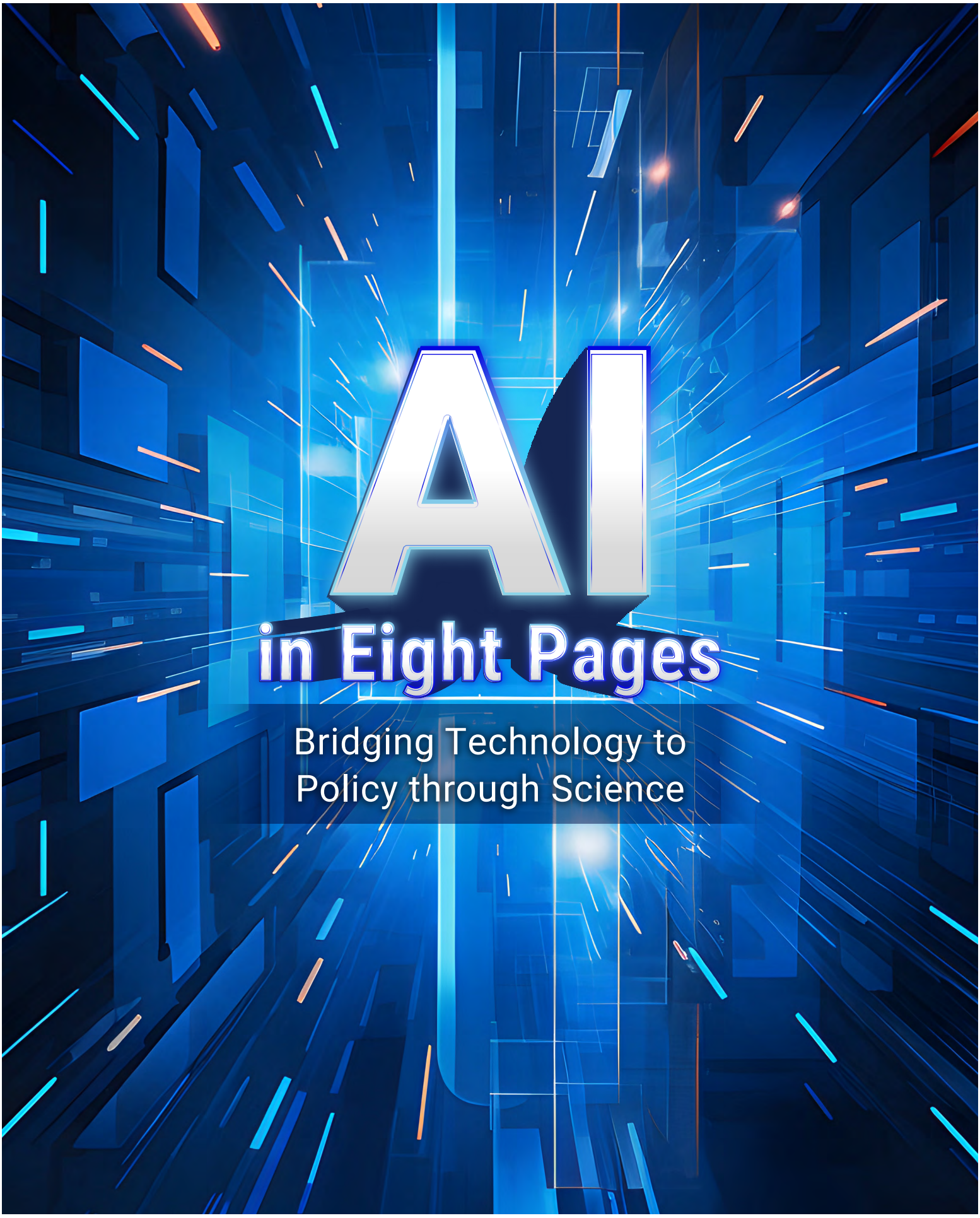
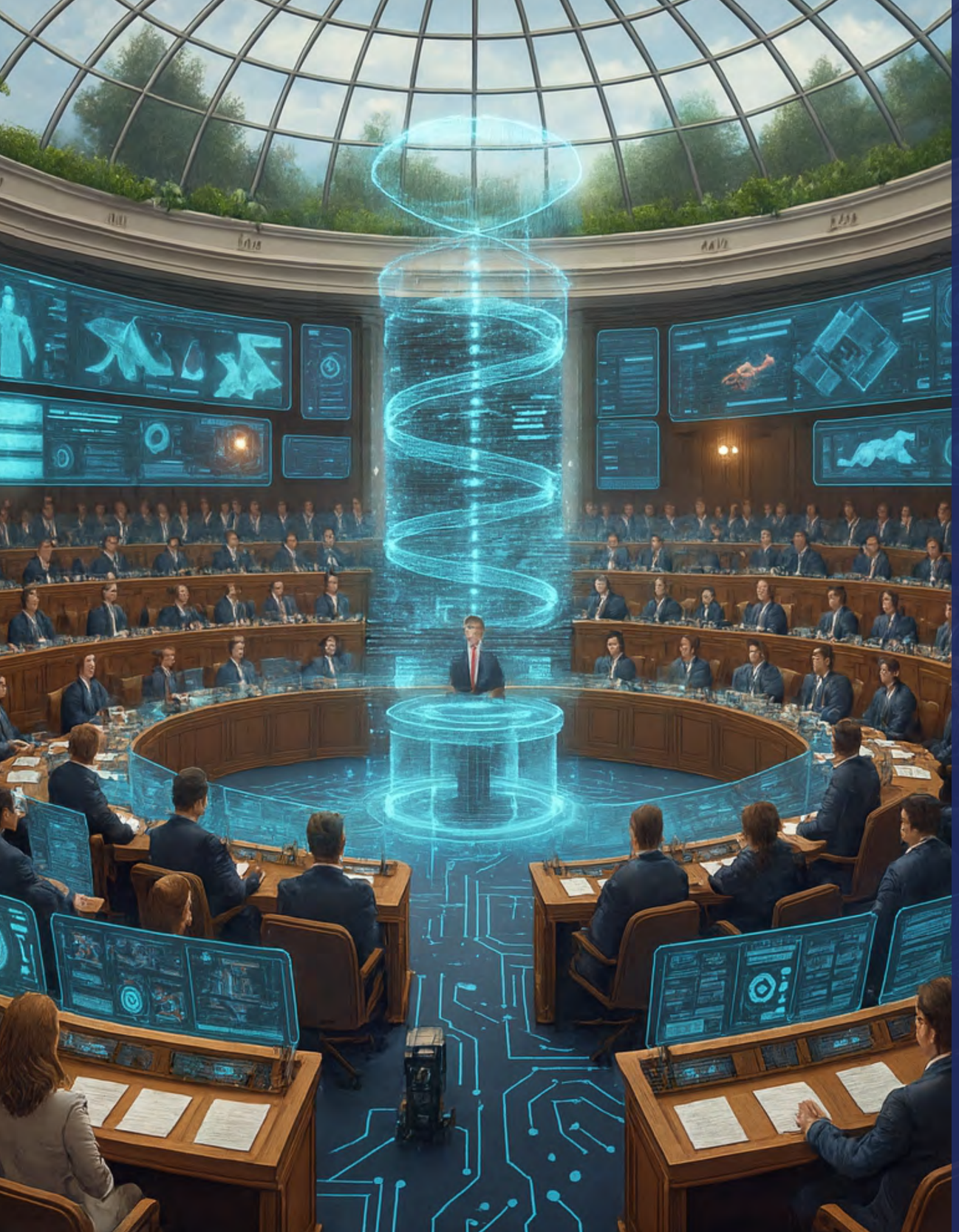


Lawrence Livermore National Laboratory

2026



AUTHORS

Brian Giera, Cindy Gonzales, and Caspar L. Donnison

Lawrence Livermore National Laboratory, Livermore, CA 94550

ACKNOWLEDGEMENTS

This work was authored by Lawrence Livermore National Laboratory, with funding from the Livermore Lab Foundation, a 501c3 nonprofit organization supporting LLNL's science and research initiatives, and the commitment of the California Foundation for Commerce and Education (CFCE).

CITE REPORT AS

Brian Giera, Cindy Gonzales, and Caspar L. Donnison, AI in 8 Pages: Bridging Technology to Policy through Science, November 2025, Lawrence Livermore National Laboratory Report, LLNL-TR-2006290.

FOR MORE INFORMATION, PLEASE VISIT:

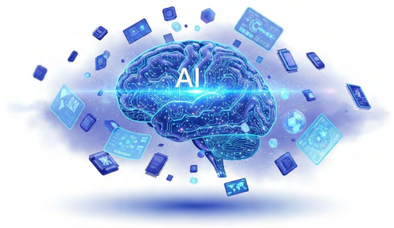
ai.llnl.gov

Table of Contents

Executive Summary

What is AI?

Artificial Intelligence (AI) describes computational tools that process data to identify patterns, solve specific problems, and generate outputs that resemble human-created work—creating the impression of reasoning, perception, or language understanding.



Why does it matter now?

AI has crossed from research labs into daily life. Rapid adoption and deployment across sectors introduces new capabilities and considerations for safety, governance, and societal impact.

Key opportunities

Job Transformation and Workforce Upskilling: AI shifts human roles to more creative and strategic tasks, creating opportunities for upskilling, greater job satisfaction, and new career paths.

Productivity and Efficiency Gains: Automation frees employees from repetitive work, driving productivity, innovation, and cost savings.

Personalization and Enhanced User Experiences: AI enables tailored services and products, improving individual outcomes and opening new business opportunities.

Key challenges

Economic Disruption and Job Displacement: Rapid AI adoption may cause job losses and wage polarization, especially in knowledge-based sectors, challenging economic stability, and workforce adaptation.

Technical Failures and Errors: AI systems are vulnerable to bias, lack of explainability, and system outages, leading to real-world risks in sensitive domains.

Misinformation and Security Risks: AI can generate false or misleading information and is susceptible to attacks and data breaches, raising concerns about safety, privacy, and trust.

Bridging promise to governance

Balanced AI governance involves collaboration between technical experts and regulators to ensure innovation is secure, ethical, and sustainable, without stifling progress.

Risk mitigation and targeted safeguards for data, compute, model design, and deployment—focusing on transparency, sustainability, safety, and accountability.

Responsible AI development to align it with the public interest and responsible use will require improving AI literacy, transparency, safety, and accountability.

What is this primer all about?

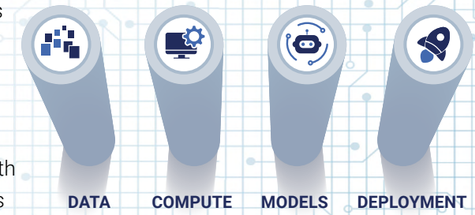
California stands at the forefront of the AI revolution, uniquely positioned with its concentration of world-leading AI companies, research institutions, and vibrant startup ecosystem. The state's academic excellence, robust public-private partnerships, and significant venture capital investment have made it a global hub for AI innovation and application across diverse sectors, from healthcare and energy to entertainment and disaster response. In 2024, Californian businesses drew in 80% ($72bn) of U.S. AI funding, according to the Bay Area Council Economic Institute.

AI is already used in a variety of applications, including self-driving cars, medical imaging, note transcription, and tools that predict wildfire risk. You likely unlocked your phone several times today using AI-powered facial recognition. There is growing evidence that AI technologies can accelerate the rate of scientific research and innovation, a potential that, if harnessed correctly, could boost economic growth. However, few perspectives bridge the promise of AI to the governance and policy frameworks needed to ensure responsible use and that its benefits are shared broadly.

Amidst the excitement and scrutiny around AI, this primer provides a scientific overview of AI in terms of technical foundations, real-world applications, and policy, with a focus on California's operating environment. It grounds the discussion in sector-specific examples of current and emerging AI use cases. The primer also presents AI failure modes and social and economic challenges, clarifying the balance between innovation and regulation, and what safeguards are currently under discussion. This perspective is evidence-driven and nonpartisan, intended to support informed decision-making on AI.

Why am I hearing about AI now?

Today's accelerated innovation in AI stems from increasing capabilities across four pillars: more **data**, more **compute** (i.e., computational power), better design of **models**, and large-scale **deployment**. Popular AI tools communicate in human language and have been downloaded onto millions of phones and computers, helping people to tackle personal and professional tasks. To demonstrate the power and breadth of AI tools, art in this primer is AI-generated and large language models have been used to generate section summaries.



What are the economic opportunities of AI?

Historically, new technologies are disruptive before creating job opportunities and spurring economic growth. The Industrial Revolution, for instance, displaced many manual trades but ultimately expanded manufacturing and engineering work, while the digital revolution automated clerical tasks and generated entire digital industries. There is debate over the extent to which AI will support or replace jobs. So far, it appears to be augmenting human labor and skills, requiring adaptation in work practices, and not necessarily the mass replacement of human jobs. Today, AI can help automate repetitive or tedious tasks, so employees have more time to focus on creative parts of their job. For example, AI scribes allow healthcare providers more time with patients, while new video editing tools allow creative industry professional more time for storytelling (see Table 1). These changes are also driving demand for non-digital and interpersonal skills, as well as trade skills, with economic opportunities expected to grow in hospitality, education, and healthcare.

Table 1: Recent and future applications of AI in California's economy present both benefits and risks. Cybersecurity and job security concerns are relevant across all sectors.

Sector (% of CA GDP)	Current Uses of AI	Future Applications	Potential Benefits	Potential Risks or Concerns
Financial Activities (20%)	Fraud detection systems use AI to find transaction data anomalies	AI can assess credit using alternative data like mobile usage, purchase history, etc.	Lower fraud detection costs, faster transactions, greater financial access	Algorithm bias, market volatility, privacy breaches, fraudulent activity
Entertainment & Information (15%)	AI-generated content for music, film, and video games; personalized streaming recommendations	Fully AI-made films, virtual performers, personalized interactive content	Faster editing, freeing time for human creativity, personalized experiences, lower production costs	Misinformation (e.g., deepfakes), erosion of creative labor markets, copyright/legal ambiguity
Professional & Technical Services (13%)	Contract review, legal analysis, client-support chatbots	AI co-pilots support complex research, real-time consulting analysis	Less repetitive work, higher accuracy, faster innovation	Confidentiality risks, accuracy of outputs, overreliance on AI in decision-making
Materials & Manufacturing (10%)	Assembly lines use computer vision to spot defects	Robots and AI workflows enable fully autonomous factories	Safer, more productive factories	Technology lock-in due to expensive implementation costs
Transportation & Trade (10%)	Driverless vehicles, real-time routing, inventory, and fleet optimization	Public transit that adapts to demand, traffic, and weather; self-driving truck convoys	Reduced accidents, traffic, and transport emissions	Safety and liability concerns
Government (including Defense, 10%)	AI detects and counters network intrusions, malware, and phishing	AI forecasts geopolitical threats by analyzing key data	Faster response; enhanced situational awareness; reduced risk	Misinformation, escalation risks, public trust loss, accountability for AI decisions
Healthcare (7%)	AI tools enhance anomaly detection, reducing errors and accelerating care	Predictive health analytics: real-time wearable data analysis supports early intervention	Improved patient outcomes, faster diagnosis, cost reduction	Medical bias, error liability, privacy and consent concerns
Energy & Utilities (6%)	Energy companies use AI to forecast demand and balance grid supply in real time	Autonomous decentralized energy markets with household-level AI trading systems for demand response	Lower costs, increased reliability, renewable energy integration	Resource demand of data centers; grid vulnerability; data security
Educational Services (3%)	AI uses real-time student data to personalize learning	Interactive tailored teaching via AI avatars and voice assistants	Personalized learning; learning efficiency, early intervention	Unequal access widens learning gaps, teacher deskilling, privacy concerns
Agriculture (2%)	AI drones and sensors monitor crops and soil, enabling targeted yield-boosting interventions	AI models predict crop-climate response, guiding optimal crop and management choices	Higher yields, resource efficiency, less environmental impact	Adoption barriers for small farms, loss of traditional knowledge, data monetization concerns

Hey AI, summarize the previous section for me please?

New technologies often disrupt existing jobs before creating new economic opportunities, as seen in past industrial and digital revolutions. As AI adoption grows, it is changing work practices, augmenting human roles, and increasing demand for both technical and interpersonal skills across multiple sectors.

What are the risks of AI?

The advancements and adoption of AI present numerous risks, some already apparent and others still emerging. These are categorized below as **technical**, **economic**, and **social** risks, which often overlap. If AI is not carefully managed, it could fail to meet expectations. It is not infallible and can fail technically (Table 2); even when these weaknesses are addressed, AI's progression presents significant economic and social challenges.

Table 2: Common ways AI goes wrong and what that means in real settings.

Failure Type	Plain Language Description	Real World Scenario
Adversarial Attacks	Intentionally tricking AI	Modifying a stop sign so a self-driving car reads it as a speed limit
Data Bias	AI learns from unfair or incomplete information	A resume screener favors familiar resumes, overlooking strong new candidates
Data Drift	The world changes, but AI has not learned from new situations	A navigation tool suggests a closed route after new roadwork starts
Hallucinations	AI presents false information as true	AI invents case law, misleading a court
Hardware or Software Failures	A computer outage or crash interrupts the AI's operation	A server failure disables an AI triage tool in an emergency room
Lack of Explainability	AI decision-making lacks transparency and clear rationale	AI forecasts lower crop yield without providing drivers or actionable information
Spoofing	Fake or copied information fools the AI	Unlocking a phone with a printed photo instead of a real face
Systemic Errors	A model design flaw causes the same mistake consistently	Emergency calls are always routed incorrectly
Overfitting	AI performs well on training examples, but struggles with new conditions	Vision-based inspection trained on one camera setup fails after lighting is upgraded
Underfitting	Overly simple AI that misses important details	A load forecaster uses daily averages, missing morning and evening peaks, causing grid overloads

Technical Risks

AI technology can fail in a myriad of ways, and those failures translate into real-world risks when they occur in sensitive settings or at scale. Industry, academia, and research institutions (e.g., national labs) are all engaged in developing technical approaches to tackle these challenges.

Economic Risks

The rapid pace of AI development and its uneven adoption have spurred conversations about the risks of wage polarization, job losses, and how to ensure that the benefits of AI are broad-based. California's economy is set for both major opportunities and disruption, as AI-driven automation reshapes employment. AI tools are particularly effective at knowledge-based tasks, with near-term disruption expected in white-collar roles in IT, finance, and legal services. In contrast, labor-intensive industries—such as agriculture in the Central Valley and skilled trades—are expected to experience more limited labor effects. Mitigating the impact of job losses could be accomplished by focusing AI on task augmentation and new role creation that raise the value of human skills.

Social Risks

As AI is integrated into daily life and various sectors, California faces salient risks in security and privacy, the large-scale production and spread of misinformation, and growing resource demands from energy and water intensive data centers. Additional societal risks stem from AI as a dual-use technology, where the same capability can be applied for benefit or harm. Emerging threats are less predictable. For example, granting AI more decision authority in workflows can present unintended consequences. Additionally, widespread use of generative AI for therapy and companionship may create dependency and harm vulnerable users. Framing risks through real-world intent and context can help align governance with practical safeguards while preserving innovation.

Hey AI, summarize this section for me please?

AI advancements introduce a range of technical, economic, and social risks, including system failures, job disruption, and increased security and privacy concerns. These risks require ongoing technical solutions and careful governance to ensure AI's benefits are realized while minimizing potential harms.

How can we bridge technology and policy for responsible AI?

Bridging AI's promise to effective governance involves linking the technical foundations of innovation with the mechanisms that safeguard responsible use. Innovation alone is not enough: robust governance can help ensure AI progress is secure, ethical, and sustainable. Collaboration between technical experts and regulators helps balance progress with risk management. Table 3 highlights trade-offs between advancing innovation and managing risks and responsibilities in each technical pillar.

Table 3: *Advancing responsible AI requires innovations across four technical pillars, supported by governance that manages risks.*

Technical Pillars for AI	DATA	COMPUTE	MODELS	DEPLOYMENT
Advancing AI Capabilities	Use large, diverse datasets to improve model performance	Secure critical mineral supply chains for AI computer chips	Develop models tailored to intended use cases	Develop methods to understand AI model and system capabilities before deployment
	Source data from public, licensed, and user-generated content	Increase data center energy and cooling efficiency	Open-source models for collaboration and innovation	Rapidly deploy AI systems to real-world environments
	Maintain high standards for data quality and reliability	Data center demand flexibility and on-site storage	Provide clear documentation on model design	Support the development of international safety/security standards
Possible Risk Mitigations	Transparency about data sources and usage	Invest in grid infrastructure and resilience	Model design and training data transparency	Establish risk assessment and incident response protocols
	Protect user privacy and ensure data security	Implement sustainability standards and benchmarking for facilities	Standards that address misuse and ensure reliability	Cybersecurity and safety requirements
	Follow and update data protection laws and ethical guidelines	Support e-waste recycling and data center waste heat use	Access controls for dual-use specialized models	Continuously monitor of deployed AI systems for ethical concerns, security threats, and vulnerabilities

Exploring safeguards through policy

Policymakers can balance innovation with practical safeguards: promoting responsible access that broadens participation while constraining misuse; improving literacy and transparency that match tools to tasks; and implementing alignment mechanisms that couple commercial activity to clear documentation, monitoring, and incident reporting.

EDUCATION

An agile, well-skilled workforce, with statewide access to digital infrastructure, will be positioned to benefit from AI development. Users will need to understand technological limitations of AI and how to verify and validate outputs. For example, it may be acceptable for AI to make a mistake when drafting a casual email, but not when updating a legal document. AI literacy can be supported through public education initiatives, updating the school curriculum (as under AB 2876, Berman, 2024)' and by encouraging AI developers to disclose system performance and limitations.

Related technical pillars ordered by importance:

TRANSPARENCY

Stanford's Foundation Model Transparency Index finds that many AI companies perform poorly. Incentivizing AI companies to be transparent, such as sharing information about their models and training data with policy makers and AI experts, can better align commercial interest with the public good. AB 2013 (Irwin, 2024) established transparency requirements for training data, effective in 2026 but applying retroactively to systems released since 2022.

Related technical pillars ordered by importance:

SAFETY

Policies that target intentional misuse of AI—such as cyber-attacks, misinformation, or worker exploitation—can protect society. Monitoring and reporting, such as "adverse event reporting" as put forward in SB 53 (Wiener, 2025), could help policymakers track real-world system performance and emerging risks.

Related technical pillars ordered by importance:

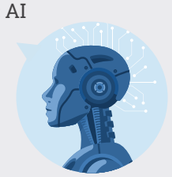
ACCOUNTABILITY

Third-party risk assessments disclosure requirements, and addressing model bias can all support accountability. SB 53, aims to enhance transparency, safety, accountability, and innovation in the creation and deployment of frontier AI models. It requires large AI developers to publicly disclose their safety and security protocols; establishes the CalCompute consortium to foster ethical and sustainable AI research; and implements mechanisms for safety-incident reporting and whistleblower protections.

Related technical pillars ordered by importance:

Hey AI, summarize this for me please?

Bridging technical foundations and policy for responsible AI means advancing AI capabilities while implementing governance measures—like transparency, safety, and accountability—to manage risks and ensure ethical, sustainable use. Collaboration between technical experts and policymakers is essential to balance innovation with safeguards that protect society and promote trust.



How do we keep up with AI?

The rapid integration of AI into daily life presents complex challenges, as policymakers weigh gains in efficiency, economic growth, and public services against risks to privacy, security, access, and workforce stability. A multifaceted approach grounded in technical understanding of how AI systems function, including their strengths, vulnerabilities, and regulatory considerations, can help frame policy choices. The policy landscape will remain dynamic because the technologies and their societal impacts are evolving quickly. Ongoing collaboration among government, industry, and the research community, supported by up-to-date expertise and real-world evidence, can inform strategies that anticipate new challenges, mitigate risks, and enable broad realization of AI's benefits across California.

This work was performed under the auspices of the U.S. Department of Energy by
Lawrence Livermore National Laboratory under Contract DEAC52-07NA27344. LLNL-TR-2013164