform simple arithmetic that could be performed on a calculator.



SCENARIO	ANNUAL GROWTH RATE	% OF 2030 ELECTRICITY CONSUMPTION
Low growth	3.7%	4.6%
Moderate growth	5%	5.0%
High growth	10%	6.8%
Higher growth	15%	9.1%

Source: EPRI Powering Intelligence 2024 White Paper. Projections of potential electricity consumption by U.S. data centers: 2023-2030. % of 2030 electricity consumption projections assume that all other (non-data center) load increases at 1% annually

New Energy Generation for Data Centers

Coinciding with the growth in energy demand is a predicted increase in power generators being taken out of service. PJM and its market monitor are forecasting 24 to 58 GW of thermal generator retirements in the region by 2030 due largely to regulations, policies, and their collective effect on the economic viability of these resources.[29]

[29] Ethan Howland. "PJM Coal, Gas Power Plant Risk of Retirement, Market Monitor." Utility Dive, 18 March 2024, https://www.utilitydive.com/news/pjm-coal-gas-power-plant-risk-retirement-market-monitor/710518/.

Data centers are becoming operational at a much quicker pace than electric infrastructure, which can be planned, approved, and built. In addition, power transmission and distribution will require infrastructure investments to accommodate the increasingly concentrated loads of data centers.[30] This surge in demand for affordable, reliable, and dispatchable generation comes at a time when the North American Electric Reliability Corporation (NERC) has repeatedly raised concerns over the grid's adequacy and reliability due to a confluence of factors, such as state and federal policies that have forced retirements of reliable generation prior to the end of the facility's expected lifetime without adequate replacement generation resources and electric infrastructure.

While much new power generation consists of wind and solar, these resources are intermittent and thus lack two characteristics essential for data centers.[31] Data centers require sufficient amounts of consistently delivered power that can rapidly adjust to fluctuating demand. Without feasible energy storage options, data centers cannot operate solely on sources like wind or solar, so they are driving up demand for coal, natural gas generation, nuclear generation, and other consistent sources like geothermal generation. Any power source that does not deliver sufficient power with the ability to adjust power levels as needed could not be considered a suitable replacement.

Innovation in AI Requires Innovation in the Energy Sector

Today, data centers remain a smaller proportion of total energy consumption than many other sources of demand. Nevertheless, barriers to sufficient and appropriate availability of energy and related grid infrastructure could constrain AI access and innovation or lead data centers to increase their presence outside of the U.S.

One concern is that data centers require continuous power. There has been increased interest in co-locating data centers with nuclear power plants and developing small modular reactors. However, despite their promise, widespread deployment of new nuclear power is not a near-term solution due to the long lead times required to license and construct a first-of-a-kind nuclear power plant.

Continued innovation in data center design, including improved cooling systems and optimization as well as improved building construction, would help increase their energy efficiency. Likewise, innovation in energy systems would decrease the burdens of growing energy demands.

Similarly, advances in chip design, packaging, interconnects, memory, and new AI accelerator architectures could improve the energy efficiency of AI systems. Many in the industry have already begun implementing solutions to reduce energy consumption.

[30] Required infrastructure investments, such as upgrades to substations, distribution lines, and transmission lines, would need to enable energy transmission and distribution infrastructure to support new loads as they are connected to the grid.
[31] Advanced manufacturing and other high-tech activities have similar requirements.

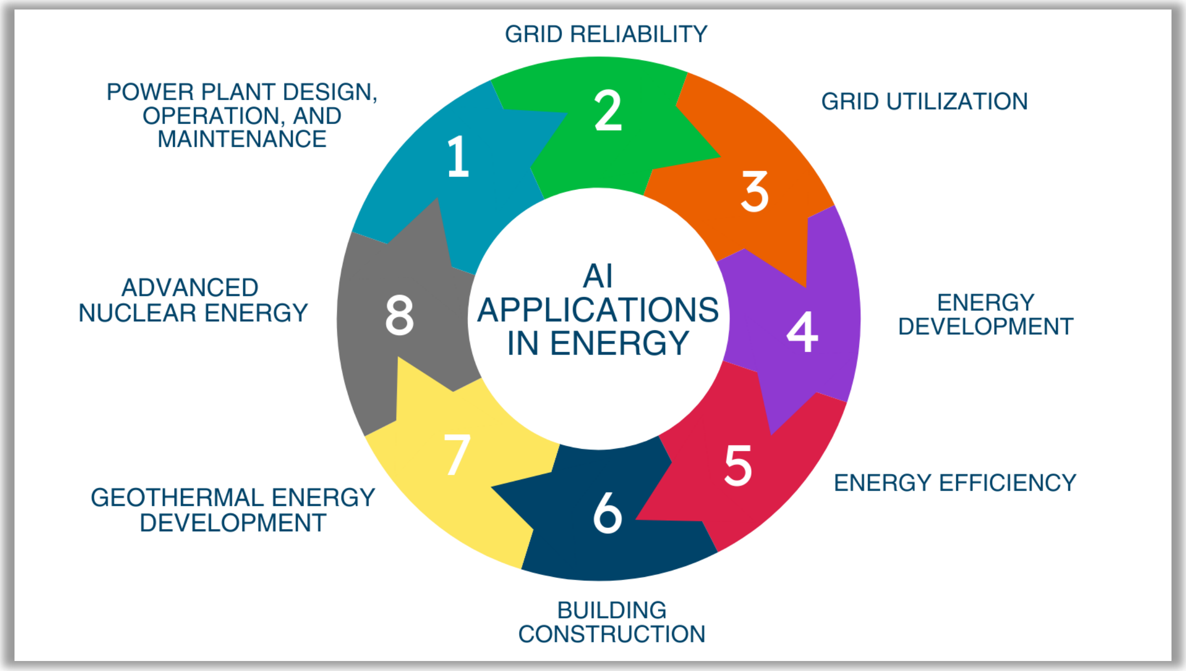
Relevant technologies include more efficient hardware and GPUs designed for AI, efficient semiconductor packaging, power capping during model training, early stopping for underperforming models, hardware-software co-optimization, improved cooling systems, model compression and pruning, more efficient model architectures, transfer learning, optimized data centers, and carbon-aware computing which trains models using grids powered by cleaner energy sources. As technology matures and research proceeds, many other solutions can be rapidly developed and implemented.

AI Enables Energy Development and Efficiency

AI will play a role in optimizing power plant design, operation and maintenance, grid reliability, and grid utilization. AI and machine learning are being progressively deployed across the energy sector to optimize energy supply and management.

Given the grid's size and complexity, it is often difficult for planners to ensure sufficient generation meets the varying load requirements needed. AI can help meet this challenge by improving coordination, optimizing infrastructure resources, unlocking underutilized assets, and organizing the enormous amount of technical data that is generated daily.[32]

With AI tools, system operators can achieve real-time operational awareness, enhancing decision-making processes and enabling rapid responses to dynamic conditions. AI can also be used to monitor the grid and protect it from potential vulnerabilities both in physical and cyber spaces.



U.S. Department of Energy. "AI for Energy: Opportunities for a Modern Grid and Clean Energy Economy." *U.S. Department of Energy,* April 2024, https://www.energy.gov/ceser/articles/doe-delivers-initial-risk-assessment-artificial-intelligence-critical-energy.

Improved modeling and simulation may facilitate more efficient resource utilization and infrastructure planning in the energy sector. Furthermore, AI's predictive maintenance capabilities and anomaly detection mechanisms can ensure energy systems' reliability and resiliency, mitigating the risk of failures and enhancing overall performance.[33]

AI is also being used to advance energy development and energy efficiency. AI technologies are increasing the energy efficiency of buildings, for example, by optimizing HVAC systems. It is also optimizing the building process to save energy and time and decrease waste.

Enabling these opportunities for AI to improve operations will be important to matching energy supply to increasing demand. AI tools are also used in finding new energy resources and in research for carbon capture, advanced nuclear energy, fusion energy systems, and energy storage technologies, including materials and design optimization for batteries and fuel cells.

For example, enhanced subsurface mapping can take the often sparse and patchwork subsurface data and provide a digital map so users can identify resources like hydrogen and geothermal energy. These tasks may be too difficult to calculate and perform without AI.

Economic Development

The growth of AI and data centers could bring economic development to both large urban areas and less-developed rural areas. Localities and states that embrace the expansion of this industry see economic benefits in the form of job creation, tax base expansion, and infrastructure development.

Economic growth from data centers is not exclusive to data centers and data center companies; benefits could spread throughout the local and regional economy. By one estimate, each job in the data center industry supports six jobs in the broader U.S. economy. For example, according to a report by the Data Center Coalition, in 2022 alone, the data center industry added over 560,000 direct jobs and supported 4.2 million total jobs across the United States.[34]

[33] U.S. Department of Energy, Office of Cybersecurity, Energy Security, and Emergency Response. "DOE Delivers Initial Risk Assessment of Artificial Intelligence for Critical Energy Infrastructure." U.S. Department of Energy, 2024, https://www.energy.gov/ceser/articles/doe-delivers-initial-risk-assessment-artificial-intelligence-critical-energy.

[34] PwC. "Economic, Environmental, and Social Impacts of Data Centers in the United States." September 2023, https://static1.squarespace.com/static/63a4849eab1c756a1d3e97b1/t/65048f92e74c956b68a419e4/1694797719030/Data+Center+Impact+Study+Executive+Summary.pdf.

This economic impact is most prevalent in Northern Virginia's Data Center Alley, which has the world's largest concentration of data centers and has been the primary source of economic growth and demand growth in that area for decades.[35, 36]

Additional Considerations

While meeting the significant energy demands of data centers is essential to economic growth and national security, it is also important to maintain affordability, reliability, and availability of electricity to customers. Protecting ratepayers from subsidizing these new large loads of technology companies should be a priority for utilities and grid operators.

Utilities should develop new rate structures to ensure that data centers pay for the necessary upgrades and electricity they demand. Risks of higher costs or electricity shortfalls are particularly pronounced in areas with heavily concentrated data center presence and growth.

> **“While meeting this greatly increased demand is essential to economic growth and national security, it is also important to maintain affordability, reliability, and availability of electricity to customers.**

Local and community stakeholders, including Tribal governments where applicable, would have insights into the full scope of these risks. Moreover, ensuring proper communication and planning by localities, data center companies, utilities, grid operators, and governments can expedite the construction of necessary infrastructure to support growing power loads.

[35] Mullin, John. "Data Centers, Big Growth, and Economic Impacts." Richmond Federal Reserve, Q2 2023, https://www.richmondfed.org/publications/research/econ_focus/2023/q2_feature2#:~:text=All%20of%20this%20has%20added,Data%20Center%20Alley%20each%20day.

[36] Loudoun County Economic Development. "Key Business Sectors: Data Centers." Loudoun County Economic Development, https://biz.loudoun.gov/key-business-sectors/data-centers/.

Key Findings

AI is critical to both U.S. economic interests and national security and maintaining a sufficiently robust power grid is a necessity.

Plentiful, consistent sources of power across the nation will enable the use and expansion of AI technologies.

The growing demands of AI are creating challenges for the grid.

The growing integration of AI technologies throughout everyday life is rapidly increasing energy demand, outpacing the addition of new power capacity. This increased energy demand from data centers can cause supply constraints and raise energy prices, creating challenges for electrical grid reliability and affordable electricity.

Continued U.S. innovation in AI requires innovations in the energy sector.

Industry innovation in AI technology produces new approaches to reducing power consumption while preserving performance. Resource efficiency improvements in data center design and energy systems could decrease the burdens of growing energy demands.

Planning properly now for new power generation and transmission is critical for AI innovation and adoption.

While new data centers take one to two years to construct, new power plants take five to ten years, and new power transmission infrastructure takes fifteen to twenty years.

AI tools will play a role in innovation and modernization in the energy sector.

AI will play a role in optimizing power plant design, operation and maintenance, grid reliability, and grid utilization. AI and machine learning are being progressively deployed across the energy sector to optimize energy supply and management.

Recommendations

Recommendation: Support and increase federal investments in scientific research that enables innovations in AI hardware, algorithmic efficiency, energy technology development, and energy infrastructure.

There is a strong impetus for AI-enabled innovation and efficiencies throughout relevant AI technologies: energy infrastructure, AI chips, algorithms, and data centers. Federally funded research and development are instrumental in driving these advancements.

Federal programs should authorize and support these research activities at the Department of Energy, the Department of Commerce, and the National Science Foundation. The government should also support public-private partnerships, both to understand the variety of industry needs and to use AI appropriately to accelerate research and development efforts. Existing interagency programs can coordinate research efforts across the federal science and technology enterprise. Previous legislative efforts can also be utilized to further research, including the National AI Initiative Act, the Energy Act of 2020, and the CHIPS and Science Act.

Recommendation: Strengthen efforts to track and project AI data center power usage.

Data center owners and operators should voluntarily report energy usage and projections to centralized bodies such as the Energy Information Administration or among relevant entities like hyperscale data centers and utilities, grid operators, and planning commissions. Collaboration among these entities will improve the accuracy of near-term demand forecasts and prevent double counting. More confident medium- and long-term projections about energy demands would likely require the cooperation of data center operators, cloud computing companies, AI hardware and software companies, researchers, and others. Transparency legislation could support these efforts.

Recommendation: Create new standards, metrics, and a taxonomy of definitions for communicating relevant energy use and efficiency metrics.

This could include creating or designating a public AI data center testbed for benchmarking energy usage with standardized hardware and software. This would allow for easier and more transparent comparisons between models or hardware components. These efforts should incentivize partnerships with the relevant federal agencies, academia, civil society, and industry stakeholders.

Recommendation: Ensure that AI and the energy grid are a part of broader discussions about grid modernization and security.

The U.S. must deal with its growing energy demands in the face of aging infrastructure and increased electrification nationwide.

Recommendation: Ensure that the costs of new infrastructure are borne primarily by those customers who receive the associated benefits.

This approach prevents an unfair allocation of costs away from technology companies and onto residential ratepayers and other customers with limited alternatives.

Recommendation: Promote broader adoption of AI to enhance energy infrastructure, energy production, and energy efficiency.

The creation and maintenance of our energy infrastructure can be enhanced with AI. AI can lower energy costs and improve energy availability by improving numerous facets of the national energy ecosystem: infrastructure, production, and efficiency. AI systems can improve the reliability and utilization of the power grid and protect it against physical and cyberattacks. Similarly, new energy development technologies can be unlocked by using AI to locate and extract resources such as oil, gas, and hydrogen. AI will also be critical in advancing fundamental research and engineering for energy generation and storage.

SMALL BUSINESS

Background

Small businesses are the backbone of the United States economy, representing 43.5% of U.S. GDP and employing 45.9% of the workforce.[1] Mature and accessible AI technologies have the potential to improve small businesses' efficiency, bandwidth, and competitiveness, allowing them to handle their work more quickly and effectively. Furthermore, small businesses play a crucial role in maintaining the United States' lead in the AI race against other world powers. Unfortunately, small businesses often lack the understanding or resources that would allow them to meaningfully adopt this critical technology.

As discussed throughout this report, AI is not a new technology. Industry, government, and academia have used some form of automation for decades. The Chamber of Commerce discussed AI use by small businesses in a recent report titled *"The Impact of Technology on U.S. Small Business."*[2]

> **"Small businesses play a crucial role in maintaining the United States' lead in the AI race against other world powers.**

[1] Ferguson, Stephanie, et al. "See the Data behind America's Small Businesses." See the Data behind America's Small Businesses. | U.S. Chamber of Commerce, 5 Sept. 2024, https://www.uschamber.com/small-business/small-business-data-center.

[2] "The Impact of Technology on U.S. Small Business." The Impact of Technology on U.S. Small Business | U.S. Chamber of Commerce 15 Sept. 2024, www.uschamber.com/technology/artificial-intelligence/the-impact-of-technology-on-u-s-small-business.

This year's survey found that AI use among small businesses is nearly universal, as 98% use legacy AI-enabled tools ranging from spam blockers to spell check to virtual assistants.

When explicitly asked about newer AI tools like generative AI, the Chamber found that 40% of companies currently identify as using generative tools, up from 23% in the previous year's survey. AI is advancing rapidly, and it is critical that Congress consider the unique challenges and needs of small businesses, including the establishment of AI guardrails that protect small businesses and ensure they have a seat at the table for federal action affecting them.

AI Literacy and Adoption

Many small businesses have expressed interest in adopting new AI tools. At a House Small Business Committee's March 2024 hearing titled *"Conducting Oversight: Testimony from the Small Business Administrator,"* Small Business Administrator Isabella Casillas Guzman testified that small business demand for AI tools and related technical assistance is "quite extensive […] as they develop their marketing, or perhaps [use them] as chatbots in their customer service."[3] AI tools can also help with data analytics and other business operations.[4]

70%
OF SMALL BUSINESSES VIEW AI POSITIVELY

However, entrepreneurs' AI literacy has not yet caught up with their interest. According to a 2023 survey of small business owners and executives conducted by the Bipartisan Policy Center and Morning Consult, while over 70% of respondents view AI positively, some of the top barriers to AI adoption include not knowing what tools to use and employees' lack of digital skills.[5] Currently, AI literacy resources are not uniform or widespread among communities.

This inequality may exacerbate the digital divide between privileged and underserved communities, including small businesses, as the former adopts AI tools while the latter remains unaware of their existence.

[3] Congressional Testimony Prepared for SBA Administrator Isabella Guzman U.S. House Committee on Small Business. 18 March 2024, https://democrats-smallbusiness.house.gov/uploadedfiles/03-20-24_administrator_guzman_testimony.pdf.

[4] Inston, Kyra, et al. "Three Ways AI Is Transforming Small Businesses." 3 Ways AI Is Transforming Small Businesses | Bipartisan Policy Center, Bipartisan Policy Center, 9 May 2024, https://bipartisanpolicy.org/blog/3-ways-ai-is-transforming-small-businesses/.

[5] Garapati, Sujan. "Poll Shows Small Businesses Are Interested in and Benefit from AI." Poll Shows Small Businesses Are Interested in and Benefit from AI | Bipartisan Policy Center, Bipartisan Policy Center, 18 March 2024, https://bipartisanpolicy.org/blog/poll-shows-small-businesses-are-interested-in-and-benefit-from-ai/.

The rapid advancement of this technology also intrinsically advantages better-resourced companies with the means and technical expertise to understand the AI landscape and adopt new tools quickly and appropriately. Most of those are large corporations. The lack of AI literacy has resulted in high-profile harmful and embarrassing incidents for businesses of all sizes.[6] These incidents could have been avoided by a thorough and critical review of the underlying AI technology involved and the relevant flaws, liabilities, and drawbacks.

The reputational and legal harm caused by flaws in AI tools deployed by businesses can be steep, negatively, and significantly impacting small businesses. This underscores the need for entrepreneurs considering adopting AI tools to critically evaluate and assess their capabilities. If they determine these tools are necessary, entrepreneurs must be equipped to adopt them effectively and to remedy any related drawbacks. Strong AI literacy promotion policies, carried out in cooperation with the Small Business Administration (SBA), would ensure that small businesses have access to the knowledge and resources needed to survive and thrive in the ongoing AI revolution.

Market Concentration

While small firms contribute greatly to AI innovation, as they have with previous technological advancements, the market around AI development is concerningly concentrated around a few large companies.[7] Large amounts of capital and computer processing power are needed to build, train, and operate AI models, posing significant barriers to entry for small AI startups to innovate, compete with big tech companies, and add to our nation's competitiveness in AI development.[8]

Some AI companies release their technology as open-source, which allows smaller companies to develop on top of it to improve capabilities, increase use cases, and discover and mitigate flaws. This collaboration between large and small AI companies is important for a diverse and robust AI ecosystem. Policies encouraging this collaboration, such as those encouraging open-source AI development, could create more opportunities for individuals and small businesses.

[6] Notopoulos, Katie. "A Car Dealership Added an AI Chatbot to Its Site. Then All Hell Broke Loose." Business Insider, 18 Dec. 2023, https://www.businessinsider.com/car-dealership-chevrolet-chatbot-chatgpt-pranks-chevy-2023-12.; see also: Melnick, Kyle. "Air Canada Chatbot Promised a Discount. Now the Airline Has to Pay It." The Washington Post, 18 Feb. 2024, https://www.washingtonpost.com/travel/2024/02/18/air-canada-airline-chatbot-ruling.; see also: Lecher, Colin, et al. "Malfunctioning NYC AI Chatbot Still Active Despite Widespread Evidence It's Encouraging Illegal Behavior." THE CITY, 2 April 2024, https://www.thecity.nyc/2024/04/02/malfunctioning-nyc-ai-chatbot-still-active-false-information/.

[7] Vipra, Jai, and Anton Korinek. "Market Concentration Implications of Foundation Models: The Invisible Hand of ChatGPT." Brookings, 7 Sept. 2023, https://www.brookings.edu/articles/market-concentration-implications-of-foundation-models-the-invisible-hand-of-chatgpt/.

[8] Liu, Henry, et al. "Generative AI Raises Competition Concerns." Federal Trade Commission, 29 June 2023, https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/06/generative-ai-raises-competition-concerns.; see also: Vipra, Jai, and Sarah Myers West. "Computational Power and AI." AI Now Institute, 27 Sept. 2023, https://ainowinstitute.org/publication/policy/compute-and-ai.

Small companies face steep challenges in accessing the capital needed to be competitive, such as from microloans, commercial lending, government grants, and private venture capital financing. This is an especially salient challenge for entrepreneurs from populations underrepresented in the technology sector, including rural populations, women, and people of color.[9]

The SBA plays a crucial role in connecting small businesses with capital by administering loans and disaster assistance; overseeing the federal Small Business Innovation and Research (SBIR), Small Business Technology Transfer (STTR), and Growth Accelerator Fund Competition (GAFC) programs; and licensing Small Business Investment Companies.

Other federal agencies support small businesses in their individual domains. For example, the National Institute of Standards and Technology (NIST) supports the Manufacturing Extension Partnership (MEP), which helps small- and medium-sized manufacturers adopt new technologies, including artificial intelligence.[10] Other public resources like the National AI Research Resource (NAIRR) could provide small businesses and start-ups with the data and computing resources necessary to compete with larger, better-resourced companies.

In addition, policymakers should protect competition in markets where AI is rapidly being adopted. Large corporations' rapid and unfettered deployment of AI tools could hurt small businesses' competitiveness across sectors, including outside the tech industry. For example, big tech firms are developing generative AI tools that produce art, music, and writing at lower prices than small content creators can compete with.

The House Committee on Small Business examined this problem in a March 2024 report.[11] As discussed in the chapter on **Intellectual Property**, there are ongoing lawsuits in this area. However, small businesses do not have the time or resources to engage in lengthy litigation without clear legal precedence. As a result, the absence of clear guidance for intellectual property creates more challenges for small businesses.

Exacerbating Existing Small Business Challenges

While AI tools could make small businesses more nimble, agile, and competitive with bigger, more well-resourced firms, they could also unintentionally or intentionally amplify existing obstacles, such as when corporations use AI to provide services to small businesses.

[9] "Diversity in High Tech." U.S. Equal Employment Opportunity Commission, 2016, https://www.eeoc.gov/special-report/diversity-high-tech.
[10] "Manufacturing Extension Partnership (MEP)." National Institute of Standards & Technology, 29 Sept. 2014, https://www.nist.gov/mep.
[11] Prepared by the House Committee on Small Business Democratic Staff. Bots Over Brushes: The Looming Competition Between Generative AI and Small Content Creators, 19 March 2024, democrats-smallbusiness.house.gov/uploadedfiles/generative_ai_report_final.pdf.

For instance, improperly tested and implemented AI tools used by financial institutions can be prone to discriminatory behavior, hindering underserved small businesses seeking loans, real estate leases, and online targeted advertising services. Additionally, AI can be used by large corporations to engage in widespread price discrimination against potential small business clients.

The use of AI by all levels of government will also affect the ease with which small businesses can access public services and assistance, as well as the government's responsiveness to their needs. Thoroughly vetted and proven AI tools could greatly speed responses to small business inquiries.[12] However, faulty and unreliable AI tools can hurt small businesses more than they help, sowing confusion and uncertainty for America's small businesses.[13]

In December 2023, the House Committee on Small Business shared these concerns with the White House Office of Management and Budget and the Commerce Department, encouraging them to include strong guardrails in deploying future federal AI tools for small businesses.[14]

Compliance Burdens

Small businesses may be disproportionately affected by federal government compliance requirements. This is why the Task Force sought input from industry and companies of all sizes to identify how to navigate potential future regulatory challenges properly. As with most small business regulatory matters, compliance efforts are time-consuming and often unfairly costly for businesses that may lack the ability to redirect resources from core business activities. Meanwhile, large firms have the infrastructure to handle unclear or lengthy compliance requirements for new and existing government regulations.

HALF OF RESPONDENTS CITED "UNCERTAINTY OVER POSSIBLE GOVERNMENT REGULATION" AS A BARRIER TO AI USE.

For example, a recent Department of Commerce's Bureau of Industry and Security (BIS) proposal aims to impose comprehensive reporting requirements on AI developers and cloud providers to the federal government.[15]

[12] "IRS Expands Use of Chatbots to Help Answer Questions on Key Notices; Expands on Technology That's Served 13 Million Taxpayers." Internal Revenue Service, 26 Sept. 2023, https://www.irs.gov/newsroom/irs-expands-use-of-chatbots-to-help-answer-questions-on-key-notices-expands-on-technology-thats-served-13-million-taxpayers.

[13] Supra 6.

[14] Shalanda, Hon, et al. "); to Advance the Safe and Responsible Use of Artificial Intelligence." Fed. Reg, vol. 75, no. 14, The White House, 2023, p. 191, democrats-smallbusiness.house.gov/uploadedfiles/231213_sbc_ds_sb_ai_chatbot_letter_closed.pdf.

[15] "Commerce Proposes Reporting Requirements for Frontier AI Developers and Compute Providers." Commerce Proposes Reporting Requirements for Frontier AI Developers and Compute Providers | Bureau of Industry and Security, 9 Sept. 2024, https://www.bis.gov/press-release/commerce-proposes-reporting-requirements-frontier-ai-developers-and-compute-providers.

While large companies have entire departments dedicated to maneuvering federal, state, and local challenges, smaller companies lose out on potential innovation and must reallocate time and resources to maintain compliance. Discussions about AI regulation are advancing quickly at the federal, state, and local levels, and worries about compliance burden are pervasive among small businesses.

The Chamber of Commerce survey found that almost a third of respondents cited “staying informed about new compliance requirements” as a key challenge for their business when it comes to AI regulation,[16] and the Bipartisan Policy Center poll found that half of respondents cited “uncertainty over possible government regulation” as a barrier to AI use.[17] This emphasizes the need to consider approaches that are not “one-size-fits-all” but tailored specifically to the type or size of audience a ruling would affect.

[16] U.S. Chamber Staff. “The Impact of Technology on U.S. Small Business.” U.S. Chamber of Commerce, 15 Sept. 2024, https://www.uschamber.com/technology/artificial-intelligence/the-impact-of-technology-on-u-s-small-business.
[17] Supra 5.

Key Findings

Small businesses often lack a full understanding of how best to adopt AI.

While small business owners are largely enthusiastic about AI, it can be challenging to understand AI enough to select appropriate AI products for tasks and avoid the pitfalls of AI use.

Small businesses can lack sufficient access to capital and AI resources.

Access to computational power and large data sets is increasingly a barrier to entry for small businesses attempting to fully utilize AI and compete with larger firms. These resources are expensive, requiring small businesses to overcome hurdles in acquiring access to sufficient capital.

Small businesses face excessive challenges in meeting AI regulatory compliance.

Any state or federal AI regulation would tend to disproportionately affect small businesses compared to large firms with greater resources. This includes the uncertainty that firms face from the possibility of future AI regulations at the federal or state level.

Recommendations

Recommendation: Support small business AI literacy.

AI literacy is necessary to adopt AI to improve a business's operations and effectiveness. Making AI training and technical assistance available to small business owners and employees would help businesses adopt AI tools. This would help new businesses start, and existing businesses enhance their productivity and more effectively serve their customers.

Recommendation: Provide resources for small business AI adoption.

Providing small business owners, including small manufacturers, with resources such as compute power and AI-ready data sets would facilitate greater AI adoption. In partnership with non-governmental entities, agencies like the Small Business Administration and NIST play a role in creating, identifying, and disseminating resources for small business owners and employees to better understand and use emerging technologies. Also, access to capital and new financing methods to address resource challenges should be explored. Additionally, the NAIRR pilot can facilitate AI adoption by small businesses by providing difficult-to-acquire data and computing resources.

Recommendation: Investigate the resource challenges of small businesses adopting AI.

A full understanding of the capital and resource challenges small businesses face when adopting AI would inform policies to support small businesses. An investigation should be conducted to specifically identify the various challenges small businesses face on the path to AI adoption, whether some challenges are industry-specific, and how specific types of business operations affect resource challenges.

Recommendation: Investigate the resource challenges of small AI businesses.

AI and small businesses are critical to the U.S. economy and national interests. Therefore, it would be advantageous to promote more startups and small businesses that create, customize, or deploy AI products and services. An investigation should be conducted into the resource challenges that make starting and running small AI businesses difficult. New financing methods to address these challenges would also better inform these findings.

Recommendation: Ease compliance burdens for small businesses.

Any relevant legislation or regulation should consider the disproportionate compliance burden on small companies and how it may unfairly reduce competition. Any AI legislation should be clear and sector-specific and could, where feasible, provide appropriate technical assistance to small businesses.

AGRICULTURE

Background

Agriculture has long served as the backbone of economies across the world, providing resources that support societal needs and drive economic advancement. By some estimates, the worldwide population will increase from 8 billion today to almost 10 billion by 2050.[1] As the global population continues to rise, the demand for food, fiber, and other agricultural products is increasing rapidly. These demands impose considerable pressure on the agricultural sector to increase productivity, improve resource management, and ensure sustainability.

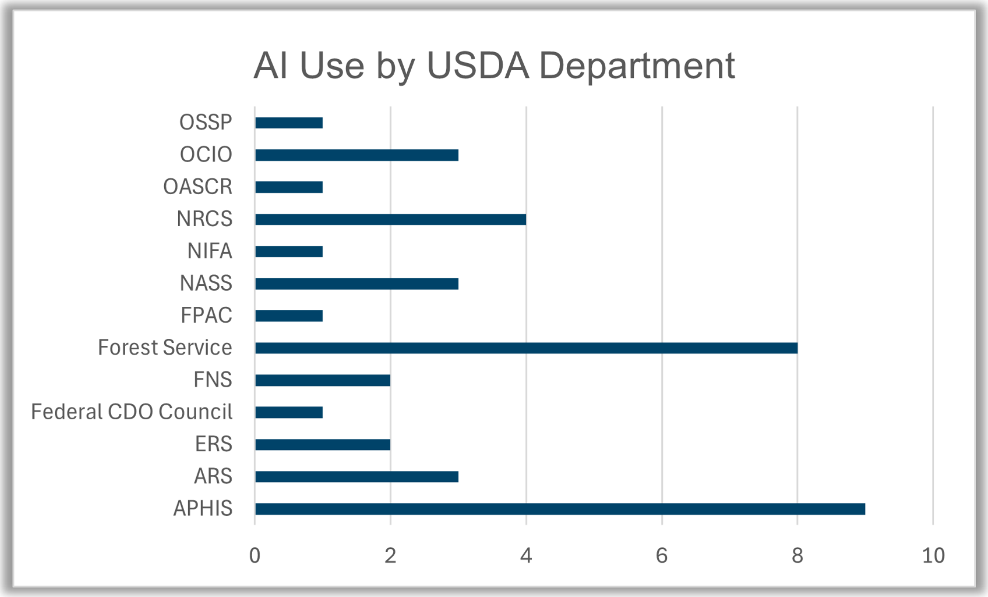
Artificial intelligence has emerged as a powerful tool capable of revolutionizing agriculture. AI can improve the processes that form the modern agricultural sector, from crop management to supply chain logistics, leading to greater aggregate efficiency and productivity. AI advancements have the potential to increase food availability, lower food prices, and bolster economic growth.

Several initiatives within the United States Department of Agriculture (USDA) have been launched to apply AI to agriculture. One prominent example is the partnership between the USDA National Institute of Food and Agriculture and the National Science Foundation (NSF) in 2020 and 2021.[2] This partnership led to the creation of five Artificial Intelligence Research Institutes at land-grant universities, thereby facilitating research on potential uses of AI to increase the efficiency and productivity of American agriculture.

[1] "9.7 Billion on Earth by 2050, but Growth Rate Slowing, Says New UN Population Report." United Nations News, 17 June 2019, https://news.un.org/en/story/2019/06/1040621.

[2] "USDA-NIFA and NSF invest $220M in Artificial Intelligence Research Institutes." USDA National Institute of Food and Agriculture, 29 July 2021, https://content.govdelivery.com/accounts/USDANIFA/bulletins/2ea829b.

Similarly, Executive Order 13960, issued by the Trump administration on December 8, 2020, mandated that federal agencies inventory and share their AI use cases with the public.[3] In compliance with this Executive Order, USDA compiled an inventory of its current and planned AI use cases consistent with its mission.[4] Some examples include using AI to strengthen crop estimates, process geospatial data, and model potential disease outbreaks.



Source: USDA - Inventory of USDA Artificial Intelligence Use Cases, May 2023

The 2023 appointment of USDA's first Chief Artificial Intelligence Officer signified a commitment to further AI development and utilization. USDA is now preparing its systems and workforce to harness AI's potential and is creating innovation incubators where USDA staff can safely test and evaluate AI technologies in a controlled environment.[5]

AI Applications for Agriculture Conservation and Natural Resources

USDA's AI inventory includes four use cases for its Natural Resources Conservation Service (NRCS).[6]

- Operational water supply forecasting for western U.S. rivers
- Ecological Site Descriptions (machine learning)
- Conservation Effects Assessment Project
- Digital Imagery (no-change) for the Natural Resources Inventory program

[3] "Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government." Federal Register, vol. 85, no. 236, 8 Dec. 2020, pp. 78939-78943. Federal Register, https://www.federalregister.gov/documents/2020/12/08/2020-27065/promoting-the-use-of-trustworthy-artificial-intelligence-in-the-federal-government.
[4] "Inventory of USDA Artificial Intelligence Use Cases." USDA, 26 July 2022, https://www.usda.gov/data/AI_Inventory.
[5] "Intelligent Automation Center of Excellence." USDA, 15 Aug. 2020, https://www.usda.gov/iacoe.
[S] Supra 4.

In addition, there are three broad areas of agricultural technology relevant to agricultural conservation:

- *Precision Agriculture:* Precision agriculture encompasses a range of technologies designed to enhance efficiencies in agricultural operations. Some prominent examples include remote sensing platforms, in-ground sensors, targeted spray systems, and automated mechanical weeders. These technologies have been incorporated into farming operations for decades, helping reduce input requirements, improve soil and water health, reduce operating costs for producers, and attain higher crop yields. As a result, the agriculture industry has become one of the most productive and efficient in the world, producing nearly 200% more food than in the 1940s, with little to no change in inputs.[7]

  Further efficiency gains may be achieved by more prominently utilizing AI technology and making AI tools more accessible and affordable to the agricultural sector. With a growing worldwide population, advances in agricultural technology, including AI, will help meet demand.

- *Water Technologies:* Water scarcity and drought remain urgent issues for producers, particularly in the Western United States. In 2021, the NRCS announced a new computer application that uses AI to forecast water supply in the West.[8] Known as the multi-model machine learning metasystem, or "M4,"[9] NRCS describes the tool as a first-of-its-kind and the largest migration of AI into river prediction programs. The application's central goal is to help producers be more efficient with water and protect the environment.

- *Land Change Analysis:* The National Agriculture Imagery Program (NAIP) collects aerial images during the agricultural growing season at a resolution of 60cm. The land change analysis tool *(LCAT*) analyzes data from the NAIP to provide high-resolution land cover maps. It is used by USDA to monitor agricultural activities, implement farm bill conservation programs, aid soil mapping and ecological investigation efforts, monitor wildlife habitat, assess disaster response, and monitor conservation easements.[10]

[7] Agricultural Productivity in the United States." U.S. Department of Agriculture, Economic Research Service, 2024, https://www.ers.usda.gov/data-products/agricultural-productivity-in-the-united-states/agricultural-productivity-in-the-united-states/.

[8] "New River Forecast Model Integrates Artificial Intelligence for Better Water Management in the West." U.S. Department of Agriculture, 4 November 2021, https://www.farmers.gov/blog/new-river-forecast-model-integrates-artificial-intelligence-better-water-management-in-west.

[9] National Water and Climate Center, M4. GitHub, 2023, https://github.com/nrcs-nwcc/M4

[10] Land Use Land Cover LCAT Metadata Map Server, USDA, 18 Nov. 2021, https://nrcsgeoservices.sc.egov.usda.gov/arcgis/rest/services/land_use_land_cover/lcat_metadata/MapServer.

> **“One persistent barrier to AI adoption is the lack of reliable network connectivity in rural and farming communities.**

The widespread adoption of advanced agricultural technology remains limited. According to the Government Accountability Office (GAO), although precision agriculture technologies have been available since the 1990s, only 27% of U.S. farms or ranches utilized such technology.[11] These technologies can be complex and have high up-front costs, making it hard for some farmers to access them.[12]

One persistent barrier to AI adoption is the lack of reliable network connectivity in rural and farming communities. Since many advanced technologies, including AI, require connections to servers or cloud providers, widespread adoption of state-of-the-art precision agriculture will be hampered until network access across agricultural communities is bolstered.

In addition, to fully leverage AI and other precision agriculture technologies, it is essential to address significant cost barriers to wider adoption. Investing in research, development, and innovation would reduce costs, enhance the commercial availability of technology, and facilitate new applications of these technologies.

Congress should continue to evaluate how precision agriculture, including AI-driven technologies, can enhance farm productivity and natural resource management. The House Committee on Agriculture should continue to address these barriers and support the deeper integration of precision technologies into production agriculture. For example, Farm Bill programs such as the Environmental Quality Incentive Program, the Conservation Stewardship Program, and the Business & Industry Loan Guarantee Program all include initiatives that promote precision agriculture.

AI Applications for Farm Production and Conservation Program Delivery

Following organizational reforms in the 1990s and 2018, three agencies that deal with agricultural producers—the Farm Service Agency (FSA), NRCS, and the Risk Management Agency (RMA)—are now organized under USDA’s Farm Production and Conservation (FPAC) mission area. All three agencies interact directly with producers to deliver assistance in various forms. However, because they lack a unified approach, producers face needless burdens in dealing with these agencies.

[11] U.S. Government Accountability Office. “Precision Agriculture: Benefits and Challenges for Technology Adoption and Use.” U.S. Government Accountability Office, 31 Jan. 2024, https://www.gao.gov/products/gao-24-105962.
[12] Id.

For example, all three agencies utilize their own differing definitions of field and land boundaries. In addition, all three continue to use various legacy information technology (IT) systems. One consequence of this approach is that producers are required to repeatedly provide the same information in applications to different programs across the same USDA mission area.

There is also an analogous burden on USDA employees, who often must manually enter data from a producer's applications for different programs within the same agency. Not only is redundant data entry inefficient, but it also increases the risk of introducing errors and inconsistencies that could void participation in certain programs.

There have been efforts at IT modernization intended to harmonize application processes and reduce duplicative information gathering from producers. However, these efforts have had limited success due to the logistical, financial, and technical challenges of migrating entire systems.[13]

Previous IT modernization efforts have attempted to migrate all FPAC customers (i.e., farmers) to a common system that utilizes similar land boundary definitions and harmonizes data collection across programs. These migration efforts have all fallen short. However, appropriately utilizing AI could circumvent these challenges by eliminating the need to migrate data.

USDA could use AI systems to better understand what data has already been collected from producers and reuse it whenever needed. For example, as producers prepare new applications, AI-enabled IT systems could automatically insert data that the producer has already provided. Unlike previous modernization attempts, an AI approach would allow existing IT systems to remain operational within their respective agencies.

Beyond enhancing program delivery, AI could help identify potential instances of waste, fraud, and abuse. One notable measure would be to use AI to identify discrepancies in data submitted by the same producer across various programs.

For example, a producer can file the same information on crop yields with both the FSA (by field) and with crop insurance through the RMA (by insurance unit). Inaccurate yield information could result in the payment of an excessive crop insurance indemnity to the producer. An AI tool that detects inconsistent crop yield information submitted over numerous years to FSA and RMA could alert RMA compliance staff to further review the matter.

[13] U.S. Government Accountability Office. "Farm Program Modernization: Farm Service Agency Needs to Demonstrate the Capacity to Manage IT Initiatives." U.S. Government Accountability Office, 18 June 2015, https://www.gao.gov/products/gao-15-506.

AI Applications for Wildfires and Forest Health

The United States continues to face a wildfire and forest health crisis on hundreds of millions of acres of private and federal lands. In January 2022, the United States Forest Service (USFS) announced a 10-year strategy to address the wildfire crisis in those areas most at risk. A prominent component of this strategy is the application of wildfire treatments to 20 million acres of National Forest System lands and 30 million acres of other federal, state, tribal, and private lands.

AI technologies are being deployed to detect wildfires earlier and manage areas affected by wildfires. AI is also providing new insights into trends in forest health and growth. Increased deployment of AI technology could decrease the threat of catastrophic wildfires, expedite responses to protect lives and property during wildfires, and improve overall forest health. Recognizing AI's potential in this area, eight AI use cases for the USFS are listed in USDA's inventory:[14]

- Ecosystem Management Decision Support System
- Wildland Urban Interface - Mapping Wildfire Loss
- Cross-Laminated Timber (CLT) Knowledge Database
- RMRS Raster Utility
- TreeMap 2016
- Landscape Change Monitoring System
- Geospatial and Remote Sensing Training Courses
- Forest Health Detection Monitoring

Forestry-Related AI Technologies

Many applications of AI can monitor and detect active wildfires, monitor trends in forest health and landscapes, and generally improve forest markets. AI broadly contributes to improving wildfire suppression by predicting wildfire events, improving the rapid detection of wildfires in real time, and guiding efforts to fight wildfires.

For example, the Forest Health Monitoring (FHM) program is a national program that assesses the status, trends, and future of forest conditions on an annual basis.[15] The program utilizes remote sensing, detection tools, aerial and ground surveys, and other technologies to collect diverse types of information on forest conditions. Using AI to analyze these data sources, USFS can make more informed decisions and interpret trends influencing forests.

[14] Supra 4
[15] "Forest Health Monitoring." U.S. Forest Service, 30 July 2022, https://www.fs.usda.gov/foresthealth/protecting-forest/forest-health-monitoring/.

Moreover, some states utilize AI to monitor and detect wildfires. The California Department of Forestry and Fire Protection (CAL FIRE) launched its AI fire detection tool in the summer of 2023.[16] This tool analyzes imagery data from over 1,000 remote cameras to monitor and identify potential fire incidents.

The technology also provides real-time data to first responders and resource providers. In some cases, firefighters are given both the estimated location of potential wildfires and the system's confidence in its wildfire prediction.

CAL FIRE uses AI to monitor and detect wildfires.

As a final example, the National Oceanic and Atmospheric Administration (NOAA) utilizes satellites equipped with advanced technology to provide real-time data for fighting wildfires.

The agency collaborates with the National Aeronautics and Space Administration (NASA) and the USFS for fire monitoring and information sharing.[17]

AI is already a powerful tool in addressing and combating the wildfire and forest health crises. Congress should explore whether further research, development, and technological innovation in technology could improve forest health and help land managers develop appropriate planning and strategies.

Additionally, policymakers should investigate whether AI can play a more prominent role in post-fire analysis and recovery promises to improve prospective measures, such as restoring land after wildfires and mitigating the damage from future fires and natural disasters.

[16] Baker, Elizabeth. "University of California San Diego's AI Fire Detection Tool Receives CENIC Innovation Award." Meteorological Technology International, 21 March 2024, https://www.meteorologicaltechnologyinternational.com/news/extreme-weather/university-of-california-san-diegos-ai-fire-detection-tool-receives-cenic-innovation-award.html.

[17] "NOAA Satellites Monitor Wildfires." National Environmental Satellite, Data, and Information Service (NESDIS), National Oceanic and Atmospheric Administration, 1 Aug. 2024, https://www.nesdis.noaa.gov/news/noaa-satellites-monitor-wildfires.

AI Applications for Specialty Crop Mechanization and Automation

Specialty crops—ranging from fruits and vegetables to tree nuts, nursery crops, and floriculture—play a crucial role in the success of American agriculture.[18] However, specialty crops are very labor-intensive, and the availability of a stable workforce has long been one of the greatest challenges facing specialty crop growers.[19]

Artificial intelligence can help improve productivity for specialty crops by detecting diseases, assessing crops' health, maturity, and quality, and predicting yields.[20] Developing mechanization and automation technologies for labor-intensive tasks on farms and in packing facilities has been a priority for the specialty crop industry.[21] However, developing these emerging technologies has been difficult. One impediment is that the broad diversity of specialty crops makes a single technological solution infeasible. Challenges such as the cost-effectiveness of equipment, unreliable network connectivity, and the need for an upskilled workforce hamper the adoption of these technologies in the agricultural sector.

Congress should explore opportunities to address these challenges and fund additional research and development into the use of AI to enhance efficiency in the specialty crop industry.

AI Applications for Derivatives Markets

The Grain Futures Act of 1922 evolved in part to the Commodity Exchange Act of 1936, which was then replaced by the Commodity Futures Trading Commission Act of 1974.[22] That act created the Commodity Futures Trading Commission (CFTC) in 1974 as an independent U.S. federal agency responsible for regulating trading in commodity futures, options, and swaps.

Today, the CFTC oversees markets for agricultural commodities such as livestock, cotton, and milk. [23] These markets not only influence the prices of food and other agricultural products but are also used by America's farmers and ranchers to reveal prices and manage associated risks.[24]

[18] "2017 Census of Agriculture Specialty Crops." National Agricultural Statistics Service, USDA, 5 Dec. 2019, https://www.nass.usda.gov/Publications/AgCensus/2017/Online_Resources/Specialty_Crops/SCROPS.pdf.

[19] "Specialty Crop Farms Have the Highest Labor Cost as a Portion of Total Cash Expenses." Economic Research Service U.S. Department of Agriculture, USDA, 20 Sept. 2022, https://www.ers.usda.gov/data-products/chart-gallery/gallery/chart-detail/?chartId=104773.

[20] Kakarla, Sri Charan et al. "Editorial: Artificial Intelligence Applications in Specialty Crops." Frontiers in plant science vol. 13 866724. 21 April 2022, doi:10.3389/fpls.2022.866724 https://pmc.ncbi.nlm.nih.gov/articles/PMC9069680/

[21] Delheimer, Sara. "Automation Helps Solve Specialty Crop Challenges." USDA, NIFA, 27 Aug. 2020, https://www.usda.gov/media/blog/2020/08/27/automation-helps-solve-specialty-crop-challenges#:~:text=Automation%20won't%20soon%20replace,efficiency%2C%20and%20reducing%20environmental%20impacts.

[22] "Agriculture: A Glossary of Terms, Programs, and Laws, 2005 Edition." CRS Reports, Congressional Research Service, 16 June 2005, https://crsreports.congress.gov/product/pdf/RL/97-905.

[23] "Protecting America's Farmers and Ranchers." CFTC, 24 Sept. 2020, https://www.cftc.gov/Agriculture.

[24] Id.

The CFTC states its mission as:

> *"Ensuring the integrity of these markets allows the agricultural sector continue to do what it does best—feed America and the world."*[25]

Today, the CFTC focuses on fostering market stability and promoting fair competition across agricultural and non-agricultural markets through a principles-based approach that focuses on market risks rather than specific technologies.[26] It also encourages the development of new financial products and trading technologies, including the use of AI, while ensuring they meet regulatory standards.

On May 1, 2024, the CFTC appointed its first Chief AI Officer to guide its efforts to increase AI adoption.[27] The CFTC has already deployed an AI model to detect data anomalies in exchange-reported data.[28] It also explores AI applications in surveillance, stress detection, and compliance enforcement.[29] Finally, the CFTC evaluates how generative AI can improve workforce performance, including upskilling staff, piloting legal research, and using machine learning to detect market manipulation.[30] This will reduce the burden on staff who must contend with an imposing 15 billion records collected daily.[31]

Derivatives Markets and Financial Institutions

For decades, market participants have relied on computers to analyze market data and automate order creation and submission. Algorithmic traders have long utilized complex algorithms to process market data and execute trades without human intervention.

> ***"AI tools represent the latest step in the long-standing use of computers and associated technology to facilitate trading in financial markets.***

Over time, exchanges and futures commission merchants (FCMs) have developed increasingly sophisticated methods to mitigate the risks associated with automated trading decisions, including multiple risk controls on the flow of orders from traders and market-wide protections such as automated trading "circuit breakers."

[25] Id.

[26] Commodity Futures Trading Commission, "2022-2026 Strategic Plan," https://www.cftc.gov/media/7081/CFTC2022_2026StrategicPlan/download.

[27] "Chairman Behnam Designates Ted Kaouk as the CFTC's First Chief Artificial Intelligence Officer." Commodity Futures Trading Commission, 1 May 2024, https://www.cftc.gov/PressRoom/PressReleases/8903-24.

[28] Kaouk, Ted, "Regulators Talk AI." U.S. Department of the Treasury's 2024 Conference on Artificial Intelligence and Financial Stability, 6 June 2024, Washington D.C., https://home.treasury.gov/policy-issues/financial-markets-financial-institutions-and-fiscal-service/financial-stability-oversight-council/2024-conference-on-artificial-intelligence-financial-stability

[29] Id.

[30] Id.

[31] Id.

AI tools represent the latest step in the long-standing use of computers and associated technology to facilitate trading in financial markets. In derivatives markets, AI tools are gradually being integrated into both front-office and back-office activities, offering several opportunities for efficiency and innovation.

Internal Efficiencies

Firms are increasingly using AI to boost internal efficiency through automation. This trend is especially prominent in software development. AI can help software developers save time by assisting in writing and debugging software. This not only saves time but can also increase the quality of the software produced.

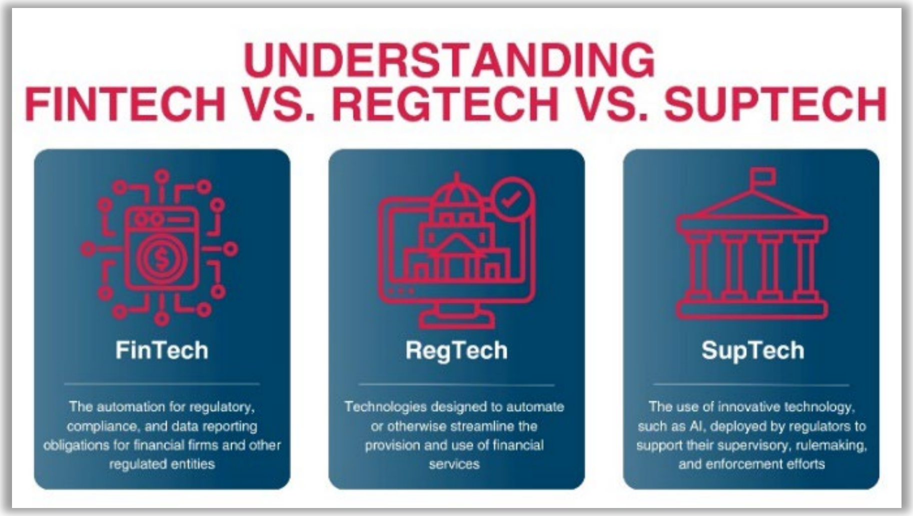
Trading (Price Discovery and Liquidity)

Machine Learning (ML) can enhance trading analytics by helping to establish a price for unique or uncommon financial products that lack comprehensive price data.

A machine-learning similarity analysis compares the characteristics of uncommon financial products to similar ones for which richer and more accurate price data is available. This benefits the over-the-counter derivatives market by improving liquidity, reducing transaction costs and enhancing efficiency in both trade execution and price discovery.

Regulatory Compliance Technology (RegTech) and Compliance

Firms could enhance compliance efforts using large language models to analyze and summarize rules and regulations. In addition, AI can augment Regulatory Compliance Technology (RegTech) by streamlining and improving compliance processes.[32]



UNDERSTANDING
FINTECH VS. REGTECH VS. SUPTECH

FinTech

The automation for regulatory, compliance, and data reporting obligations for financial firms and other regulated entities

RegTech

Technologies designed to automate or otherwise streamline the provision and use of financial services

SupTech

The use of innovative technology, such as AI, deployed by regulators to support their supervisory, rulemaking, and enforcement efforts

[32] Financial Stability Board. Artificial Intelligence and Machine Learning in Financial Services: Market Developments and Financial Stability Implications. 1 Nov. 2017, https://www.fsb.org/uploads/P011117.pdf.

Natural Language Processing (NLP), a subfield of AI, could be particularly useful in interpreting and understanding the vast amount of text in relevant regulations and regulatory reforms. A regulated entity could strengthen its compliance functions by augmenting its risk models and reporting systems with AI that understand the language of regulations.

Managing Risks

Financial institutions have established best practices, governance documents, and risk management frameworks to guide their use of AI.[33] For instance, many require a “human-in-the-loop” to ensure that AI models do not operate entirely autonomously.

As uses of generative AI in the financial services industry continue to evolve, additional risks may accompany novel uses that emerge.

Congress should continue to evaluate the current and future roles of AI in the U.S. derivatives markets by actively engaging with stakeholders such as AI service providers, financial institutions, and regulators.

Congress should continue to explore the general opportunities and risks associated with AI in these markets, with particular emphasis on the CFTC’s initiatives for AI upskilling and hiring, its data and AI strategy, and the need for regulation that balances innovation with oversight.[34]

[33] Buehler, Kevin, et al. “Scaling Gen AI in Banking: Choosing the Best Operating Model.” McKinsey & Company, 22 March 2024, https://www.mckinsey.com/industries/financial-services/our-insights/scaling-gen-ai-in-banking-choosing-the-best-operating-model.

[34] CFTC TAC. “Artificial Intelligence in Financial Markets.” Commodity Futures Trading Commission, 3 May 2024, https://www.cftc.gov/media/10626/TAC_AIReport050224/download.

Key Findings

AI-driven precision agriculture could enhance farm productivity and natural resource management.

Precision agriculture technologies have been used in farming for decades to reduce input requirements, enhance soil and water health, lower operating costs, and increase crop yields. Incorporating AI into technologies such as remote sensing platforms, in-ground sensors, targeted spray systems, and automated mechanical weeders could reduce input requirements, improve soil and water health, reduce operating costs for producers, and attain higher crop yields.

Increased AI integration could enable mechanization and automation technologies and enhance efficiency within the specialty crop industry.

Because of the unique nature of the specialty crop industry, the development of specialized equipment is required to mechanize and automate many of the time-consuming and labor-intensive tasks for the production, harvest, and processing of specialty crops. By utilizing AI to accelerate the research and development of these technologies, the specialty crop industry could become more efficient.

Lack of reliable network connectivity in rural and farming communities impedes AI adoption in the agricultural sector.

According to the Government Accountability Office (GAO), although precision agriculture technologies have been available since the 1990s, only 27% of U.S. farms or ranches utilize such technology. Internet connections that are slow or inconsistent may not meet the bandwidth requirements for data-intensive applications, including AI, and may further limit advanced precision agriculture practices.

AI is already a powerful tool in addressing and combating the wildfire and forest health crises.

AI can monitor and detect active wildfires, monitor trends in forest health and landscapes, and generally improve forest markets.

Greater adoption of AI at USDA could enhance delivery of numerous agriculture programs and reduce costs for farmers and others.

USDA could leverage AI systems to gain deeper insights from the data collected, facilitating easier access to this information when needed.

The CFTC's principles-based approach allows for flexibility in addressing new technologies.

The CFTC is committed to fostering market stability and promoting fair competition. By focusing on the precise risks posed by technology rather than the technology itself, the CFTC can monitor and address the specific risks AI might pose to derivatives markets.

Recommendations

Recommendation: Assess existing programs to identify opportunities for advancing AI in precision agriculture.

AI technologies enable farmers to optimize resource use, reduce waste, and increase crop yields, contributing to both food security and environmental stewardship. However, there are significant barriers to adoption, including the up-front cost of equipment and its complexity. Efficiency gains could be achieved by making AI tools more accessible and affordable to the agricultural sector. Evaluating current USDA programs to advance AI in precision agriculture can further strengthen the agricultural sector's competitiveness and resilience, ensuring it meets the demands of a growing global population.

Recommendation: Pursue further AI research and development to enhance efficiency in specialty crops.

Increased utilization of AI could accelerate the research and development of mechanization and automation technologies for specialty crops. With further research and development— and ultimately adoption—of these tools, the specialty crop industry will benefit from reduced manual labor requirements, increased production efficiency, and improved resource management.

Recommendation: Continue to explore how research and innovation in AI technology could aid land managers in improving forest health through better planning and strategies.

AI can potentially enable more effective, data-driven approaches to forest management. By evaluating tools for precise monitoring and predictive modeling, land managers can proactively address challenges such as wildfires, ultimately preserving forest ecosystems, protecting biodiversity, and promoting sustainable resource use for future generations.

Recommendation: Direct USDA to better utilize AI in program delivery.

By utilizing AI technologies, USDA could potentially improve data analysis, streamline processes, and provide more accurate support to farmers and producers. This could lead to better resource allocation, timelier assistance, and more targeted programs, ultimately strengthening the agricultural sector. There is also potential for USDA to reduce waste, fraud, and abuse by utilizing AI to conduct detailed data analyses. When collecting information and deploying AI technologies, the USDA should employ responsible government AI principles. For more information, please see the Government Use chapter.

Recommendation: Continue to review the application of the CFTC’s principles-based framework to ensure it captures unique risks posed by AI in financial markets.

Ongoing oversight of the CFTC’s implementation of the principles-based framework for market regulation will encourage the thoughtful integration of market activities utilizing AI tools into the existing framework. Congress should also continue to evaluate the statutory principles to identify regulatory gaps and mitigate systemic risks that may arise from using AI and are not captured by existing requirements.

HEALTHCARE

Background

Artificial intelligence (AI) technologies have the potential to improve multiple aspects of healthcare research, diagnosis, and care delivery. AI can quickly analyze large data sets and, in so doing, has the potential to improve diagnostic accuracy, streamline operations, and automate routine tasks, all of which can improve efficiency and efficacy in treatment and reduce the burden for healthcare practitioners, freeing up more time for patient care. At the same time, it is also important to be mindful of AI’s potential to possess bias due to data limitations, which may lead to misallocation of resources and inaccurate diagnoses and treatment, particularly for populations underrepresented in the data. Additionally, AI health systems require large amounts of sensitive patient data, such as medical records, personal information, and payment information, which can be vulnerable to abuse and breaches.

AI technologies have been used in healthcare in some form for decades under the names “clinical informatics,” “health information technology,” or “Software as a Medical Device (SaMD).” The use of AI in clinical settings has been touted as a means to alleviate administrative burdens and allow clinicians to focus more on providing care. This contrasts with fully automated decision-making tools, which do not pair with clinicians’ care.

The continual evolution in AI capabilities and integration has raised new policy issues. Some of the most prominent challenges involve data availability and quality, incomplete or inaccurate responses, non-individualized recommendations, decision transparency, data privacy and cybersecurity, interoperability between existing systems and AI, liability for errors made or enabled by AI models, and biased decision-making as well as the deployment of these models in a way that promotes financial gain over patient care and safety.

These issues can have serious implications for a patient's health and the healthcare sector at large. In December 2023, the Biden Administration announced voluntary commitments to the safe, secure, and trustworthy use and purchase of AI in healthcare from twenty-eight providers and payers.[1]

AI Adoption in the Healthcare System

Use of AI in Drug Development

AI is already used in drug development, where it promises to significantly expedite the discovery, design, and testing of drug candidates.

> ***"The average cost of developing a new drug and bringing it to market is estimated to be between $314 million and $2.8 billion, including expenditures on failed trials. Use of AI technologies may decrease the time and cost required to get a drug to market.***

It takes an average of twelve years for a developed drug to transition from preclinical testing to receiving approval by the U.S. Food and Drug Administration (FDA). This period does not account for the additional years required to research and develop the drug before preclinical testing.[2]

The average cost of developing and bringing a new drug to market is estimated to be between $314 million and $2.8 billion, including expenditures on failed trials.[3] The use of AI technologies may decrease the time and cost required to get a drug to market.[4]

This potential efficiency could reduce the price of drugs and speed their market entry. It could also make it cost-effective to invest in producing orphan drugs and drugs for rarer diseases.[5,6]

[1] The White House. "Delivering on the Promise of AI to Improve Health Outcomes." The White House, The United States Government, 14 Dec. 2023, www.whitehouse.gov/briefing-room/blog/2023/12/14/delivering-on-the-promise-of-ai-to-improve-health-outcomes/.

[2] Van Norman, Gail A. "Drugs, Devices, and the FDA: Part 1: An Overview of Approval Processes for Drugs." JACC: Basic to Translational Science, vol. 1, no. 3, 25 April 2016, pp. 170–179, https://doi.org/10.1016/j.jacbts.2016.03.002.

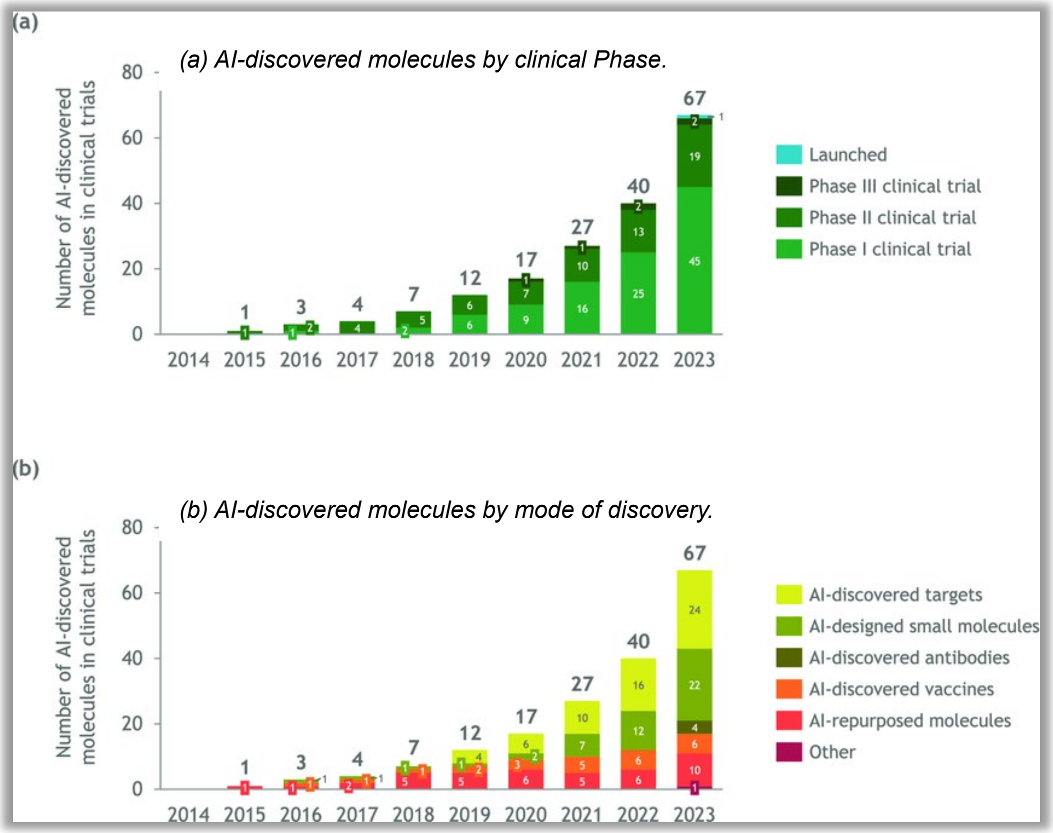
[3] Wouters, Olivier J., et al. "Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009-2018." JAMA, vol. 323, no. 9, 3 March 2020, p. 844, https://doi.org/10.1001/jama.2020.1166.

[4] "AI's potential to accelerate drug discovery needs a reality check." Nature, vol. 622, no. 7982, 10 Oct. 2023, pp. 217–217, https://doi.org/10.1038/d41586-023-03172-6.

[5] An orphan drug is a drug for a rare diseases or condition that was "orphaned" or discontinued because there was not enough financial incentive to continue the costs of development or production. Office of the Commissioner. "Rare Diseases at FDA." U.S. Food and Drug Administration, U.S. Food and Drug Administration, December 2022, www.fda.gov/patients/rare-diseases-fda.

[6] Irissarry, Carla, and Thierry Burger-Helmchen. "Using Artificial Intelligence to Advance the Research and Development of Orphan Drugs." Businesses, vol. 4, no. 3, September 2024, pp. 453–472, https://doi.org/10.3390/businesses4030028.

The FDA recently published a discussion paper on the use of AI in drug development and manufacturing.[7] According to the FDA, an increasing number of investigational drug applications contain AI components, with AI most commonly being included in the clinical development or research phase of drug development.[8] While only one application in 2016 relied on AI elements, nearly 130 applications in 2021 and over 300 in 2024 integrated AI into drug development, drug discovery, and post-market safety monitoring across various therapeutic areas.[9] As of 2024, nearly 70 drugs with some form of AI involvement had made it into clinical trials with patients, as represented in the figure immediately below.[10]



(a) AI-discovered molecules by clinical Phase.

(b) AI-discovered molecules by mode of discovery.

Number of molecules discovered by AI-first Biotechs that have entered clinical trials. The analysis includes molecules that were partnered with pharmaceutical companies and excludes COVID-19-related molecules.

[7] Cavazzoni, Patrizia. "FDA Releases Two Discussion Papers to Spur Conversation about Artificial Intelligence and Machine Learning in Drug Development and Manufacturing." U.S. Food and Drug Administration, U.S. Food and Drug Administration, May 2023, www.fda.gov/news-events/fda-voices/fda-releases-two-discussion-papers-spur-conversation-about-artificial-intelligence-and-machine.

[8] FDA, et al. "Using Artificial Intelligence & Machine Learning in the Development of Drug & Biological Products." U.S. Food and Drug Administration, U.S. Food and Drug Administration, 2023, www.fda.gov/media/167973/download.

[9] Eglovitch, Joanne S. "FDA Plans to Release AI Drug Development Guidance This Year." RAPS, 30 May 2024, www.raps.org/news-and-articles/news-articles/2024/5/fda-plans-to-release-ai-drug-development-guidance.

[10] KP Jayatunga, Madura, et al. "How successful are AI-discovered drugs in clinical trials? A first analysis and emerging lessons." Drug Discovery Today, vol. 29, no. 6, June 2024, p. 104009, https://doi.org/10.1016/j.drudis.2024.104009.

Researchers use machine learning (ML) and generative AI throughout the first three phases of drug development: drug discovery, preclinical trials, and clinical trials. For example, during the drug discovery phase, machine learning can be used to screen drug compounds by reviewing the effects of the chosen compound on a target, such as a protein.[11] Conducting virtual screening before physical screening via biological tests can help researchers narrow in on promising compounds rather than run expensive testing across all available compounds.[12]

Researchers are also using generative AI in drug development. These systems can be trained on existing biological and chemical data and recognize unique interactions, patterns, or connections to create novel molecular structures with desired properties for drug candidates.[13] The use of generative AI allows this process to be completed faster than a human manually checking the data could. Generative AI may also check interactions or patterns not immediately apparent to the human mind.[14] Additionally, generative AI has the potential to find alternative uses for current drugs on the market by reviewing the way the components of the drug interact with the body.[15]

In preclinical trials, before humans are included in testing, machine learning can predict ways the drug may interact with the human body, including ways in which it may be toxic.[16] Generative AI may assist in simulation testing of the drug's performance, such as varying dosages or at different stages of tumor progression or illness.[17]

On average, less than fourteen percent of drugs that enter Phase I of clinical trials ultimately receive FDA approval.[18] Machine learning can assist researchers in increasing the probability of a successful clinical trial by adjusting the number of design variables to best suit the trial.[19] Generative AI may assist in faster, cheaper, and more efficient trials by helping to streamline design, find eligible patients, recruit and retain them, and serve as a resource for patients with questions regarding the trial.[20]

[11] U.S. Government Accountability Office. "Artificial Intelligence in Health Care: Benefits and Challenges of Machine Learning in Drug Development." U.S. Government Accountability Office, 6 Feb. 2020, www.gao.gov/products/gao-20-215sp.
[12] Id.
[13] U.S. Government Accountability Office. "Science & Tech Spotlight: Generative AI in Health Care." U.S. Government Accountability Office, 9 Sept. 2024, www.gao.gov/products/gao-24-107634.
[14] Id.
[15] Yan, Chao, et al. "Leveraging generative AI to prioritize drug repurposing candidates for Alzheimer's disease with real-world clinical validation." Npj Digital Medicine, vol. 7, no. 1, February 2024, https://doi.org/10.1038/s41746-024-01038-3.
[16] Supra 11.
[17] Maria Bordukova, Nikita Makarov, Raul Rodriguez-Esteben, Fabian Schmich & Michael P. Menden, Generative artificial intelligence empowers digital twins in drug discovery and clinical trials, Expert Opinion on Drug Discovery, 19, 33-42, October 27, 2023, https://www.tandfonline.com/doi/full/10.1080/17460441.2023.2273839#d1e544.
[18] Chi Heem Wong, Kien Wei Siah, & Andrew W Lo, Estimation of clinical trial success rates and related parameters, Biostatistics 20, 273-286, 9 April 2019. https://pubmed.ncbi.nlm.nih.gov/29394327/.
[19] Supra 11.
[20] Matthew Hutson, How AI is being used to accelerate clinical trials, Nature, 13 March 2024, https://www.nature.com/articles/d41586-024-00753-x.

While strides are being made in using AI in drug development, challenges have hindered its adoption or impact. These include a limited understanding of how AI makes decisions, limited high-quality data for training the AI model, a need for researchers trained in AI skills, and uncertainty regarding regulatory limitations on the use of AI in the medical field.[21]

For example, because decisions in drug development can vastly affect people's health and well-being, it will be important to understand how AI may affect the decisions on what drugs to make and what clinical trials to conduct. Further, as AI systems require increasing amounts of personal information to support research in this field, there are significant concerns related to data privacy and security.[22]

Fundamental Biomedical Research Accelerates AI Innovation

AI, machine learning, and informatics have been used for decades to drive discoveries in fundamental biomedical research. Researchers are increasingly leveraging data from biobanks, electronic health records, wearable sensors, and genomic and microbiome sequences to better investigate the biological mechanisms that represent the complexity of human health and disease.[23]

The discovery and understanding of these biological mechanisms often termed basic or fundamental research, provides scientific data for AI to leverage into useful findings and outcomes.[24] For example, AlphaFold's breakthrough in using AI to predict protein structure resulted from decades of fundamental research on the biophysical forces behind protein structure.[25] A broad portfolio of such fundamental research is funded by the National Institutes of Health (NIH), National Science Foundation (NSF), Department of Defense (DOD), and Department of Energy (DOE). It is critical to continue support for fundamental and curiosity-driven research to maintain America's edge and status as a world leader in biomedical AI research.

[21] Data quality is a measure of how accurate, complete, valid, consistent, and tailored it is to the needs at hand. Some aspects of high-quality data may include that the data is: 1) complete and not missing values, 2) is free from bias, 3) is accurate, 4) is consistent and there are not differing formats or units of measure, and 5) free from duplicates.

[22] Blanco-González, Alexandre et al. "The Role of AI in Drug Discovery: Challenges, Opportunities, and Strategies." Pharmaceuticals (Basel, Switzerland) vol. 16,6 891. 18 Jun. 2023, https://doi.org/10.3390/ph16060891.

[23] Acosta et al., Multimodal biomedical AI, Nature, 15 Sept. 2022, https://www.nature.com/articles/s41591-022-01981-2.

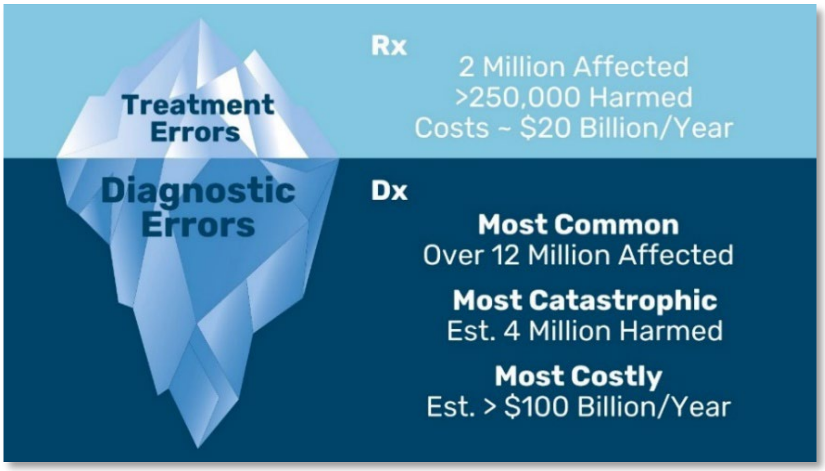
[24] Lorsch, Tabak, and Bertagnolli, Applied research won't flourish without basic science, Elife, 24 Sept. 2024, https://elifesciences.org/articles/102368.

[25] Jumper et al., Highly accurate protein structure prediction with AlphaFold, Nature. July 2021, https://www.nature.com/articles/s41586-021-03819-2.; see also: Ewen Callaway, The huge protein database that spawned AlphaFold and biology's AI revolution, Nature. 18 Oct. 2024, https://www.nature.com/articles/d41586-024-03423-0.

Use of AI in Diagnostics

AI has the potential to significantly improve diagnostic capabilities. Some AI technology is already being utilized to assist with imaging, such as X-rays or MRIs for certain cancers, heart disease, and Alzheimer's disease.

Diagnostic errors are among the most catastrophic, common, and costly medical errors, likely exceeding $100 billion a year in aggregate costs.[26] The graphic representing the magnitude of treatment and diagnostic errors shows that diagnostic errors are the overwhelming share of patient safety errors.[27] Using AI to assist providers may augment clinical diagnostic practices to reduce errors, detect diseases earlier, and more consistently analyze medical data.[28]



Machine learning technologies are not intended to provide a diagnosis. Instead, they may be able to help medical professionals in the decision-making process that eventually leads to a diagnosis.

For example, machine learning diagnostic technologies can screen patients more quickly than a human doctor, identifying cases of concern that, if the medical professional agrees are valid, could move patients forward in priority amidst long referral wait times.[29]

Moreover, machine learning may consistently apply criteria to diagnostic images across all cases regardless of patient or location. This has potential utility in locations where medical care is less readily available. For example, machine learning technologies can detect signs of diabetic retinopathy by interpreting images from a specialty camera in less than a minute.[30]

[26] Society to Improve Diagnosis in Medicine, The Roadmap for Research to Improve Diagnosis, Part 1: Converting National Academy of Medicine Recommendations into Policy Action, 7 Feb. 2018, https://pmc.ncbi.nlm.nih.gov/articles/PMC6971119/.
[27] Id.
[28] Id.
[29] Id.
[30] Id.

Generative AI can also assist in medical imaging. Images from MRI machines or other devices may have noise or graininess that can hinder the ability of medical professionals to interpret the images.[31] Rather than submit a patient to another round of radiation via a new scan, generative AI has shown the ability to identify areas of noise and then generate a cleaner image free of noise.[32] While this technology has the potential for research into radiation reduction techniques, it can also potentially introduce errors such as fake lesions or blurs in areas without lesions.[33]

It remains essential for a medical professional to use professional judgment in interpreting AI-augmented images. Beyond accuracy issues, the use of AI systems in diagnosis can have significant implications related to data privacy, security, transparency, and fairness. These issues transcend the health industry and are addressed in greater depth in other chapters of this report.

Clinical Decision-Making

AI tools have shown some promise in augmenting patient care in clinical applications, such as predicting the health trajectories of patients, recommending treatments, and supporting population health management. AI-driven computer-assisted image visualization can also be valuable in areas with a shortage of medical specialists.

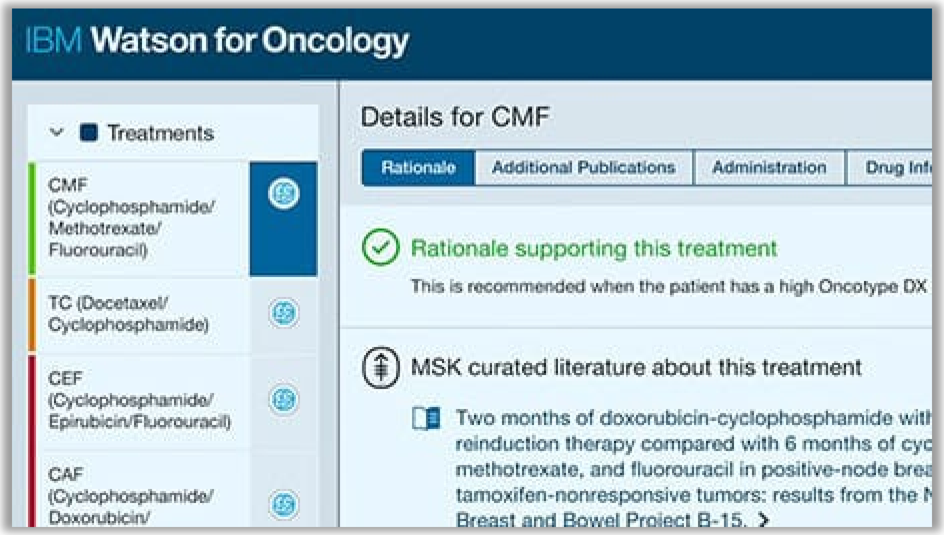
Researchers are evaluating and validating machine learning-based tools that can use existing data from previous patients to predict the health outcomes of current patients. AI machine learning tools have also been implemented for accurate surveillance, such as at the Center for Disease Control (CDC) with the National Vital Statistics System using MedCoder, which can code 90% of mortality records automatically.[34]

For example, several hospital systems have developed AI tools that can use real-time data to predict sepsis—a life-threatening blood infection—before obvious signs occur.[35] The COMPOSER model by the University of California San Diego Health continuously monitors patients across more than 150 variables throughout their time in the emergency room to recognize any changes indicative of sepsis.[36] Using COMPOSER has resulted in a 17% reduction in mortality.[37]

[31] Supra 13.
[32] Id.
[33] Id.
[34] Center for Disease Control, "Artificial Intelligence and Machine Learning Applying Advanced Tools for Public Health, July 2023, https://www.cdc.gov/surveillance/data-modernization/technologies/ai-ml.html.
[35] Laura Cech, AI to detect sepsis, John Hopkins Magazine, 2022, https://hub.jhu.edu/magazine/2022/winter/ai-technology-to-detect-sepsis/.; see also: Jeanna Vazquez, Study: AI surveillance tool successfully helps to predict sepsis, saves lives, UC San Diego Health, 23 Jan. 2024, https://health.ucsd.edu/news/press-releases/2024-01-23-study-ai-surveillance-tool-successfully-helps-to-predict-sepsis-saves-lives/.
[36] Id.
[37] Id.

Another example of an AI-enabled clinical decision support tool is one that may be able to recommend specific treatments to healthcare professionals based on the patient's symptoms and medical history.[38] These machine learning-based tools use data from other patients with similar conditions or histories to determine which treatment options led to the best outcomes. For example, one company's clinical decision support system analyzes a patient's data against historical case data and data supported by literature and journals to provide oncology treatment options to clinicians.[39] Clinicians can then consider the recommendations from these systems when deciding how to treat a patient.

An example of an artificial intelligence assistant decision system for oncology treatment options.[40]



IBM Watson for Oncology

CMF (Cyclophosphamide/ Methotrexate/ Fluorouracil)

TC (Docetaxel/ Cyclophosphamide)

CEF (Cyclophosphamide/ Epirubicin/Fluorouracil)

CAF (Cyclophosphamide/ Doxorubicin/

Details for CMF

Rationale supporting this treatment

This is recommended when the patient has a high Oncotype DX

MSK curated literature about this treatment

Two months of doxorubicin-cyclophosphamide with reinduction therapy compared with 6 months of cyc methotrexate, and fluorouracil in positive-node brea tamoxifen-nonresponsive tumors: results from the N Breast and Bowel Project B-15.

While clinical decision support tools have been found to help address health disparities in some cases—such as improving quality and access to care—they have also been found to perpetuate health disparities. For example, some of these tools incorporate racial biases that have detrimental effects on medical and clinical education and patient health outcomes.[41] However, research also shows that AI can reduce these biases if applied correctly.[42]

Further, using these systems has significant liability and provides autonomy concerns. Questions remain about who is held liable for an incorrect health decision recommended by clinical decision support tools.

[38] Id.
[39] IBM, 5725-W51 IBM Watson for Oncology, 1 Aug. 2023, https://www.ibm.com/docs/en/announcements/watson-oncology?region=CAN.
[40] Jasimine Pennic, "Jupiter Medical Center to Implement Watson for Oncology for Data-Driven Cancer Treatment Decisions", HIT Consultant, February 2017, https://hitconsultant.net/2017/02/01/jupiter-medical-center-watson-for-oncology/.
[41] Vyas, Darshali A et al. "Hidden in Plain Sight - Reconsidering the Use of Race Correction in Clinical Algorithms." The New England journal of medicine vol. 383, 2020: 874-882. https://doi.org/10.1056/nejmms2004740
[42] Green, B Lee et al. "Accelerating health disparities research with artificial intelligence." Frontiers in digital health vol. 6 1330160. January 2024. https://pmc.ncbi.nlm.nih.gov/articles/PMC10844447/.

Population Health Management

AI-enabled tools may also support population health management activities outside of clinical decision support tools.[43] Using population-level health data, such as data across the entire United States or even within a community, can help medical professionals identify health risks for a group of individuals or the entire community and identify the most at-risk.[44] This, in turn, may allow these individuals to be targeted for specialized care programs to address chronic diseases or complications preemptively.[45] AI systems may use population-level health data to create a model to predict how different population groups may respond to varying healthcare initiatives or strategies and optimize currently available programs.[46]

Administrative Clinical Uses

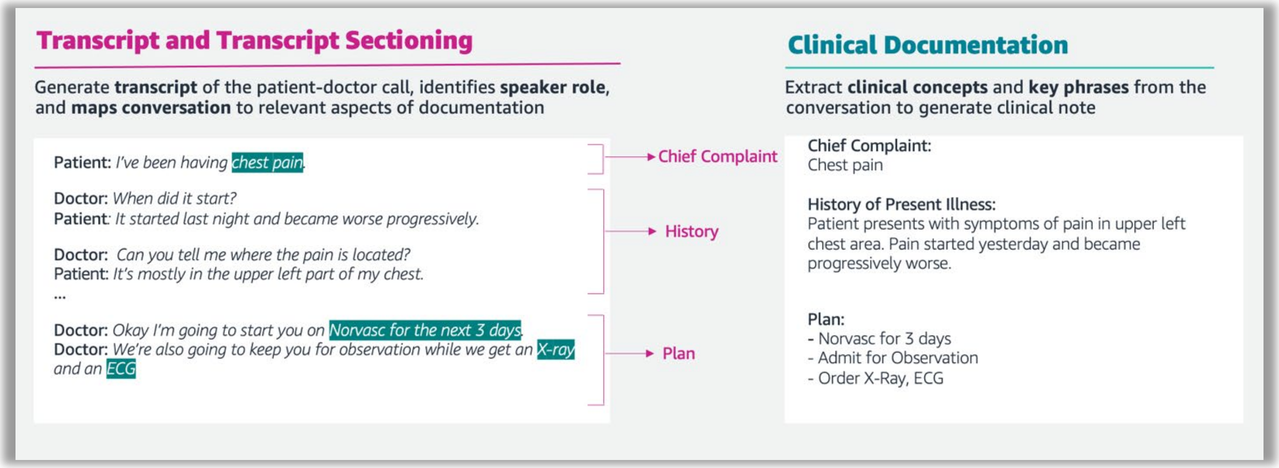
Healthcare facilities use electronic health records (EHR)[47] to record patient demographics, medical history, diagnoses, immunizations, notes, laboratory and radiology data, vitals, and more. They are a clear source of high-volume and medically pertinent data that could serve as training data for health AI tools, such as predictive diagnostic support. However, EHR data may not be high-quality. Since medical personnel have limited time with each patient to listen, document the interaction, and provide care, their notes may be brief or written later based on memory.

Some healthcare facilities are beginning to deploy generative AI tools for note-taking and patient portal messaging to assist healthcare professionals. For example, generative AI tools are available that record the interaction between medical professionals and patients, transcribe those conversations into clinical notes, and provide summaries of the interactions via formatted clinical notes.[48] This allows the provider to spend more time interacting with the patient without concern for missing information that must be entered into the EHR.

This can also help with physician burnout, as physicians consistently cite the administrative burden related to EHR documentation as a top contributor to burnout.[49] It can also increase the quality of the clinical notations, improving later review of patient records and creating data amenable for use in AI training.

[43] Supra 35.
[44] Id.
[45] Id.
[46] PwC, AI-powered healthcare: Shaping the future of population health, 3 March 2024, https://www.pwc.com/m1/en/publications/documents/2024/shaping-the-future-of-population-health.pdf.
[47] EHRs are often used interchangeably with electronic patient records (EPR) or electronic medical records (EMR).
[48] Supra 13.
[49] Tania Tajirian et al., The influence of electronic health record use on physician burnout: cross-sectional survey, Journal of Medical Internet Research, 2020, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7392132/.

An example of how a software vendor can use generative AI for clinical documentation.[50]



Transcript and Transcript Sectioning

Generate **transcript** of the patient-doctor call, identifies **speaker role**, and **maps conversation** to relevant aspects of documentation

Patient: *I've been having chest pain.*

Doctor: *When did it start?*
Patient: *It started last night and became worse progressively.*

Doctor: *Can you tell me where the pain is located?*
Patient: *It's mostly in the upper left part of my chest.*
...

Doctor: *Okay I'm going to start you on Norvasc for the next 3 days.*
Doctor: *We're also going to keep you for observation while we get an X-ray and an ECG*

Clinical Documentation

Extract **clinical concepts** and **key phrases** from the conversation to generate clinical note

Chief Complaint:
Chest pain

History of Present Illness:
Patient presents with symptoms of pain in upper left chest area. Pain started yesterday and became progressively worse.

Plan:
- Norvasc for 3 days
- Admit for Observation
- Order X-Ray, ECG

Generative AI can also assist with the administrative burden of documentation preparation for insurance submission (discussed in more detail below). Physicians can use AI systems to compile the information needed for the insurance preauthorization documentation that medical professionals must send to insurance companies. This form of AI automation would free up the time doctors and staff spend on this task.

However, if implemented poorly, these same tools could degrade EHRs and other medical documentation.[51] Medical professionals will need additional tools, including evaluations, to ensure AI systems are not inadvertently diminishing the quality of care or efficiency in healthcare.

Use of AI in the Development of Medical Devices and Software

AI is changing the field of medical products by playing a crucial role in the research and development of innovative therapeutics. The FDA's Center for Devices and Radiological Health (CDRH) regulates AI-ML-enabled medical devices. CDRH has authorized over 800 non-generative AI/ML-enabled devices under its existing medical device authorities.[52]

[50] Jason Mark, et al., "Introducing AWS HealthScribe – automatically generate clinical notes from patient-clinician conversations using AWS HealthScribe", Amazon Web Services, July 2023, https://aws.amazon.com/blogs/industries/industries-introducing-aws-healthscribe/.

[51] Liam McCoy, et al., "Large Language Models and the Degradation of the Medical Record", the New England Journal of Medicine, 26 Oct. 2024, https://www.nejm.org/doi/full/10.1056/NEJMp2405999.

[52] U.S. Food and Drug Administration (FDA), Artificial intelligence and machine learning (AI/ML)-enabled medical devices, https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices.

AI has the potential to significantly improve specific medical devices and software, enhancing their functionality, accuracy, efficiency, and overall impact on healthcare and patient outcomes. By leveraging AI in medical devices and software, healthcare providers can improve diagnostic accuracy, personalize treatments, enhance surgical precision, optimize operational efficiency, and ultimately deliver better patient care.

Health Insurance Decisions

Payers of healthcare services are also critical stakeholders for AI use in the U.S. healthcare system, both for the coverage of AI-provided services and devices and for the use of AI tools in the health insurance industry. Payment for using AI in healthcare services remains an unanswered question, both in its implementation within health systems and in how reimbursement occurs for its use.

Some AI applications, like IDx-DR (a diabetic retinopathy diagnostic), have received traditional Centers for Medicare and Medicaid Services (CMS) coverage codes. CMS makes coverage and payment determinations related to items and services provided to Medicare beneficiaries once the FDA has determined the safety and efficacy of the item or service, where relevant. To get Medicare coverage, an item or service must be "reasonable and necessary for the diagnosis or treatment of illness or injury or to improve the function of a malformed body member."[53]

> ***"As more evidence is developed regarding the application of certain tools in healthcare settings, particularly among Medicare populations, further evaluation of current CMS payment systems will be necessary.***

CMS has been deliberate in deciding whether to cover a given software product or service once it has received FDA clearance or approval. CMS's current AI coverage framework focuses on FDA-approved software that helps clinicians make decisions through algorithms or predictive modeling.[54] CMS has allowed for limited Medicare coverage of AI technologies in cases where the services meet Medicare's coverage criteria.

Some technologies, such as medical devices with AI technologies, may already be covered under CMS codes, while others, such as administrative technologies, may not need reimbursement. Questions linger regarding whether current Medicare policies will suit all appropriate AI technologies in healthcare. As more evidence is developed regarding applying certain tools in healthcare settings, particularly among Medicare populations, further evaluation of current CMS payment systems will be necessary.

[53] Social Security Administration, "Exclusions from Coverage and Medicare as Secondary Payer: Section 1862(a)(1)(A) of Title XVIII," https://www.ssa.gov/OP_Home/ssact/title18/1862.htm.
[54] Department of Health and Human Services, Medicare first defined "Artificial Intelligence in its Outpatient Prospective Payment System (OPPS) rules in 2023. Federal Register, July 2024, https://www.federalregister.gov/documents/2024/07/22/2024-15087/medicare-and-medicaid-programs-hospital-outpatient-prospective-payment-and-ambulatory-surgical

The National Health Care Anti-Fraud Association, a public-private partnership of insurers and federal and state agencies, estimates healthcare fraud costs the United States up to $300 billion annually and represents an estimated 3 to 10% of total healthcare spending.[55]

With respect to Medicare, the U.S. Department of Health and Human Services (HHS) Office of Inspector General (OIG) has found that fraudulent billing and prescription writing schemes often follow specific patterns.[56]

In January of 2023, HHS announced that it would launch an AI-enabled pilot program to "streamline fraud identification" by CMS.[57]

However, stakeholders have criticized the implementation of AI tools by health insurers for insurance decisions for a lack of transparency in coverage decisions. While Medicare Advantage insurers have flexibility in Medicare benefit design, questions have been raised about the use of AI systems created to predict estimated lengths of stay based on statistical metrics and then rejecting patient requests for care that exceeded this length, even if supported by caregiver opinion.[58]

In another example, when AI-driven denials of elderly patients' claims for extended care were appealed to federal administrative law judges, approximately 90% were reversed.[59]

In response to these growing instances, CMS adopted a final rule for 2024 that Medicare Advantage plans must make medical necessity determinations "based on the circumstances of the specific individual…as opposed to using an algorithm or software that doesn't account for an individual's circumstances" and those determinations "must be reviewed by a physician or other appropriate healthcare professional."[60]

[55] National Health Care Anti-Fraud Association. "The Challenge of Health Care Fraud." National Health Care Anti-Fraud Association, https://www.nhcaa.org/tools-insights/about-health-care-fraud/the-challenge-of-health-care-fraud/.

[56] U.S. Department of Health and Human Services (HHS) Office of the Secretary and U.S. Department of Justice (DOJ) Office of the Attorney General, Annual report of the Departments of Health and Human Services and Justice: Health Care Fraud and Abuse Control Program FY 2022, November 2023, https://oig.hhs.gov/publications/docs/hcfac/FY2022-hcfac.pdf.

[57] Nihal Krishan, HHS CIO Mathias says tree-based AI models helping to combat Medicare fraud, FEDSCOOP, 18 Jan. 2023, https://fedscoop.com/hhs-cio-mathias-says-tree-based-ai-models-helping-to-combat-medicare-fraud/; see also: Casey Ross & Bob Herman, Denied by AI: How Medicare Advantage plans use algorithms to cut off care for seniors in need, STAT, 13 March 2023, https://www.statnews.com/2023/03/13/medicare-advantage-plans-denial-artificial-intelligence/.

[59] Laney, Douglas B. "AI Ethics Essentials: Lawsuit Over AI Denial Of Healthcare." Forbes, 16 Nov. 2023, www.forbes.com/sites/douglaslaney/2023/11/16/ai-ethics-essentials-lawsuit-over-ai-denial-of-healthcare/.

[60] Medicare Program; Contract Year 2024 Policy and Technical Changes to the Medicare Advantage Program, Medicare Prescription Drug Benefit Program, Medicare Cost Plan Program, and Programs of All-Inclusive Care for the Elderly, 42 CFR Parts 417, 422, 423, 455, and 460 [CMS-4201-F], 12 April 2023, https://www.federalregister.gov/documents/2023/04/12/2023-07115/medicare-program-contract-year-2024-policy-and-technical-changes-to-the-medicare-advantage-program.

These concerns are not isolated to Medicare Advantage but extend across health coverage programs. For example, the Medicaid and Children's Health Insurance Program (CHIP) Payment and Access Commission (MACPAC) included several recommendations in its March 2024 Report to Congress aimed at increasing the transparency and improving the monitoring of denials and appeals in Medicaid-managed care and in October 2024, MACPAC announced that it will expand this body of work, including by examining the extent to which states and MCOs use artificial intelligence to automate parts of the prior authorization process.[61]

There is potential to use AI as a medical management tool in some instances, but there are concerns that these applications could create unnecessary denials and lack of access to necessary treatments when AI produces inaccurate or biased results.

Policy Challenges Confronting AI Adoption in Healthcare

Data Availability, Utility, and Quality

AI systems must be trained on large data sets to achieve performance levels necessary for successful healthcare applications. The data of the required type must exist, be of high quality, and be able to be transferred and/or combined with other data. For example, data for AI models may come from EHRs, radiology images, patient surveys, and scientific data such as genetic sequencing. Epic's EHR system "Cosmos" represents data for 238 million patients, potentially serving as an immense resource for AI training.

Trends in storing and managing healthcare data, including via EHRs, have developed an ecosystem rich in healthcare data. However, this ecosystem lacks efficient mechanisms for integrating and merging data sets outside of their silos, making it challenging to utilize the data fully. For example, various EHRs may have different formats and designs that render the data inoperable across domains.[62]

If health data from different EHR systems cannot be integrated because formats are not standardized, they cannot be combined to train AI models. Consequently, those models may be trained only for small populations and lack generalizability to broader populations.[63] Furthermore, many federal research agencies now possess legacy datasets without an established sunset date. How research agencies organize, manage, and share their biomedical data with researchers will be critically important for downstream AI innovation.

[61] Medicaid and CHIP Payment and Access Commission (MACPAC), Report to Congress on Medicaid and CHIP 2024, https://www.macpac.gov/wp-content/uploads/2024/03/MACPAC_March-2024-WEB-Final-508.pdf.
[62] U.S. Gov't Accountability Office (GAO), GAO-21-7SP, Artificial Intelligence in Health Care: Benefits and Challenges of Technologies to Augment Patient Care (2020), https://www.gao.gov/products/gao-21-7sp.
[63] U.S. Gov't Accountability Office (GAO), GAO-20-215SP, Artificial Intelligence in Health Care: Benefits and Challenges of Machine Learning in Drug Development (2020), https://www.gao.gov/products/gao-20-215sp.

Cultural, societal, and regulatory issues further complicate data availability, utility, and quality issues. Health information is private, so existing protections may prevent its sale to or use by private companies. However, there are ways to use patient data once it is sufficiently deidentified, i.e., altered to protect the patient’s identity.

Unfortunately, the information removed to deidentify the data—such as gender or race—may be information necessary to train the AI model. For example, a model may need to be trained on data representative of a population, which in turn requires revealing the characteristics of that population to ensure representativeness.

This additional information may also be necessary to ensure that the model is not biased or otherwise flawed by the characteristics of the training data. A biased model can unknowingly perform poorly or cause inequitable outcomes for different groups of people. Moreover, there continue to be issues with insufficient data quality. For instance, the information entered into the EHRs by a provider may be incomplete, lack sufficient detail, or contain biased information, making it unsatisfactory for use.

Transparency

Transparency in AI decision-making is required for medical professionals to be confident in its use and for patients to trust and accept AI use in their healthcare. A lack of transparency, interpretability, and explainability regarding how an AI model makes decisions can have grave implications in the healthcare sector.

> **“Transparency in AI decision-making is required for medical professionals to be confident in its use and for patients to trust and accept AI use in their healthcare.**

Some uses of AI in healthcare—such as the summarization of patient and provider interactions—may not require an understanding of how the model completed the task. In other cases, it is essential to interpret, at least to some extent, how a model arrived at its decision, such as denying medical coverage or a warning that a medical image is likely suspicious and merits further scrutiny.

This lack of transparency can have various causes. Some can be rooted in the AI developer declining to share model information due to intellectual property concerns, compounded by a failure to understand how the algorithms work in more complex systems.

This is particularly prominent in generative AI applications, in which outputs can contain some degree of skewed or misleading information, especially when describing information that is rare or underrepresented in the training data set.

Another issue is that medical professionals may lack the training to understand whether an error occurred in the AI decision-making process. This not only undermines trust in the output of the AI system but also thwarts the ability of medical professionals and regulators to determine if an AI tool is safe and effective.[64]

This limited understanding may also cause medical professionals to either underutilize or become over-reliant on the AI tool, compromising its effectiveness or bypassing the human-based safeguards, respectively.[65] Hospitals and medical professionals will also need transparency into AI-enabled healthcare tools to understand their efficacy, compare them, and determine what best fits their patients' needs.

Bias

One major set of data-related risks involved with AI systems is the potential for bias, which occurs when an algorithm produces results that are systemically skewed. Bias can be found in an AI system during development or as the system is being deployed.

Biased AI can stem from algorithmic biases, such as when the training data exhibit skew, bias, or underrepresentation/overrepresentation of specific populations. Alternatively, bias in an AI system could potentially be rooted, either intentionally or unintentionally, in human bias that manifests in the design of the AI system.[66]

For example, certain computer-aided diagnosis systems were found to be less accurate for Black patients than White patients because the training data used had insufficient representation of minority groups, leading to skewed predictions.[67] These challenges are especially acute in healthcare, as incorrect assumptions about particular patient populations can result in inappropriate care and worse outcomes, including death. Such issues may cause individuals to lose trust in healthcare technologies and applications. As with other sectors, AI in healthcare could benefit from standards and evaluations to detect and mitigate biased and erroneous outputs by these systems.

Privacy and Cybersecurity

AI tools require large amounts of data—often patient data—that may be used by or shared between various groups, increasing the risk to patient data privacy. There are concerns among providers, healthcare systems, and patients about who has access to data, how it is being used, and if it is safely secured. EHRs contain not only information about a patient's medical conditions and medications, but they also contain addresses, social security numbers, and even, at times, billing information.[68]

[64] Supra 35.
[65] Id.
[66] Norori, Natalia et al. "Addressing bias in big data and AI for healthcare: A call for open science." Patterns (New York, N.Y.) vol. 2,10 100347. 8 Oct. 2021, https://pubmed.ncbi.nlm.nih.gov/34693373/.
[67] IBM Data and AI team, Shedding light on AI bias with real world examples, 16 Oct. 2023, https://www.ibm.com/blog/shedding-light-on-ai-bias-with-real-world-examples/.
[68] Supra 35.

Healthcare payment systems also contain personal information that is a prime target for criminals who attempt to sell the data on the black market, commit medical fraud, or engage in identity theft.[69] Data breaches in the health industry are numerous, both from cyber-attacks and inside leaks. Recently, Change Healthcare was the target of a ransomware attack that led to widespread disruptions, affecting thousands of medical practices, hospitals, pharmacies, and more.[70]

One concern regarding privacy is who has access to, use of, and control of the data. Training AI requires large datasets from various patient types and locations, encouraging developers to gather as much data as possible. Many AI advances are undertaken by the private sector via large technology companies outside the confines of a medical system, so patient data is transferred to private firms via information sharing agreements. Some of these partnerships—such as the partnership between DeepMind and the Royal Free London National Health Service Trust in 2016—have met sharp criticism.

Among the leading concerns were that patients did not maintain control over the use of their data and that the data was transferred to another company once DeepMind was acquired.[71] Some patients have filed violation-of-privacy lawsuits due to data-sharing between health systems and AI developers.[72] AI continues to pose a privacy risk even after the AI development phase.

For example, for generative AI tools that record and summarize interactions between the provider and patient, the data is sometimes transferred to an AI company for processing rather than remaining within the health system network.

Depending on what organization controls it, health data may be protected by the Health Insurance Portability and Accountability Act of 1996 (HIPAA) and its implementing regulations and other federal and state laws.[73] HHS has developed two rules under HIPAA to address these concerns: the Privacy Rule and the Security Rule.

[69] Department of Health and Human Services, “Health Industry Cybersecurity Practices: Managing Threats and Protecting Patients (HICP),” 28 Dec. 2018. https://asprtracie.hhs.gov/technical-resources/resource/6624/health-industry-cybersecurity-practices-managing-threats-and-protecting-patients.

[70] Tom Murphy, Change Healthcare cyberattack was due to a lack of multifactor authentication, UnitedHealth CEO says, Associated Press News 1 May 2024, https://apnews.com/article/change-healthcare-cyberattack-unitedhealth-senate-9e2fff70ce4f93566043210bdd347a1f.

[71] Blake Murdoch, Privacy and artificial intelligence: challenges for protecting health information in a new era, BMC Med. Ethics 2021, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8442400/.

[72] I. Glenn Cohen & Michelle M. Mello, Big data, big tech, and protecting patient privacy, JAMA, 9 Aug. 2019, https://jamanetwork.com/journals/jama/fullarticle/2748399.

[73] The Health Insurance Portability and Accountability Act (HIPAA) of 1996 and its implementing regulations, the Privacy and Security Rules, protect individually identifiable health information that is used within the patient and provider relationship. The HIPAA Security Rule establishes national standards to protect individuals’ electronic PHI that is created, received, used, or maintained by a covered entity. 45 C.F.R. pt. 160 and pt. 164, subpts. A and C.

The HIPAA Privacy Rule contains several requirements for protecting the privacy and anonymity of health data and requirements to de-identify protected health information. Concerns remain about the limitations of HIPAA, including when health information or entities controlling this information are not covered by HIPAA protections.[74]

The HIPAA Security Rule establishes a national set of security standards for protecting certain health information held or transferred in electronic form.[75] Specifically, the rule requires physical safeguards such as restricting access to offices, technical safeguards such as computer logins and encryption, administrative safeguards such as strict policies and procedures, and a training program. These rules may need to be updated to meet challenges created by AI systems deployed in health contexts.

Interoperability

While AI-enabled tools have the potential to change the healthcare landscape, they must be able to integrate with healthcare systems, including EHR systems. Adding to this challenge is the fact that medical systems may use different EHR systems or methods of collecting and storing data.

For many years, healthcare systems have struggled with interoperability between EHRs. Even when different health systems use the same EHR, they may not share a common data storage repository, making information exchange difficult.[76] AI tools that cannot connect with all relevant systems could stifle their adoption and use of these tools. EHR vendors may also avoid facilitating data sharing.

Two vendors—Epic Systems and Oracle Cerner—account for more than half of the EHR market. Oracle Cerner has integrated generative AI into its EHR and has also included a digital assistant tool that automates notetaking, schedules appointments, and automatically responds to specific patient questions.[77] Epic Systems recently partnered with Microsoft to integrate OpenAI into their EHR.[78]

Liability

There is limited legal and ethical guidance regarding accountability when AI produces incorrect diagnoses or harmful recommendations. In 2024, the Federation of State Medical Boards released guidance stating that doctors who use AI in clinical decision support have accepted responsibility for their responses to AI recommendations.[79]

[74] Thomas Germain. "Guess What? HIPAA Isn't a Medical Privacy Law." Consumer Reports, June 2022, https://www.consumerreports.org/health/health-privacy/guess-what-hipaa-isnt-a-medical-privacy-law-a2469399940/.
[75] U.S. Department of Health and Human Services (HHS), The Security Rule, https://www.hhs.gov/hipaa/for-professionals/security/index.html.
[76] Supra 26.
[77] Marium M. Raza, Kaushik P. Venkatesh & Joseph C. Kvedar, Generative AI and large language models in healthcare pathways to implementation, npj Digital Medicine, 2024, https://www.nature.com/articles/s41746-023-00988-4.
[78] Id.
[79] Federation of State Medical Boards, Navigating the responsible and ethical incorporation of artificial intelligence into clinical practice, April 2024, https://www.fsmb.org/siteassets/advocacy/policies/incorporation-of-ai-into-practice.pdf.

Determining liability becomes increasingly complex as multiple parties become involved in developing and deploying an AI system. The contours of a liability policy may also affect clinical decision-making. For example, the risk of penalties may change depending on whether a provider relies on their judgment or defers to an algorithm.

In April 2024, the HHS Office of Civil Rights (OCR) finalized its rule addressing AI liability as part of its nondiscrimination rule under Section 1557 of the Affordable Care Act.[80] The rule aims to prevent discrimination when using "patient support decision tools," including AI tools. The rule obliges providers to make efforts to understand the foundation of the tools they use and whether those tools might contribute to discrimination. This squarely places the responsibility for some AI-related actions on healthcare providers rather than AI developers.

[80] U.S. Department of Health and Human Services (HHS), Section 1557 of the Patient Protection and Affordable Care Act, https://www.hhs.gov/civil-rights/for-individuals/section-1557/index.html.

Key Findings

AI's use in healthcare can potentially reduce administrative burdens and speed up drug development and clinical diagnosis.

When used appropriately, these uses of AI could lead to increased efficiency, better patient care, and improved health outcomes.

The lack of ubiquitous, uniform standards for medical data and algorithms impedes system interoperability and data sharing.

If AI tools cannot easily connect with all relevant medical systems, their adoption and use could be impeded.

Recommendations

Recommendation: Encourage the practices needed to ensure AI in healthcare is safe, transparent, and effective.

Policymakers should promote collaboration among developers, providers, and regulators in developing and adopting AI technologies in healthcare where appropriate and beneficial. For example, through workshops and conferences, policymakers could convene multidisciplinary experts to design and develop these technologies.[81] Policymakers could also develop or expand high-quality data access mechanisms that ensure the protection of patient data. Examples include voluntary standards for collecting and sharing data, creating data commons, and using incentives to encourage data sharing of high-quality data held by public or private actors.[82] More specifically, Congress should continue to monitor the use of predictive technologies to approve or deny care and coverage and conduct oversight accordingly.

Recommendation: Maintain robust support for healthcare research related to AI.

Sustained, strategic investments in research and development will be critical to maintaining U.S. leadership in AI across disciplines and use cases, especially in sectors that stand to benefit significantly from this technology. The research supported by the National Institutes of Health (NIH) has the potential to enable improvements in all of the healthcare applications discussed in this chapter. For more information on recommendations to improve AI-related skills across sectors, please see the Research, Development, and Standards chapter of this report.

Recommendation: Create incentives and guidance to encourage risk management of AI technologies in healthcare across various deployment conditions to support AI adoption and improve privacy, enhance security, and prevent disparate health outcomes.[83]

To promote the responsible use of AI systems in the healthcare sector, stakeholders would benefit from standardized testing and voluntary guidelines that support the evaluation of AI technologies, promote interoperability and data quality, and help covered entities meet their legal requirements under HIPAA. This includes using de-identification techniques and privacy-enhancing technologies to protect patient privacy.

[81] Supra 26.
[82] Supra 35.; See also Supra 26.
[83] Supra 26.

For example, the Department of Commerce, through its work developing general standards for AI risk management and evaluation, could work with HHS and relevant stakeholders to establish best practices (such as standards for data and algorithms) to facilitate the development, implementation, and use of AI technologies.[84]

One critical area for improved guidance or regulation is industry post-market surveillance and self-validation. Similar to pre-market review conditions, manufacturers have the responsibility to monitor the post-market performance and safety of their regulated medical devices. This responsibility may include actions such as implementing robust quality management systems, conducting post-market surveillance studies, monitoring user feedback, and building a mechanism to report significant events. Congress should explore whether the current laws and regulations need to be enhanced to help the FDA's post-market evaluation process ensure that AI technologies in healthcare are continually and sufficiently monitored for safety, efficacy, and reliability.

Recommendation: Support the development of standards for liability related to AI issues.

Limited guidance exists on constructing legal and ethical frameworks for determining who bears responsibility when AI models produce incorrect diagnoses or make erroneous and harmful diagnostic recommendations. Currently, most providers are expected to use AI tools as supplementary devices while still relying on their own judgments, thus placing liability on the providers themselves. As AI's use continues to increase in everything from EHRs to transcription services to diagnosis, Congress should examine liability laws to ensure patients are protected.

Recommendation: Support appropriate payment mechanisms without stifling innovation.

CMS calculates reimbursements by accounting for physician time, acuity of care, and practice expense. Considering that AI tools streamline these practices and reduce time spent on services, current payment mechanisms cannot adequately reimburse these tools. Certainly, there will be no "one size fits all" reimbursement policy for every AI technology, and developing appropriate payment mechanisms requires recognition of varying kinds of technology and clinical settings. For example, many AI technologies may fit into existing benefit categories or facility fees. Congress should continue to evaluate emerging technologies to ensure Medicare benefits adequately recognize appropriate AI-related medical technologies.

[84] Supra 35.; See also: Supra 11.

FINANCIAL SERVICES

Background

The financial services sector has employed artificial intelligence technologies for decades. Despite this experience, the latest advancements in AI have brought several issues to the forefront. AI can be used by financial regulators, the firms they regulate, and malicious third parties who could compromise the integrity of the financial services system. The numerous uses of AI in the financial services sector merit consideration of the range of benefits and potential risks that AI technology poses, as well as the hurdles that impede the adoption of AI technology.

Given the rapid change occurring within the financial services and housing industries, it is critical to understand the latest technological developments and how market participants are utilizing them. Congress and financial regulators must ensure that financial service regulations take into account the potential benefits and risks associated with the use of various AI technologies to protect consumers and maintain market integrity.

AI has the potential to transform the financial services sector, which is highly reliant on digital technology. A regulatory framework that supports the responsible deployment of AI is essential to foster innovation while safeguarding consumers and maintaining market stability. By prioritizing efficiency, customer experience, and financial inclusion, AI can improve accessibility and operational effectiveness within the sector. However, realizing these benefits will require both activity-based regulation and collaboration between public and private entities. Further, improper implementation of AI may worsen existing challenges, such as cybersecurity risks and erroneous system outputs, thereby undermining trust and stability within the financial system.

Previous Committee Work

In the 116th and 117th Congresses, the Financial Services Committee established a Committee Task Force on Artificial Intelligence (“committee task force”) to better understand the opportunities and challenges surrounding the use of AI and technology in financial services.[1] The committee task force held numerous hearings on the policy implications of the use of AI, including how AI can be used to assess risks, predict outcomes, and allocate capital across the financial system in a more efficient manner than traditional human assessments.

In January 2024, Chairman Patrick McHenry and Ranking Member Maxine Waters established the bipartisan Financial Services Committee AI Working Group (Working Group) comprised of twelve members.[2]

The House Financial Services Committee held six roundtables and an off-site visit to the Massachusetts Institute of Technology (MIT), all focused on identifying existing and growing AI use cases across the financial services and housing industries.

History of AI in Financial Services

Over the past several decades, the field of AI has experienced significant growth and investment and has been highly pertinent to financial services. In the 1980s, AI experienced renewed interest and exploration, particularly in Japan, the United Kingdom, and the United States.[3] From the early to mid-80s, the popularity of personal computers and computer hardware exploded following the release of the Apple II, TRS-80 Model I, Commodore PET, and later through the release of Lotus 1-2-3 and the Apple Macintosh.[4]

During this period, the financial services industry began considering how to automate decisions with the help of the prevalent AI technologies. General Electric (GE) used rules-based systems and heuristics to analyze the quality of commercial loans.[5]

[1] Some of the content in these sections are drawn from the bipartisan work of the U.S. House of Representatives Committee on Financial Services, and their July 18, 2024 report “AI Innovation Explored: Insights into AI Applications in Financial Services and Housing,” https://financialservices.house.gov/uploadedfiles/bipartisan_working_group_on_ai_staff_report.pdf.

[2] Financial Services Committee, "McHenry, Waters Announce Creation of Bipartisan AI Working Group." U.S. House of Representatives, 11 Jan, 2024, https://financialservices.house.gov/news/documentsingle.aspx?DocumentID=409108.

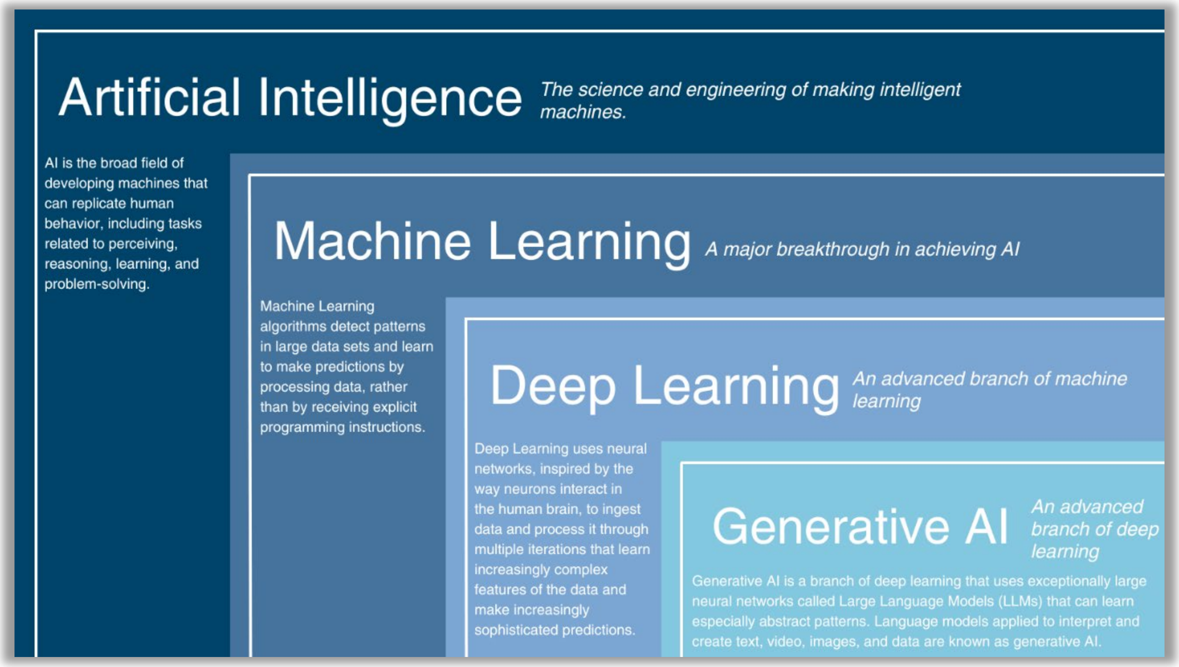
[3] Buchanan, Bonnie G. Artificial Intelligence in Finance: Turing Report. Turing Institute, April 2019, https://www.turing.ac.uk/sites/default/files/2019-04/artificial_intelligence_in_finance_-_turing_report_0.pdf.

[4] Giacaglia, Giuliano. “Making Things Think: How AI and Deep Learning Power the Products We Use.” Holloway, November 2022, www.holloway.com/g/making-things-think/sections/the-ai-boom-19801987.

[5] Peter Duchessi, Hany Shawky, and John P. Seagle, "A Knowledge-Engineered System for Commercial Loan Decisions," Financial Management vol. 17, no. 3 1988: 57-65. https://www.jstor.org/stable/3666072.

In 1989, the Fair Isaac Corporation developed the FICO credit-scoring algorithm, which used a combination of factors, including payment history, credit utilization, and length of credit history, to assess the borrowers' creditworthiness.[6]

During the 1980s, Edward Feigenbaum, a computer science professor at Stanford University, developed the concept of "expert systems," also known as "knowledge systems," which focused on mimicking human reasoning within a specific domain.[7] This technique enabled companies to make tailored financial plans for consumers, as well as "investment planning, debt planning, retirement planning, education planning, life-insurance planning, budget recommendations, and income tax planning."[8] Wall Street began using these expert systems through program trading, also known as algorithmic trading, to automatically execute trades at high speeds based on predetermined conditions and without human intervention.[9]



Artificial Intelligence *The science and engineering of making intelligent machines.*

AI is the broad field of developing machines that can replicate human behavior, including tasks related to perceiving, reasoning, learning, and problem-solving.

Machine Learning *A major breakthrough in achieving AI*

Machine Learning algorithms detect patterns in large data sets and learn to make predictions by processing data, rather than by receiving explicit programming instructions.

Deep Learning *An advanced branch of machine learning*

Deep Learning uses neural networks, inspired by the way neurons interact in the human brain, to ingest data and process it through multiple iterations that learn increasingly complex features of the data and make increasingly sophisticated predictions.

Generative AI *An advanced branch of deep learning*

Generative AI is a branch of deep learning that uses exceptionally large neural networks called Large Language Models (LLMs) that can learn especially abstract patterns. Language models applied to interpret and create text, video, images, and data are known as generative AI.

Source: McKinsey: What is AI?

[6] Kaufman, Rob. "The History of the FICO® Score." myFICO, 25 July 2018, www.myfico.com/credit-education/blog/history-of-the-fico-score.; see also: Ebert, Sven, and Philipp Immenkötter. "Machine Learning in Financial Markets: Come to Stay." Flossbach von Storch Research Institute, 27 Feb. 2023, www.flossbachvonstorch-researchinstitute.com/en/studies/machine-learning-in-financial-markets-come-to-stay/.; see also: Allen, Brian. "Navigating Artificial Intelligence in Banking." Bank Policy Institute, 8 April 2024, bpi.com/navigating-artificial-intelligence-in-banking/#:~:text=In%20fact%2C%20the%20integration%20of,variety%20of%20use%20cases%20since.

[7] Expert systems can be understood as "a computer program that, after having been properly instructed by a professional, is able to deduce information from a set of data and starting information. See: Brown, Carol E., et al. Expert Systems for Personal Financial Planning, The Institute of Certified Financial Planners, July 1990, prism.ucalgary.ca/server/api/core/bitstreams/726f43f5-d1a6-4984-982d-e8cd2300184d/content.

[8] Supra 3.

[9] In 1982, James Simons, a renowned mathematician and investor, founded the quantitative hedge fund Renaissance Technologies and in the late 1980s, the firm began to explore algorithmic trading. See: Veiga, Alex. "James Simons, Mathematician, Philanthropist and Hedge Fund Founder, Has Died." Associated Press News, 10 May 2024,

In 1982, the mathematician and investor James Simons founded Renaissance Technologies, a quantitative hedge fund that explored algorithmic trading in the late 1980s.[10] Through vast market data and pattern analytics, these algorithms can execute trading decisions at high speeds without human intervention.

Gradually, algorithmic trading became more popular among institutional investors and large trading firms due to benefits like faster execution time and reduced costs. On Monday, October 19, 1987, also known as "Black Monday," global stock exchanges plummeted, with the Dow Jones Industrial Average (DJIA) falling 22.6%[11]—an amount exacerbated by algorithmic trading.[12]

Between the late 1980s and the mid-1990s, AI experienced diminished interest and a lack of development. This period was caused by a variety of factors, including the limitations of these advanced systems and unmet expectations, which resulted in an unwillingness to fund AI projects.[13]

AI started to emerge again in the late 1990s with the development of internet search engines and better hardware.[14] The first internet banking solution was offered by the Stanford Federal Credit Union in 1994.[15]

The increases in digitized data and computing hardware performance both contributed to the growth of the next generation of AI.[16] With these advancements, AI experts began to focus more on ML and neural networks.[17] Neural networks attempt to mimic how living things process information and identify complex patterns.

This shift in the AI paradigm first occurred in the early 1990s when IBM developed Deep Blue, a chess-playing computer system that could search up to 200 million options per second. In 1996, Deep Blue defeated Garry Kasparov, a Russian grandmaster, in one of six games.[18]

apnews.com/article/james-simons-renaissance-technologies-simons-foundation-9f97b19939806f970bdaa09878e382da; see also: Miller, Rena S., and Gary Shorter. "High Frequency Trading: Overview of Recent Developments." Congressional Research Service, 2016, www.crs.gov/Reports/R44443.

[10] Id.

[11] Bernhardt, Donald, and Marshall Eckblad. "Stock Market Crash of 1987." Federal Reserve History, 16 Oct. 2013, www.federalreservehistory.org/essays/stock-market-crash-of-1987.

[12] Id.; Ruder, David S. "The October Market Break: A Stimulant To United States-Japanese Cooperative Securities Regulation." U.S. Securities and Trade Commission, 2006, www.sec.gov/news/speech/1988/021888ruder.pdf.

[13] Werner, John. "3 Lessons Learned From The Second AI Winter." Forbes, 9 April 2024, www.forbes.com/sites/johnwerner/2024/04/09/three-lessons-learned-from-the-second-ai-winter/?sh=56e8b0b9c3cd.

[14] Id.

[15] "About Us." Stanford Federal Credit Union, 26 Sept. 2019, www.sfcu.org/about/.

[16] Supra 13.

[17] Id.

[18] "Deep Blue." IBM, 8 Dec. 2023, https://www.ibm.com/history/deep-blue.; See also: Yao, Deborah. "25 Years Ago Today: How Deep Blue vs. Kasparov Changed AI Forever." IOT World Today, 11 May 2022, www.iotworldtoday.com/iiot/25-years-ago-today-how-deep-blue-vs-kasparov-changed-ai-forever.

Deep Blue's success demonstrated the capabilities of AI systems and inspired a new wave of research to create supercomputers that could mine data, conduct risk analysis in finance, and more.[19]

In addition, the then-newly established Financial Crimes Enforcement Network (FinCEN) employed a unique application of AI technology to flag suspicious activity. This AI application assisted analysts in searching internal database reports to combat money laundering.[20] Banks, payment processors, and core providers also implemented AI fraud detection systems in the following years.

In the 2010s, advances in graphical processing units (GPUs) enabled numerous layers of neural networks to be trained more efficiently on large amounts of data. This method, called "deep learning," enabled neural networks to recognize much more complex patterns and continually learn from data in ways somewhat analogous to human learning.[21] For example, neural networks enhance foreign exchange trading by utilizing simulated data from various market conditions to select the best order placement and execution style that minimizes market impact.[22]

The most recent advance in AI technology, generative AI, enables systems to respond to natural language inquiries and generate poems, essays, document summaries, and other high-quality text. The first version of GPT was launched by OpenAI in 2018 and was trained on 40 gigabytes of internet data.[23] In 2021, OpenAI created DALL-E, an ML model trained on internet data that generates images from text descriptions provided by the user.[24] OpenAI's launch of ChatGPT in 2022 led to "a rare moment when an AI/ML technology became directly accessible by the broad public,"[25] as well as significant new interest and investment in AI technology by a broad range of sectors.

Generative AI models hold enormous potential and can streamline the regulatory examination process of investigating suspected market manipulation and insider trading activity. For example, generative AI systems can be used to produce a consolidated table of the company's regulatory filings, news articles, their associated tone or sentiment, and other factors that may impact any given security.[26]

[19] Id.
[20] Senator, T E, et al. "FinCEN Artificial Intelligence System: Identifying Potential Money Laundering From Reports of Large Cash Transactions." U.S. Department of Justice, 1995, www.ojp.gov/ncjrs/virtual-library/abstracts/fincen-artificial-intelligence-system-identifying-potential-money.
[21] Rigano, Christopher. "A Brief History of Artificial Intelligence." National Institute of Justice, 30 Sept. 2018, nij.ojp.gov/topics/articles/brief-history-artificial-intelligence.; See also: Gleyzer, Sergei, et al. "The Rise of Deep Learning." CERN Courier, 9 July 2018, cerncourier.com/a/the-rise-of-deep-learning/.
[22] Editorial team Vention. "Neural Networks in Financial Trading and Analysis." Vention, 7 March 2023, https://ventionteams.com/blog/neural-networks-in-trading.
[23] Marr, Bernard. "A Short History of Chatgpt: How We Got to Where We Are Today." Forbes, 19 May 2023, www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/.
[24] Ramesh, Aditya, et al. "Dall·E: Creating Images from Text." OpenAI, 5 Jan. 2021, https://openai.com/index/dall-e/.
[25]Tierno, Paul. "Artificial Intelligence and Machine Learning in Financial Services." Congressional Research Service, 3 April 2024, crsreports.congress.gov/product/pdf/R/R47997.
[26] "Nasdaq to Enhance Global Market Surveillance Offering with Generative AI." Nasdaq, 15 May 2024, www.nasdaq.com/press-release/nasdaq-to-enhance-global-market-surveillance-offering-with-generative-ai-2024-05-15.

Generative AI technologies differ from traditional AI tools, such as predictive ML models, which have been used in the housing and financial services sectors for decades.[27] A 2022 survey found that over 75% of financial services companies employ at least one use case of AI-related computing.[28] However, new AI capabilities present new challenges and risks of using AI inappropriately. These risks can affect the customers of financial services firms, other groups, and potentially the wider financial services sector.

AI Issues in Financial Services

AI can be used by financial regulators, financial services firms, and malicious actors seeking to attack financial services firms, their customers, or the financial services system itself. Seven key issues are addressed below.

Financial Regulators' Use of AI for Supervision and Regulation

Supervisory technology (SupTech) refers to using innovative technology, such as AI, deployed by regulators to support their supervisory, rulemaking, and enforcement efforts.[29] Through SupTech, regulators have improved their supervisory capabilities, helped financial institutions meet regulatory requirements,[30] and supported their efforts to collect and analyze data.[31] Regulatory technology (RegTech) typically refers to automation for regulatory, compliance, and data reporting obligations for financial firms and other regulated entities.[32]

It is important to understand the extent to which SupTech keeps pace with advances in the technology used by businesses, the extent to which regulators deploy AI to enhance their oversight responsibilities, whether regulators have the resources necessary to oversee the rapid adoption of AI among the entities they regulate, and the challenges in hiring and recruitment programs to ensure there are sufficient staff members with technology backgrounds to help regulators both utilize AI and monitor its developments.

[27] Supra 25.

[28] NVIDIA Corporation. "State of AI in Financial Services." *NVIDIA*, 25 Jan. 2022, www.nvidia.com/content/dam/en-zz/Solutions/industries/finance/ai-financial-services-report-2022/fsi-survey-report-2022-web-1.pdf.

[29] "What Is SupTech? (A Market Overview of Supervisory Technology)." Stellex, 4 July 2018, www.stellexgroup.com/blog/suptech-supervisory-technology.

[30] "The Use of Supervisory and Regulatory Technology by Authorities and Regulated Institutions." Financial Stability Board, 9 Oct. 2020, www.fsb.org/uploads/P091020.pdf.; see also: Innes, Kirsty, and Rosie Beacon. "Government by Algorithm: The Myths, Challenges and Opportunities." Tony Blair Institute for Global Change (TBI), 25 Jan. 2021, institute.global/insights/tech-and-digitalisation/government-algorithm-myths-challenges-and-opportunities.; see also: Tricentis Staff. "AI in Software Testing: Rule-Based Testing vs. Learning Systems." Tricentis, 8 Jan. 2019, www.tricentis.com/learn/ai-approaches-rule-based-testing-vs-learning.; and Brynjolfsson, Erik, et al. "What Can Machines Learn, and What Does It Mean for Occupations and the Economy?" American Economic Association, May 2018, www.aeaweb.org/articles?id=10.1257%2Fpandp.20181019.

[31] Supra 29.

[32] Regtech has been used for about ten years and has played an important role in assisting institutions with their national security and illicit finance programs, including detecting, preventing, and reporting illicit financial activities. See: Supra 25; see also: Broeders, Dirk, and Jermy Prenio. "Innovative Technology in Financial Supervision (Suptech) – the Experience of Early Users." Financial Stability Institute, July 2018, www.bis.org/fsi/publ/insights9.pdf.

The Working Group learned the Treasury Department is in discussions with large technology companies to create a pipeline of qualified individuals into government service. The Consumer Financial Protection Bureau (CFPB) is building interdisciplinary teams and augmenting technical expertise and talent in ML, data science, and analytics.

Other agencies cited challenges working within current funding levels and attracting staff with technological backgrounds. The Federal Reserve emphasized that its use of AI would not replace staff. Instead, its use of AI is intended to enhance the staff's abilities and allow them to focus on things that better use their time and expertise.

The use of AI does not absolve regulated entities from complying with applicable laws and regulations, including anti-discrimination laws and consumer protection laws.

On the contrary, regulated firms are expected to follow all laws in a technology-neutral manner. For example, if a lender cannot explain an adverse outcome based on its use of AI, the CFPB considers that to be a violation of the Equal Credit Opportunity Act (ECOA). Regulators can leverage their oversight and enforcement authorities to ensure existing obligations are met and examine alternative compliance processes where appropriate.

> **“*Regulated firms are expected to follow all laws in a technology-neutral manner.***

Regulators can also use AI to identify non-compliance with regulations. The Treasury Department's Office of Terrorism and Financial Intelligence emphasized that the benefits of AI integration in the financial system are readily apparent in anti-money laundering (AML), countering the financing of terrorism (CFT), and sanctions compliance. When properly calibrated, this technology could streamline efficiency in meeting compliance obligations and monitoring transactions to identify suspicious activity.

Tasked with reviewing the financial technology expertise of the prudential regulators[33] and the CFPB, GAO released a report[34] that examines (1) the technological skills or expertise related to financial technology policymaking and oversight that regulators' staff possess; (2) regulators' workforce planning processes to ensure their staff are sufficiently knowledgeable to engage in policymaking and oversight related to FinTech products and services, as well as the extent to which those processes are consistent with leading practices; (3) how regulators address innovation in FinTech and measure the results of innovation efforts; and (4) how regulators use technology to improve their supervisory capabilities.

[33] Prudential regulators include the Federal Reserve, the Office of the Comptroller of the Currency, and the Federal Deposit Insurance Corp. See: Stackhouse, Julie L. "Why Are There So Many Bank Regulators?" Federal Reserve Bank of St. Louis, 24 April 2017, https://www.stlouisfed.org/on-the-economy/2017/april/why-many-bank-regulators.

[34] U.S Government Accountability Office. "Agencies Can Better Support Workforce Expertise and Measure the Performance of Innovation Offices." U.S. Government Accountability Office, 6 Sept. 2023, https://www.gao.gov/assets/gao-23-106168.pdf.

It is critical that regulators keep up with rapid innovation and utilize new technologies to enhance the efficiency of federal programs and improve the monitoring of financial markets and institutions.

Regulators must continue to foster the development of and bolster oversight of new products and services in financial services. As one example, regulatory “sandboxes” could provide regulators with controlled experimentation in AI applications, enabling them to observe their impacts and adapt regulations accordingly.

While risks accompany the use of AI, as with any emerging technology, financial regulators can continue to ensure market participants comply with existing laws. At the same time, they should focus on fostering an environment where firms can reap the benefits of AI technology.

Regulators must also examine the associated benefits and risks of AI in the financial services and housing sectors, including the risks of improper decision-making, and continue to ensure consumer and investor protection and market integrity. Finally, it is essential to continually assess any potential legislative or regulatory gaps or limitations concerning AI applications in the financial services and housing industries.

Consumer Data Privacy and AI Models

Because AI technologies rely on large amounts of data, data privacy has become a prominent issue. Both large and small financial service entities utilize large language models (LLMs) or other models that are trained on significant amounts of consumer data. While collecting, using, and sharing consumer financial data is necessary for financial firms to improve how their services are delivered, it also creates an area of risk for policymakers to address. The high volumes and wide range of data used by AI, especially generative AI, underscore the importance of controls around the quality, security, and privacy of data, and the importance of safeguarding consumer data.

Several harms can arise from inadequate and improperly sourced data. One such harm is the significant risk of revealing confidential or personally identifiable information incorporated into the data used to train AI models. This risk is only exacerbated if the model is trained in an environment without adequate security and privacy controls.

Another data security concern is vulnerability in intellectual property protections. A third party might be capable of reverse engineering the data on which a model was trained to acquire proprietary data or curated data sets. For this reason, federal agencies should only collect consumer financial data when necessary, especially if the data is stored in a centralized location that is vulnerable to potential cyberattacks.

Data privacy concerns will only grow as data becomes more widely used in training AI. Strengthening data privacy requirements, especially in light of advancements in generative AI, is an important policy concern.

Congress will continue to review and work to update federal laws that apply to financial institutions and financial data, like the Gramm-Leach-Bliley Act (GLBA)[35] and the Fair Credit Reporting Act (FCRA)[36], to strengthen data privacy protections.

Balancing the protection of personal financial information with the benefits of AI-driven advancements and innovation is essential to maintain consumer trust and foster a competitive, innovative financial services sector.

Potential Risks of Using AI in Decision-Making

Because AI models are data-driven, problems with training data can lead to models not performing as designed or expected. Specifically, training data can be imbalanced, incomplete, or otherwise limited in ways that are subtle and difficult to detect, including by reflecting historical patterns. These flaws in training data can fail to account for important nuances or misrepresent particular groups or types of decisions.

It is essential that training data is representative and high-quality. If not, AI technologies may reproduce or even exacerbate biased or discriminatory outcomes due to the use of biased data inputs, particularly in light of historical segregation and discrimination in the housing sector.

If the outputs of a flawed model are used in decisions by financial services firms, there are risks of harming customers or other parties.[37] These harms could have wide-ranging effects beyond a few customers; they may impact large groups and even risk market instability.

> **“*The risks and liability of using AI in decision-making can be reduced through increased transparency in the development and use of AI.***

Not only can AI models produce flawed outputs, but those flaws can be skewed or uneven in a way that disproportionately affects one or more groups of people. A financial services firm that uses such flawed outputs in its decisions risks engaging in bias and discrimination against protected classes such as race, sex, or veteran status. The chance of producing such a skewed model is amplified when the training data reflects historical bias or does not adequately represent certain groups.[38]

[35] Office of the Federal Register, National Archives and Records Administration. Public Law 106 - 102 - Gramm-Leach-Bliley Act. U.S. Government Printing Office, 11 Nov. 1999, https://www.govinfo.gov/app/details/PLAW-106publ102.

[36] Liu, Henry, and Staff at the FTC. “Fair Credit Reporting Act.” Federal Trade Commission, 19 July 2013, www.ftc.gov/legal-library/browse/statutes/fair-credit-reporting-act.

[37] Financial Stability Board. “The Financial Stability Implications of Artificial Intelligence.” Financial Stability Board, 14 Nov. 2024, www.fsb.org/2024/11/the-financial-stability-implications-of-artificial-intelligence/#:~:text=AI%2Drelated%20vulnerabilities%20that%20stand,risk%2C%20data%20quality%20and%20governance.

[38] Nicoletti, Leonardo, and Dina Bass. “Humans Are Biased. Generative AI Is Even Worse.” Bloomberg, 9 June 2023, www.bloomberg.com/graphics/2023-generative-ai-bias/.

For example, AI can be used in various housing and insurance products and services. While this can provide new benefits and conveniences to consumers, it also presents challenges to fair housing, consumer protection, and other aspects of housing and insurance markets. Errors in the use of AI and other housing and property technology (PropTech), such as automated valuation models (AVMs), online housing platforms, tenant screening companies, and rent-setting companies, could lead to increased housing costs, discrimination, and other barriers to fair and affordable housing.[39]

Similarly, the use of AI in underwriting, credit evaluation, or mortgage approval could lead to incorrect decisions or decisions that discriminate against a protected class.

If AI risks are perceived as pervasive, they can undermine trust in AI-driven processes in the financial services sector.[40] Likewise, if there is widespread adoption of the same flawed AI model, multiple firms could make the same errors, resulting in herd-like behavior affecting and influencing large portions of capital markets or housing markets.[41]

Panelists on the Financial Services Committee Working Group noted that the risks of erroneous or harmful outputs were particularly salient with newer generative AI systems. As a result, some firms stated they were delaying using generative AI in critical areas of their operations until these problems were sufficiently addressed. Many market participants have also developed AI governance bodies within their organizations.

The risks and liability of using AI in decision-making can be reduced through increased transparency in the development and use of AI. Risks related to biased data and discriminatory AI models can be further mitigated by ensuring diverse groups are part of model engineering, development, testing, and deployment phases.

Market participants must be able to understand and explain the limitations of the model and how to correct errors to avoid compounding problems. AI outputs are often difficult to explain and interpret, which can increase the risk that improper decisions are not detected. Finally, keeping a human in the loop to help monitor and check for quality and accuracy in outputs and decisions can help.[42]

[39] United States, House of Representatives, Committee on Financial Services. “Waters Issues Statement on SEC’s Landmark AI Proposal to Protect Investors.” House Committee on Financial Services Democrats, 16 Nov. 2024, https://democrats-financialservices.house.gov/news/documentsingle.aspx?DocumentID=410767.
[40] Bhaskar Chakravorti. “AI’s Trust Problem”, Harvard Business Review, May 2024, https://hbr.org/2024/05/ais-trust-problem.
[41] United States, House of Representatives, Committee on Financial Services. “AI Innovation Explored: Insights into AI Applications in Financial Services and Housing.” House Committee on Financial Services, 18 July 2024, https://financialservices.house.gov/uploadedfiles/bipartisan_working_group_on_ai_staff_report.pdf.
[42] Holistic AI Team, “Human in the Loop AI: Keeping AI Aligned with Human Values.” Holistic AI, October 2024, https://www.holisticai.com/blog/human-in-the-loop-ai#:~:text=Benefits%20of%20Human%20in%20the,%2C%20ethical%2C%20and%20adaptable%20solutions.&text=Quality%20Control%3A%20Human%20intervention%20helps,that%20automated%20systems%20might%20overlook.

In summary, the risks of bias and other inaccuracies by AI can be reduced through appropriate awareness of the risks when creating and deploying AI systems, as well as routine monitoring and implementing other safeguards.

Use of AI in accessing financial products and services

Innovation in FinTech has introduced new ways to bank, lend, invest, trade, and conduct payments. ATMs, online banking, and online brokerage accounts were once considered novel financial technologies. They expanded access to the financial sector and are now the preferred methods for millions of Americans to access financial services.

Notwithstanding the potential risks discussed in the previous section, AI has the potential to increase access to services and become the norm for more Americans as it becomes woven into our financial activities.

AI technologies are already being deployed across the financial services sector in areas including fraud detection, underwriting, debt collection, customer onboarding, real estate, investment research, property management, and customer service. Continued adoption and further automation of services could result in significant cost reductions and greater access to financial services for more Americans.[43]

Additionally, in underbanked and underserved communities, AI-powered banking solutions can provide crucial financial services and help bridge the gap in financial inclusion. Alternative data underwriting, multilingual 24/7 automated call centers, and enhanced investment advice make it less expensive for institutions to offer more customers financial services with less infrastructure.[44]

AI can also streamline regulatory compliance, reducing the burden on financial institutions and allowing them to focus more on serving their customers.

Use of AI to enhance employee efficiency and productivity.

AI technology has the potential to drastically change the nature of work in the financial services sector by significantly enhancing employee efficiency and productivity. In our capital markets, AI is increasingly utilized for research and analysis, enabling financial professionals to process and interpret large amounts of data with unprecedented speed and accuracy.[45]

AI-driven algorithms can analyze market trends, enhance forecasting, and evaluate investment opportunities, allowing analysts and traders to make more informed decisions.

[43] United States, House of Representatives, Committee on Financial Services. “Letter to Secretary Yellen: Request for Information on Uses, Opportunities, and Risks of Artificial Intelligence in the Financial Services Sector.” House Committee on Financial Services, 14 Aug. 2024, https://financialservices.house.gov/uploadedfiles/2024-08-14_fsc_comment_letter_to_treasuryrfi__ai_final.pdf.

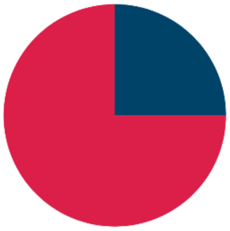
[44] LexisNexis Risk Solutions, “New LexisNexis Risk Solutions Report Reveals Financial Inclusion is Accelerating the Adoption of Alternative Data Across Financial Institutions.” February 2023, https://risk.lexisnexis.com/about-us/press-room/press-release/20230215-alternative-data-across-financial-institutions.

[45] Azati Team, “Real-Time Data Analysis: How AI is Transforming Financial Market Predictions.” Azati, June 2024, https://azati.ai/real-time-data-analysis/.

AI can accelerate the research process and increase the precision of market insights, thereby enhancing the overall productivity of employees in research roles.

Recent AI advances have also facilitated a major shift in housing products and services. AI tools streamline operations for property managers by automating tasks such as rent collection, maintenance scheduling, and tenant communication.

AI-powered platforms can analyze market conditions to optimize rental pricing, predict maintenance needs through data analytics, and improve tenant satisfaction by providing more responsive and personalized services. However, the rise in third-party tenant screening and rent-setting AI technologies can create risks for consumer access to fair and affordable housing.[46]



75%

OF FINANCIAL SERVICES COMPANIES EMPLOY AT LEAST ONE USE CASE OF AI-RELATED COMPUTING.

In customer service, AI chatbots and virtual assistants transform how financial institutions interact with clients. These AI systems can handle a wide range of inquiries, from account balances to transaction histories, freeing customer service representatives to deal with more complex issues that require human intervention and skill. This improves response times and customer satisfaction and allows employees to focus on higher-value tasks.[47]

As AI continues to be integrated into applications, the role of employees is evolving from task execution to strategic oversight and decision-making. This shift enhances productivity and increases job satisfaction by allowing employees to focus on more complex aspects of their jobs that require their expertise. Financial firms that effectively leverage AI could enjoy significant improvements in both operational efficiency and the quality of service they provide to clients.

AI is both leveraged by malicious actors to compromise financial firms and by financial firms to respond to threats.

Many risk assessments and monitoring systems used by financial institutions today are still rules-based, meaning they look for defined activities or anomalies and only produce alerts if established patterns are recognized. Integrating AI into these financial and cybercrime monitoring systems would detect unusual or suspicious activities in transactions utilizing large data sets, behavioral analysis, and other means.

[46] U.S. Department of Housing and Urban Development, "HUD Issues Fair Housing Act Guidance on Applications of Artificial Intelligence." 2 May 2024, https://www.hud.gov/press/press_releases_media_advisories/hud_no_24_098.

[47] Anca Dunavete, "AI in Workplace Technology: Improving Productivity with Chatbots and Virtual Assistants." Yarooms, May 2024, https://www.yarooms.com/blog/ai-in-workplace-technology.

These AI systems would automate tasks that are either difficult or impossible for humans to perform without such augmentation. AI-driven models could enable a transaction monitoring system, for example, to continuously learn from prior processed transactions and re-train the model to identify anomalous activity for human review.

Financial institutions also use AI to verify and authenticate their customers' identities. For instance, financial institutions can use AI to analyze the "liveness" of a customer's voice or picture to determine whether it is a real human. AI can also assist in reducing the likelihood that customer activity is flagged as suspicious when it should not. This, in turn, bolsters firms' efficiency in Bank Secrecy Act[48] AML compliance. It also strengthens fraud, sanctions, and cyber intrusion screening and potentially improves financial services access for those who might otherwise be "de-risked" from the banking system due to incomplete or inaccurate risk assessments.

There are additional national security-focused benefits of AI for financial institutions. AI can enable small, community-based financial institutions, which typically have less robust in-house IT and cybersecurity competencies than the largest multinational firms, to detect phishing, fraudulent identity, and other tactics to penetrate or fool system defenses. Additionally, AI can streamline investigative processes by aggregating key information for analysts to complete investigations more efficiently and in a manner that facilitates legal and compliance reviews.

One Working Group participant provided examples of how utilizing AI to detect fraudsters—who themselves may be using AI—is the most effective way to determine that a voice is authentic rather than computer-generated.

Importantly, a company must have the knowledge and ability to explain to examiners and others how AI is being used, its capabilities and deficiencies, and the security environment surrounding it.[49]

Financial institutions must constantly assess the efficacy of AI deployment and ensure that AI systems are adding value to their risk management plans, especially in light of increased AI-enabled attacks and fraud. It is unclear how many financial institutions have truly incorporated AI into their compliance programs for financial and cybercrime prevention and detection. The innovations that AI could offer include better transaction monitoring systems, more accurate customer risk assessments, improved and automated compliance reporting, and risk-based case management for financial institution investigations.

AI, and generative AI in particular, has also armed criminals with a new tool that can increase the frequency and sophistication of attacks against or through the financial services sector—both domestically and globally.

[48] "The Bank Secrecy Act." Financial Crimes Enforcement Network, 1 March 2011, https://www.fincen.gov/resources/statutes-and-regulations/bank-secrecy-act.
[49] Steve Ellis, "How Scammers Can Use Your Voice Against You." Office1, April 2024, https://www.office1.com/blog/voice-cloning-scam-protection.

To counter these threats, financial services regulators and agencies can work with foreign jurisdictions to understand cross-border applications of AI in financial services and ensure American principles are at the forefront of the discussion.

This is especially important given the efforts by authoritarian governments like the Chinese Communist Party to use AI to spread repression, curb democracy, and further their anti-American interests.[50]

Finally, financial services firms may be better able to address illicit actors' use of AI if they can understand and harness advanced cybersecurity technology, including AI-enabled defenses, to defend themselves and their customers. With the use of AI, firms can allocate their resources more effectively to meet modern threats from malicious actors.

Barriers to competition between small and larger financial firms seeking to utilize AI

Integrating AI into the financial services sector presents both opportunities and challenges, particularly when it comes to the competitive dynamics between small and larger financial firms. While AI technologies offer the potential to transform financial services, disparities in resources, access to data, and technical expertise create significant barriers to competition between financial firms of different sizes.[51]

Large financial institutions have substantial resources to invest in AI development, including the ability to build proprietary AI models, hire specialized technical talent, and integrate advanced technologies across their operations. In contrast, smaller firms often lack the financial and other resources to develop, acquire, and scale sophisticated AI tools. This disparity creates a competitive advantage, allowing larger firms to leverage AI more effectively to enhance their services, optimize operations, and reduce costs.[52]

AI models thrive on vast amounts of high-quality data. Larger financial institutions typically have access to more extensive datasets, enabling them to train and refine AI models more effectively. Smaller firms may struggle to access the volume and diversity of data needed to develop robust AI applications. This data gap further entrenches the competitive advantage of larger firms, as they can use their data to create more accurate and reliable AI-driven insights and services, generating more data for future use.[53]

[50] Adrian Shahbaz, "The Rise of Digital Authoritariansim." Freedom House, https://freedomhouse.org/report/freedom-net/2018/rise-digital-authoritarianism.

[51] El Bachir Boukherouaa, et al., "Powering the Digital Economy: Opportunities and Risks of Artificial Intelligence in Finance." International Monetary Fund, 2021, https://www.elibrary.imf.org/view/journals/087/2021/024/article-A001-en.xml#:~:text=In%20the%20financial%20sector%2C%20advances,automated%20processes%20(Box%202).

[52] McKinsey & Company, "Building the AI bank of the future." McKinsey & Company, May 2021, https://www.mckinsey.com/~/media/mckinsey/industries/financial%20services/our%20insights/building%20the%20ai%20bank%20of%20the%20future/building-the-ai-bank-of-the-future.pdf.

[53] Id.

The development and deployment of AI technologies require specialized technical expertise, which is more readily available to larger firms with the resources to attract top talent. Smaller firms may face challenges in hiring and retaining AI specialists, limiting their ability to innovate and compete in the rapidly evolving AI landscape.

Smaller firms can acquire relevant expertise and other benefits from bank-fintech partnerships. To ensure the long-term viability of such partnerships, third-party risk management guidance should recognize the increased importance of these relationships to smaller firms.

By leveraging these partnerships, entrepreneurs and the innovative products and services they develop could be supported by processes that foster the next generation of entrepreneurs.

To remain competitive, smaller financial firms often rely on third-party AI solutions. While these partnerships allow smaller firms to access advanced technologies without significant in-house development, they also introduce dependencies and potential vulnerabilities. For instance, smaller firms may be constrained by the limitations of the third-party solutions they adopt, or they may face increased risks related to data privacy, cybersecurity, and regulatory compliance from tools they do not completely control.

The regulatory environment is critical in shaping competition within the financial services sector. Current regulations may inadvertently favor larger firms with the resources to navigate complex regulatory landscapes and implement compliance measures effectively. Smaller firms with limited resources may find it more challenging to meet regulatory requirements, particularly when adopting new AI technologies.

Key Findings

AI presents an opportunity to transform the financial services sector.

AI can reduce the price and improve the quality of products and services offered by financial services firms. AI adoption can also enhance the efficiency and productivity of employees in financial services firms.

Data quality and data security are paramount in financial service AI models.

Training data for AI models must be representative and high-quality—if not, the models may be skewed and give erroneous outputs. Models must be analyzed and monitored for bias and other adverse effects, including compliance with existing anti-discrimination laws.

AI can expand access to financial products and services.

In underbanked and underserved communities, AI-powered solutions can provide crucial financial services and help bridge the gap in financial inclusion.

AI technologies are already deployed across the financial services sector.

AI is used in activities such as fraud detection, underwriting, debt collection, customer onboarding, real estate, investment research, property management, and customer service. For example, financial services firms use AI to augment firms' financial and cybercrime monitoring systems, including for AI-enabled threats by malicious actors.

Some regulators use AI to identify non-compliance with regulations.

AI is already being utilized as a regulatory tool and has been successfully used in anti-money laundering, countering the financing of terrorism, and compliance with sanctions.

Small financial services firms can be at a disadvantage in AI adoption.

Large financial institutions have substantial resources to invest in AI development, while smaller firms often lack sufficient resources to adopt AI. This disparity allows larger firms to leverage AI more effectively.

Recommendations

Recommendation: Foster an environment where financial services firms can responsibly adopt the benefits of AI technology.

Appropriate uses of AI by participants in the financial services sector can provide numerous benefits, including reduced costs, higher-quality services, and greater access for a wide array of consumers. While risks accompany the use of AI, as with any emerging technology, financial regulators should ensure that market participants using AI continue to comply with existing laws and regulations or examine alternative compliance processes, where appropriate.

Recommendation: Encourage and resource regulators to increase their expertise with AI.

Regulators should adopt AI to improve their efficiency and productivity and to gain a deep understanding of how AI can be used by market participants and across the financial services sector. Greater expertise would ensure financial regulators have the appropriate focus and tools to effectively oversee new AI products and services, apply and enforce existing laws, and assess regulatory gaps as market participants adopt AI.

Recommendation: Maintain consumer and investor protections in the use of AI in the financial services and housing sectors.

Congress should also examine the benefits and risks of AI use in the financial services and housing sectors and ensure existing consumer and investor protections, including those on discrimination in decision-making, are maintained and appropriately monitored by regulators.

Recommendation: Consider the merits of regulatory "sandboxes" that could allow regulators to experiment with AI applications.

Regulating the use of AI in financial services requires regulators to have a comprehensive understanding of AI technologies. This, in turn, requires regulators to develop practical expertise with AI tools through hands-on experience. Regulatory sandboxes could allow regulators to experiment with different uses of AI in a highly controlled manner, managing the possible risks of that use. Regulators could observe the impact of AI and adapt accordingly.

Recommendation: Support a principles-based regulatory approach that can accommodate rapid technological changes.

Primary regulators understand their respective fields, markets, and AI use cases within those markets. A sectoral, technology-neutral approach to financial services regulation would permit primary regulators to leverage their expertise and retain their existing authority, even when it intersects with artificial intelligence.

Recommendation: Ensure that regulations do not impede small firms from adopting AI tools.

Large financial institutions have substantial resources to invest in AI development, including the ability to build proprietary AI models, hire specialized technical talent, and integrate advanced technologies across their operations. In contrast, smaller firms often lack the resources to develop, acquire, and scale sophisticated AI tools. This disparity creates a competitive advantage, allowing larger firms to leverage AI more effectively to enhance their services, optimize operations, and reduce costs.

Current regulations may inadvertently favor larger firms with the resources to navigate complex regulatory landscapes and implement relevant compliance measures effectively. This regulatory burden can stifle innovation and limit the ability of smaller firms to compete on an equal footing with larger institutions. Regulatory frameworks should be tailored to the diverse range of market participants and their needs.

Appendix I: AI Task Force Members

Co-Chairman Jay Obernolte (R-CA)

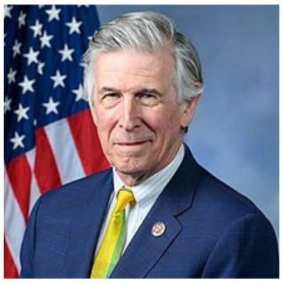
Science, Space, and Technology Committee,
Energy and Commerce Committee

Co-Chairman Ted Lieu (D-CA)

Judiciary Committee,
Foreign Affairs Committee

Ami Bera (D-CA)

Foreign Affairs Committee,
Permanent Select Committee on Intelligence



Don Beyer (D-VA)

Ways and Means Committee,
Joint Economic Committee

Suzanne Bonamici (D-OR)

Science, Space, and Technology Committee,
Education and the Workforce Committee

Eric Burlison (R-MO)

Education and the Workforce Committee,
Oversight and Accountability Committee,
Transportation and Infrastructure Committee

Kat Cammack (R-FL)

Agriculture Committee,
Energy and Commerce Committee

Yvette Clarke (D-NY)

Energy and Commerce Committee,
Homeland Security Committee

Ben Cline (R-VA)

Appropriations Committee,
Budget Committee,
Judiciary Committee

Michael Cloud (R-TX)

Oversight and Accountability Committee,
Appropriations Committee

Neal Dunn (R-FL)

Energy and Commerce Committee,
Select Strategic Competition Between the U.S. and the Chinese Communist Party Committee

Anna Eshoo (D-CA)

Energy and Commerce Committee

Bill Foster (D-IL)

Financial Services Committee

Valerie Foushee (D-NC)

Science, Space, and Technology Committee,
Transportation and Infrastructure Committee

Scott Franklin (R-FL)

Science, Space, and Technology Committee,
Appropriations Committee,
Veterans' Affairs Committee

French Hill (R-AR)

Foreign Affairs Committee,
Financial Services Committee,
Permanent Select Committee on Intelligence

Darrell Issa (R-CA)

Science, Space, and Technology Committee,
Foreign Affairs Committee,
Judiciary Committee

Sara Jacobs (D-CA)

Foreign Affairs Committee,
Armed Services Committee

Laurel Lee (R-FL)

Homeland Security Committee,
Judiciary Committee,
Administration Committee

Rich McCormick (R-GA)

Science, Space, and Technology Committee,
Foreign Affairs Committee,
Armed Services Committee

Alexandria Ocasio-Cortez (D-NY)

Oversight and Accountability Committee,
Natural Resources Committee

Brittany Pettersen (D-CO)

Financial Services Committee

Michelle Steel (R-CA)

Education and the Workforce Committee,
Ways and Means Committee,
Select Strategic Competition Between the U.S. and the Chinese Communist Party Committee

Haley Stevens (D-MI)

Science, Space, and Technology Committee,
Education and the Workforce Committee,
Select Strategic Competition Between the U.S. and the Chinese Communist Party Committee

Appendix II: AI Task Force Events

On February 20, 2024, Speaker Mike Johnson and Democratic Leader Hakeem Jeffries publicly announced the formation of the House Bipartisan Task Force on AI. The Task Force convened regularly to interpret how Congress should pursue AI innovation and address concerns facing Americans. Panels consisted of officials from academia, industry, and government, as described below:

AI Generated Content (Deepfakes), Harms, and Remediations

- Santiago Lyon, Head of Advocacy and Education for the Content Authenticity Initiative (CAI)
- Soheil Feizi, Associate Professor of Computer Science, University of Maryland, and Director of the Reliable AI Lab
- Alexandra Reeve Givens, President & Chief Executive Officer, Center for Democracy and Technology
- Michael Marando, Content Policy Director, Meta

Workforce, Education, Skill Gaps, Training, and Augmentation

- Simon Johnson, Ronald A. Kurtz Professor of Entrepreneurship and Head of Global Economics and Management, MIT
- Allyson Knox, Senior Director of Education and Workforce Policy, Microsoft
- Pat Yongpradit, Chief Academic Officer of Code.org and Lead of TeachAI
- Amanda Ballantyne, Director, Technology Institute at AFL-CIO

Standards and Government Use

- Elham Tabassi, Associate Director for Emerging Technologies in the Information Technology Laboratory (ITL) and Chief Technology Officer of the U.S. AI Safety Institute, National Institute of Standards and Technology (NIST)
- Cindy Martinez, Acting Director for AI Policy, Office of Management & Budget (OMB)

Federal Preemption of State Law

- Jim Harper, Senior Fellow, American Enterprise Institute (AEI)
- Adam Thierer, Senior Fellow, R Street Institute
- Woodrow Hartzog, Professor of Law, Boston University School of Law
- Lori Wallach, Director of the Rethink Trade Program, American Economic Liberties Project

Open-Source Models vs Closed-Source Models

- Nik Marda, Technical Lead, AI Governance, Mozilla
- Irene Solaiman, Head of Global Policy, Hugging Face
- Tim Fist, Research Fellow, Artificial Intelligence, Institute for Progress
- Helen Toner, Director of Strategy and Foundational Research Grants, Georgetown's Center for Security and Emerging Technology

Healthcare Applications

- Dr. Taha Kass-Hout, Chief Technology Officer, GE HealthCare
- Dr. Bradley Malin, Accenture Professor of Biomedical Informatics, Biostatistics, and Computer Science, Vanderbilt University
- Shannon Curtis, Assistant Director, Division of Federal Affairs, American Medical Association (AMA)
- Dr. Sara Murray, M.D., Chief Health AI Officer, University of California, San Francisco

Healthcare Applications with GAO's Science, Technology Assessment, and Analytics Team

Privacy, Identity, Transparency

- Keir Lamont, Director of the Future of Privacy Forum
- Pat Kinsel, Chief Executive Officer, Proof
- Jennifer Huddleston, Senior Fellow, CATO
- Dr. Sorelle Friedler, Shibulal Family Professor of Computer Science, Haverford College
- Brandon Pugh, Senior Fellow, R Street Institute

Civil Rights & Civil Liberties

- Maya Wiley, President and Chief Executive Officer, The Leadership Conference on Civil and Human Rights
- Damon T. Hewitt, President and Executive Director, Lawyers' Committee for Civil Rights Under Law
- Neil Chilson, Head of AI Policy, Abundance Institute

Energy Usage & Hardware

- Joshua Parker, Director of Sustainability, Nvidia
- Alicia Ruckteschler, Chief Procurement Officer, Equinix
- Dr. Arman Shehabi, Staff Scientist, Lawrence Berkeley National Laboratory
- Dr. Valerie Taylor, Director of the Mathematics and Computer Science Division, Argonne National Lab

Financial Services and Banking

- Ari Tuchman, Chief Executive Officer, Quantifind
- John Morgan, Senior Director, Capital One
- Peter Licursi, Chief Strategy Officer, S&P Global
- Jennifer Chien, Senior Policy Counsel for Financial Fairness, Consumer Reports

Intellectual Property and Copyright

- Aaron Cooper, Vice President, Global Policy, BSA | The Software Alliance
- Ken Doroshow, Chief Legal Officer, Recording Industry Association of America
- Keith Kupferschmid, President & Chief Executive Officer, Copyright Alliance
- Jennifer Rothman, Nicholas F. Gallicchio Professor of Law, University of Pennsylvania Law School
- Corey Salsberg, Vice President & Global Head of IP Affairs, Novartis
- Ben Sheffner, Senior Vice President & Associate General Counsel, Law & Policy, Motion Picture Association

Small Business and Entrepreneurship

- Tyrance Billingsly, Chief Executive Officer, Black Tech Street
- Michael Richards, Director of Policy for Technology Engagement, U.S. Chamber of Commerce
- Nathan Lindfors, Policy Director, Engine

National Security

- Department of Defense
- Office of the Director of National Intelligence

Roundtables with AI CEOs & Experts

- Sam Altman, Chief Executive Officer, OpenAI
- Jack Clark, Co-Founder and Head of Policy, Anthropic
- Alexandr Wang, Founder and Chief Executive Officer, ScaleAI
- Tom Siebel, Chairman and Chief Executive Officer, C3 AI
- Aidan Gomez, Co-Founder and Chief Executive Officer, Cohere
- Marc Andreessen, Co-Founder and General Partner, Andreessen Horowitz
- Max Tegmark, Professor of Physics, Massachusetts Institute of Technology, and President, Future of Life Institute

Executive Branch Meetings

- Dr. Arati Prabhakar, the Director of the White House Office of Science and Technology Policy (OSTP)
- Saif Kahn, Sr. Advisor to the Secretary of Commerce for Critical and Emerging Technologies, Department of Commerce (DOC)
- Helena Fu, the Director of the Office of Critical and Emerging Technologies under the Department of Energy (DoE)
- Travis Hall, Associate Director at the Office of Policy Analysis and Development under the National Telecommunications and Information Administration

Appendix III: Key Government Policies

Executive Order 13859: Maintaining American Leadership in Artificial Intelligence[1]

This Executive Order directs federal agencies to prioritize investments and initiatives that promote AI innovation. The order is intended to ensure American leadership in AI, protect economic and national security, and establish guidelines for responsible and ethical AI development.

Executive Order 13960: Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government[2]

This order sets forth policies and principles for federal agencies to follow when designing, developing, acquiring, and implementing AI technologies. The Executive Order also highlights AI's role in enhancing government operations and services, prioritizing safety and efficiency, and encouraging innovation while safeguarding civil liberties and privacy.

National Artificial Intelligence Initiative Act of 2020[3]

The law establishes a federal strategy to advance AI research, development, and adoption. The law is intended to strengthen AI competitiveness and leadership in AI, foster workforce development, and promote ethical standards in AI use.

AI in Government Act[4]

This law requires OMB to issue government-wide guidance on agency use of AI and agency AI governance plans. The law also created an AI Center of Excellence within the General Services Administration (GSA) to provide technical expertise, coordinate AI initiatives, and assist agencies in adopting AI solutions.

[1] Executive Office of the President. "Maintaining American Leadership in Artificial Intelligence." Federal Register, 14 Feb. 2019, www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence.

[2] Executive Office of the President. "Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government." Federal Register, 8 Dec. 2020, www.federalregister.gov/documents/2020/12/08/2020-27065/promoting-the-use-of-trustworthy-artificial-intelligence-in-the-federal-government.

[3] Office of the Federal Register, National Archives and Records Administration. Public Law 116 - 283 - William M. (Mac) Thornberry National Defense Authorization Act for Fiscal Year 2021. U.S. Government Publishing Office, 31 Dec. 2020, https://www.govinfo.gov/app/details/PLAW-116publ283.

[4] Office of the Federal Register, National Archives and Records Administration. Public Law 116 - 260 - Consolidated Appropriations Act, 2021. U.S. Government Publishing Office, 26 Dec. 2020, https://www.govinfo.gov/app/details/PLAW-116publ260.

The CHIPS and Science Act of 2022[5]

This law appropriated $50 billion to the Department of Commerce to restore advanced semiconductor manufacturing in America and reauthorized the government research enterprise for 5 years. In addition, the bill authorized various federal science agencies to explore artificial intelligence research and applications, including the Department of Energy (DOE), the National Oceanic and Atmospheric Administration (NOAA), and the National Institute of Standards Technology (NIST). It further expanded NIST's AI responsibilities with provisions to establish testbeds and technical standards for promoting the safety and trustworthiness of AI systems.

Advancing American AI Act[6]

This law requires specified federal agencies to take steps to promote responsible AI acquisition and use while protecting privacy, civil rights, and civil liberties.

Executive Order 14110: Executive Order on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence[7]

This Executive Order outlines the Administration's policy goals regarding AI. The order is intended to promote competition, protect civil liberties, and maintain U.S. global competitiveness in AI.

[5] Office of the Federal Register, National Archives and Records Administration. Public Law 117 - 167 - An act making appropriations for Legislative Branch for the fiscal year ending September 30, 2022, and for other purposes. U.S. Government Publishing Office, 8 Aug. 2022, https://www.govinfo.gov/app/details/PLAW-117publ167.

[6] Office of the Federal Register, National Archives and Records Administration. Public Law 117 - 263 - James M. Inhofe National Defense Authorization Act for Fiscal Year 2023. U.S. Government Publishing Office, 22 Dec. 2022, https://www.govinfo.gov/app/details/PLAW-117publ263.

[7] The White House. "Fact Sheet: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence." The White House, 30 Oct. 2023, https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/.

Appendix IV: Areas for Future Exploration

As discussed in the introduction, this report is not the final word on AI issues for Congress. While the House AI Task Force has engaged in a robust process of interviews, meetings, and stakeholder roundtables, many issues of significant relevance to AI were not fully explored by the Task Force or this report. The House AI Task Force encourages members, committees of jurisdiction, and future congresses to continue to investigate opportunities and challenges related to AI.

The following is a list of potential areas of future exploration related to AI that a future Congress may consider investigating:

1. Global Development and International Cooperation
2. Export Control Policy
3. Manufacturing, Supply Chain, and Industrial AI
4. Antitrust and Competition Policy
5. Critical Infrastructure and Security
6. Environmental Impact of AI
7. Law Enforcement
8. The Intelligence Community
9. Transportation
10. Election Integrity
11. State and Local Governments
12. Biotechnology
13. Law and the Courts
14. AI Adoption Across Sectors, including Entertainment, and Business to Business.
15. Other Industries

Appendix V: Overview of AI Technology

Term	Definition
Artificial Intelligence (AI)	Software systems capable of performing tasks typically expected to require human intelligence, e.g., voice recognition, image analysis, and language translation. The field of AI encompasses various subfields, including machine learning, natural language processing, and computer vision.
Machine Learning (ML)	The subfield of artificial intelligence that involves software learning and improving from data. ML algorithms can analyze large amounts of data, identify patterns in that data, and, based on those patterns, make predictions or decisions without being explicitly programmed how to do so. Generally, using more data in training results in better performance on the task the software is trained for.
AI Model	A software program that receives input data, such as text, images, or numbers, and processes those inputs to produce specific types of outputs, such as predictions, recommendations, or generated content. Many AI models today use ML to “learn” how to produce outputs from inputs. The larger the model and the larger the training data set, the better the model performs.
Generative AI	AI systems that can generate new content, such as text, images, video, and music, with minimal or no human guidance on how exactly to create that content. Some generative AI systems allow the user to specify the general nature or characteristics of the content to generate. A generative AI system is designed to produce content that is novel rather than copied from existing data. Generated content is also intended to be realistic in that it resembles human-created content. Typically, the content of the training data determines the types of content that can be generated.
Large Language Models (LLMs)	A powerful generative AI model trained on vast amounts of text data, giving it the capability to understand text and generate human-like text. LLMs are useful for a wide range of natural language processing tasks, such as chatbots, text summarization, and language translation.
Neural Network	A type of machine learning model consisting of interconnected nodes, or "artificial neurons," typically organized in one or more layers. Once the neural network is trained, the nodes cooperate to transform input data into outputs such as predictions, classification decisions, or generated content.
Deep Learning	An ML paradigm that uses neural networks with many layers. In general, the more layers in the neural network, the greater the performance of the neural network. Deep learning systems allow more sophisticated patterns to be recognized and more complex tasks to be performed. Deep learning has been responsible for many of the breakthroughs in AI over the past decade.

Natural Language Processing (NLP)	The subset of AI that involves processing human language, such as written text. NLP enables machines to understand, interpret, and generate human language, facilitating tasks like language translation, text summarization, and chatbots.
Computer Vision	The subset of AI that involves understanding and interpreting visual information from images or videos. Computer vision allows machines to recognize objects, identify faces, and analyze various types of visual content.
Foundation Models	A type of AI model that, through training on vast amounts of data, is general purpose enough to be used in a wide variety of different tasks.
Compute	The computational resources that are required to train and run AI models efficiently. It encompasses the computer hardware, memory, and other resources needed to create and use AI models. With the increasing complexity and size of AI models, compute has become a crucial resource.

Appendix VI: Definitional Challenges of AI

If artificial intelligence is to be addressed by government activity, such as through legislation or agency rulemaking, it must first be defined appropriately so that the nature of government activity is transparent with clearly delineated boundaries.[8] A poorly crafted definition risks being overinclusive or underinclusive and may have unintended effects.

Unfortunately, it can be surprisingly difficult to define the term “artificial intelligence.” Subtle word choices in the definition can significantly influence its scope.[9] When New York City established a task force to investigate the use of AI and other types of Automated Decision Systems, they quickly concluded that the definition in the statute posed challenges; it was so broad it could include general tools like internet searches or spreadsheets.[10]

The UN Educational, Scientific and Cultural Organization (UNESCO) avoids defining AI altogether and instead focuses on the particular consequences of using AI.[11] Similar challenges exist in defining related terms, such as “frontier AI.”[12] For example, different definitions of AI are used in the National Artificial Intelligence Initiative Act of 2020[13] (NAIIA) and the John S. McCain National Defense Authorization Act for Fiscal Year 2019 (NDAA).[14]

[8] Defining AI is not merely an academic exercise, particularly when drafting legislation. AI research and applications are evolving rapidly. Thus, congressional consideration of whether to include a definition for AI in a bill and, if so, how to define the term or related terms, necessarily includes attention to the scope of the legislation and the current and future applicability of the definition. See: Laurie Harris, Congressional Research Service, “Artificial Intelligence: Background, Selected Issues, and Policy Considerations,” Congressional Research Service, 19 May 2021, https://crsreports.congress.gov/product/pdf/R/R46795

[9] Matt O'Shaughnessy, “One of the Biggest Problems in Regulating AI Is Agreeing on a Definition”, Carnegie Endowment, 6 Oct. 2022, https://carnegieendowment.org/posts/2022/10/one-of-the-biggest-problems-in-regulating-ai-is-agreeing-on-a-definition?lang=en.

[10] “New York City Automated Decision Systems Task Force Report”, New York City Automated Decision Systems (ADS) Task Force, November 2019, https://www.nyc.gov/assets/adstaskforce/downloads/pdf/ADS-Report-11192019.pdf.

[11] UNESCO, “Ethics of Artificial Intelligence”, *UNESCO*, https://www.unesco.org/en/artificial-intelligence/recommendation-ethics#:~:text=The%20Recommendation%20interprets%20AI%20broadly,make%20future%2Dproof%20policies%20infeasible.

[12] Helen Toner and Timothy Fist, “Regulating the AI Frontier: Design Choices and Constraints”, *Georgetown Center for Security and Emerging Technology*, 26 Oct. 2023, https://cset.georgetown.edu/article/regulating-the-ai-frontier-design-choices-and-constraints/.

[13] Supra 3.

[14] Office of the Federal Register, National Archives and Records Administration. Public Law 115 - 232 - John S. McCain National Defense Authorization Act for Fiscal Year 2019. U.S. Government Publishing Office, 12 Aug 2018, https://www.govinfo.gov/app/details/PLAW-115publ232.

Section 5002(3) of the NAIIA defines artificial intelligence as:

> *"a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. Artificial intelligence systems use machine and human-based inputs to— (A) perceive real and virtual environments; (B) abstract such perceptions into models through analysis in an automated manner; and (C) use model inference to formulate options for information or action."*

In contrast, section 238(g) of the NDAA defines artificial intelligence as including any of the following:

> (1) Any artificial system that performs tasks under varying and unpredictable circumstances without significant human oversight or can learn from experience and improve performance when exposed to data sets.
> (2) An artificial system developed in computer software, physical hardware, or other context that solves tasks requiring human-like perception, cognition, planning, learning, communication, or physical action.
> (3) An artificial system designed to think or act like a human, including cognitive architectures and neural networks.
> (4) A set of techniques, including machine learning, is designed to approximate a cognitive task.
> (5) An artificial system designed to act rationally, including an intelligent software agent or embodied robot that achieves goals using perception, planning, reasoning, learning, communicating, decision-making, and acting.

The recently introduced Artificial Intelligence Research, Innovation, and Accountability Act of 2023 (S.3312)[15] defines an artificial intelligence system as:[16]

> an engineered system that—
> (A) generates outputs, such as content, predictions, recommendations, or decisions for a given set of human-defined objectives; and
> (B) is designed to operate with varying levels of adaptability and autonomy using machine and human-based inputs.

The variety of definitions suggests that no single definition is appropriate for all situations. It may be best to narrowly tailor the definition of artificial intelligence depending on the applicable sector and policy objectives of the government activity.

[15] "S. 3312 – 118th Congress (2023-2024): "Artificial Intelligence Research, Innovation, and Accountability Act of 2023," Congress.gov, Library of Congress, 15 Nov. 2023, https://www.congress.gov/bill/118th-congress/senate-bill/3312

[16] Id. Sec. 101.

Acknowledgments

This AI Task Force report is the result of extensive bipartisan efforts, insights, and hard work. We are proud to share this work as a reflection of our collective efforts and shared commitment. We extend our gratitude to all contributors whose unwavering commitment to excellence has set a high standard for this work.

First, we offer special thanks to Speaker Johnson and Democratic Leader Jeffries for their leadership and for establishing the Bipartisan Task Force on Artificial Intelligence.

We next offer our sincerest gratitude to the 22 members who joined us on the AI Task Force and dedicated long hours, thoughtful consideration, and valuable policy expertise to the many complex issues we considered this year.

As members of Congress, we all benefitted greatly from distinguished panelists who offered invaluable insights and experience to help shape the discussion of our report.

We also thank the many committee staff who generously contributed their time, knowledge, and expertise throughout the preparation of this report. Your contributions have enriched this work, and your collaborative spirit is deeply appreciated.

Specifically, we would like to thank the House Committees on: Agriculture; Armed Services; Education and Workforce; Energy and Commerce; Financial Services; Intelligence; Judiciary; Oversight and Accountability; Science, Space, and Technology; Small Business; and Ways and Means.

Next, we would like to thank the staff who dedicated time and effort to develop and finalize this report: Lorissa Bounds, Rob Hicks, Marc Cevasco, Zach Atran, Leah Uhrig, Josh Mathis, Heather Vaughan, Cate Johnson, Bridget Dunn, Alex Murglin, Cody Elkins, Troy Hall, Dahlia Sokolov, Albert Hinman, Steffani Silva, Joseph Orellana, Adam Rosenberg, John Piazza, and Andrew Rubin.

Finally, we would also like to acknowledge the lead staff of this report: Dean Alderucci, Holten Stringer, Alan McQuinn, and Connie Hsueh.

Their extraordinary efforts went above and beyond in leading this project from inception to completion. Their tireless commitment has been instrumental in crafting the report's narrative and analysis. Without their dedication, this report would not have been possible.