BETTER TECH TOMORROW

PROMISE AND PERIL:
HOW TEXAS CAN LEAD ON RESPONSIBLE ARTIFICIAL INTELLIGENCE POLICY

WRITTEN BY
Zach Whiting and David Dunmoyer
November 2024

TABLE OF CONTENTS

PROMISE AND PERIL: HOW TEXAS CAN LEAD ON RESPONSIBLE ARTIFICIAL INTELLIGENCE POLICY

WRITTEN BY **Zach Whiting and David Dunmoyer**

KEY POINTS

- **Humanity is at** an inflection point. Artificial intelligence (AI) offers boundless promises while also presenting critical challenges to civilization.
- **AI as a field is replete** with technical components that must be neatly defined and understood to legislate prudently—both in terms of the enforceability and scope of public policy.
- **AI applications** vary from sector-to-sector. This paper provides numerous case studies to illustrate how AI is being used and its impacts on healthcare, law enforcement, national security, kids' online safety, education, finance, and more.
- **Industry and legislators** must balance technological innovation with utmost respect for human dignity, privacy, transparency, and accountability. Ultimately, light-touch, values-driven, state-based, legislative guardrails for AI are necessary to propel humanity forward.

EXECUTIVE SUMMARY

Humanity is at an inflection point. Artificial intelligence (AI) offers boundless promises while also presenting critical challenges to civilization. It presents what media theorist Neil Postman would call an ecological change—that is, it does not add to or subtract from, but fundamentally transforms. For the numerous occasions in which AI provides positive developments, there are also many concerns associated with its growth and proliferation. What is AI? How should AI be used? Are there ways it should not be used? For example, should it be allowed to replace, or just simply augment, human reason, judgment, and labor? Are civil liberties enhanced or put at risk by its widespread adoption? These are some of the questions that lawmakers must consider when assessing AI regulation, and these are among the questions the authors intend to address in this paper.

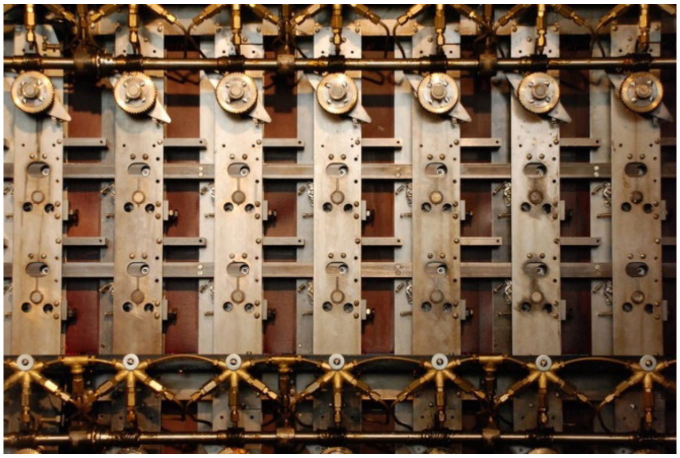
This technological frontier provides the United States an opportunity to innovate and out-maneuver its geopolitical foes while also developing technologies that enhance the general welfare of its citizens. Furthermore, Texas continues to lead the nation as a hub of technological innovation and economic growth. As such, it is incumbent upon the Texas Legislature, especially considering federal inaction, to provide responsible guardrails and uniform standards for AI development, deployment, and use.

This paper will provide a brief history of AI, a description of the technology, examples of various use cases, the regulatory landscape, and policy considerations. With human dignity as a guiding ethos, the authors maintain that technologies like AI are tools which should serve humanity, not the other way around. This principle will serve as the North Star for our discussion.

OVERVIEW OF ARTIFICIAL INTELLIGENCE

Brief History

Despite the seemingly frenetic pace at which discussions about AI went from esoteric theoretics in university laboratories to mainstream ubiquity, the field, study, and technology that is AI has been around for decades. Like many other contemporary technologies, its origin can be tied to the urgency for innovation precipitated by global war. During World War II, the Germans developed an Enigma Machine that allowed the Axis alliance to encrypt and send sensitive messages about key strategies, plans, and other intelligence integral to the Axis' push for global dominance. The machine was considered unbreakable by a human. Alan Turing, the fabled British mathematician and computer scientist, agreed with this assessment, opting to invest precious time developing a machine that would ultimately crack the Enigma code. The machine, "Bombe," would prove invaluable in the Allies' eventual triumph. But as time progressed, this technology would also lay the foundation for the development of modern computing and AI. Just a few years after the end of World War II, Turing (1950) published a seminal article, "Computing Machinery and Intelligence," in which he explored the possibility of intelligent machines and how we might gauge their intelligence—this would later be understood in common parlance as the "Turing Test." He distilled his article into one question: "Can machines think?" ("The Imitation Game" section). Turing pondered whether, like humans, machines can be fed information to solve problems and make decisions. Even as computing technology needed to improve and economize before mass research and development in this frontier could fully take root, Turing's paper captivated the minds of researchers across the globe and catalyzed focused exploration in the realm of machines capable of human-like problem solving abilities. Through this work, Turing became an indispensable spark in the push for research and development in computing.



Note. A photo of the Turing-Welchman Bombe, a cryptography machine used to decode German messages during World War II. Photo from *The Turing-Welchman Bombe*, The National Museum of Computing, n.d. (https://www.tnmoc.org/bombe).

A few years later, in the summer of 1956, a historic conference took place at Dartmouth College, the Dartmouth Summer Research Project on Artificial Intelligence. This served as an inflection point for the research community. The presentation of Logic Theorist, a computer program designed to emulate the problem-solving abilities of humans, established a proof of concept for Turing's musings (Anyoha, 2017). By the end of the conference, which assembled an array of researchers from various disciplines, a name was given to this nascent field: artificial intelligence (Moor, 2006). Further, the participants were resolute in their agreement that AI was an attainable technology. Thus, a second major catalyst emerged that would drive developments in this field.

Over the next two decades, several important advancements occurred in the field of AI. Initially, the cost of computing came down dramatically,[1] allowing for the more efficient storage of information and an increase in speed and accessibility. Additional AI prototypes emerged in the late 1950s and early 1960s—including machine learning algorithms that solved general problems and interpreted spoken language—providing further evidence of promising progress to entice new investors. Such developments were impressive enough to persuade the Defense

[1] Garner (2015) noted that early 1950s computer leasing cost about $200,000 per month (in 2010 inflation-adjusted dollars). These numbers slowly declined over the next two decades to $130,000, $32,000, $25,000, $20,000, $6,500, $5,000, $2,500, $1,900, and finally $1,000 per month to lease an IBM System/3 in 1969.

Advanced Research Projects Agency (DARPA) in 1963 to fund the Massachusetts Institute of Technology (MIT) with a $2.2 million grant to further the study of AI (AIWS, 2021). At the time, "thinking big" with this technology largely revolved around language transcription and translation machines. By 1970, Marvin Minsky, AI-research legend and co-founder of MIT's AI laboratory, asserted in an interview that "in from three to eight years we will have a machine with the general intelligence of an average human being" (Darrach, 1970, p. 58d). Steady progress led to high levels of optimism in the research community at the turn of the decade, considering that only three years prior Minsky had much more conservatively predicted that creating baseline AI might be possible within a generation (p. 2).

Despite breakthroughs in the theory behind AI and computer science, computational power remained too weak to store and process enough data to enable machine intelligence. This aligns with the "Pre-Deep Learning Era" as defined by Moore's Law: the notion that training computation doubled approximately every two years (Intel, 2023). Private and public funding and fanfare surrounding AI experienced something of a lull for the next two decades, with limited tangible progress made aside from (what would later prove important) research endeavors that invested in younger researchers who would eventually devote their lives to groundbreaking work in the field.

The first grand and publicly witnessed breakthroughs in the field came during the 1990s and 2000s. IBM's Deep Blue shocked the world in 1997 when chess world champion and grandmaster Garry Kasparov was defeated by the computer program. While this proof of concept was an important milestone, this illustration of computing abilities captivated the attention of the public and inspired researchers, funders, and the culture at large. Developments only mushroomed from there, with the 2000s punctuated by advancements in new language models, image recognition algorithms, and creative ways to train machines using graphics processing units (GPUs) and unsupervised learning (Karjian, 2023). Importantly, the 2010s was when applications of AI became consistently more visible to the public. Examples include Apple's launch of Siri in 2011, IBM defeating *Jeopardy!* legend Ken Jennings in 2011, DeepMind's AlphaGo defeating a top Go player in 2016, and OpenAI releasing the first iteration of GPT in 2018 (Karjian, 2023).

Ultimately, the history of AI can be best understood as a tension between theoretical developments outpacing computational power due to the limitations imposed by Moore's Law. Thus, once the Deep Learning Era commenced and training computation accelerated exponentially faster than the Pre-Deep Learning Era of the 1950s to 2010, the theoretical rubber finally met the road—and it moved at a breakneck speed. In essence, "we saturate the capabilities of AI to the level of our current computational power (computer storage and processing speed), and then wait for Moore's Law to catch up again" (Anyoha, 2017, para. 8). Even though there are still theoretical limits on computational power, such staggering advancements have occurred in computing power that companies developing training models are encountering more prohibitive constraints to data and energy access. Unlike the previous "hurry up and wait" mentality evinced by the Pre-Deep Learning Era, the current era introduces heightened threats over the ownership of personal data and more competitive demands for scarce dispatchable energy.

This history, and the contemporary convergence of theory and computational power, is well-summed in Neil Postman's ecology of technology theory. Postman is best known for his social critiques of the medium of television and its (negative, in his view) effect on how we process information. As Postman 1998) noted,

> technological change is not additive; it is ecological. I can explain this best by an analogy. What happens if we place a drop of red dye into a beaker of clear water? Do we have clear water plus a spot of red dye? Obviously not. We have a new coloration

> to every molecule of water. That is what I mean by ecological change. A new medium does not add something; it changes everything. In the year 1500, after the printing press was invented, you did not have old Europe plus the printing press. You had a different Europe. After television, America was not America plus television. Television gave a new coloration to every political campaign, to every home, to every school, to every church, to every industry, and so on. (p. 4)

Postman's ecological argument can reasonably be applied to AI. That is, AI does not just add to or subtract from our jobs, schools, and politics, but it fundamentally transforms society. Indeed, "it changes everything" (p. 4).

AI and Culture

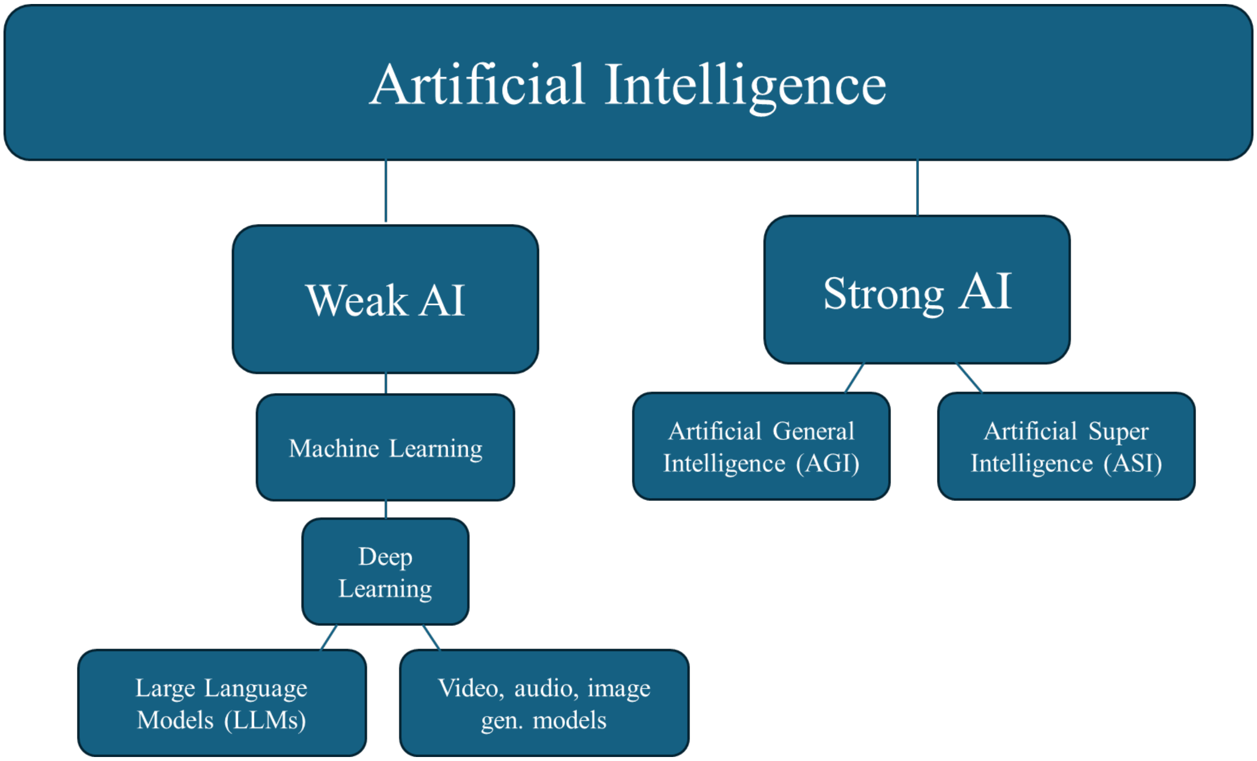
As discussed in the previous section, many of the formative advancements in AI went unnoticed by the public. AI defeating world champions in chess and *Jeopardy!* was culturally remarkable, which played into the well-established narrative of computers and robots outsmarting humans. Take, for example, the blossoming of the science fiction genre during the 20th century. The 1920s and 1930s gave rise to the "Pulp Era," with science-fiction and fantasy magazines being published and distributed en masse for the first time. Next, 1938–1946 marked the Golden Age of Science Fiction, in which realism and psychological depth permeated the genre, shifting from a focus on "super science" gizmos to characters employing them. And finally, the "New Wave Era" transpired during the 1960s and 1970s, with an even more experimental genre that emphasized the psychological and social sciences in relation to the physical sciences. Certainly, fans of the genre were captivated by futuristic depictions of society. But as the negativity bias dictates, humans are wont to latch onto the dystopian, catastrophic depictions of robots and computers with greater salience than the more optimistic applications. This would be on even greater display in Hollywood depictions of AI. The *Terminator* series, *2001: A Space Odyssey*, *Ex Machina*, *I, Robot*, and *Blade Runner* are just a few movies from an extensive menu of examples. To test the veracity of the Breitbart Doctrine, which asserts that "politics is downstream of culture" (Rothschild, 2021, para. 1), one need not look any further than the catastrophizing that is all too common in discussions about AI. While the "Turing Test" as a concept can be hard for the layperson to grasp while reading Alan Turing's seminal work, *Ex Machina* explains this concept much more potently in story form—albeit a story that ends by suggesting AI poses a severe and perhaps unavoidable civilizational threat to humanity. And while works of science fiction have positively influenced the minds of young thinkers to look to the stars with thoughts of *ad astra per aspera*,[2] they have also given many a fundamental misunderstanding of what the field of AI truly is and how the technology works, in addition to a fear of a hypothetical future.

Note. Movies like the *2001: A Space Odyssey* and the *Terminator* franchise are part of a long line of literary and cinematic portrayals of machines and emerging technologies that present a threat to humanity. Photo from *Arnold Schwarzenegger Believes AI Has Made Terminator's Dystopian Future 'A Reality,'* by L. Miller, CBR, July 5, 2023 (https://www.cbr.com/arnold-schwarzenegger-terminator-ai-reality/).

What is AI?

According to the landmark Organisation for Economic Co-operation and Development (OECD) (n.d.) definition, an artificial intelligence system is "a machine-based system that, for explicit or implicit

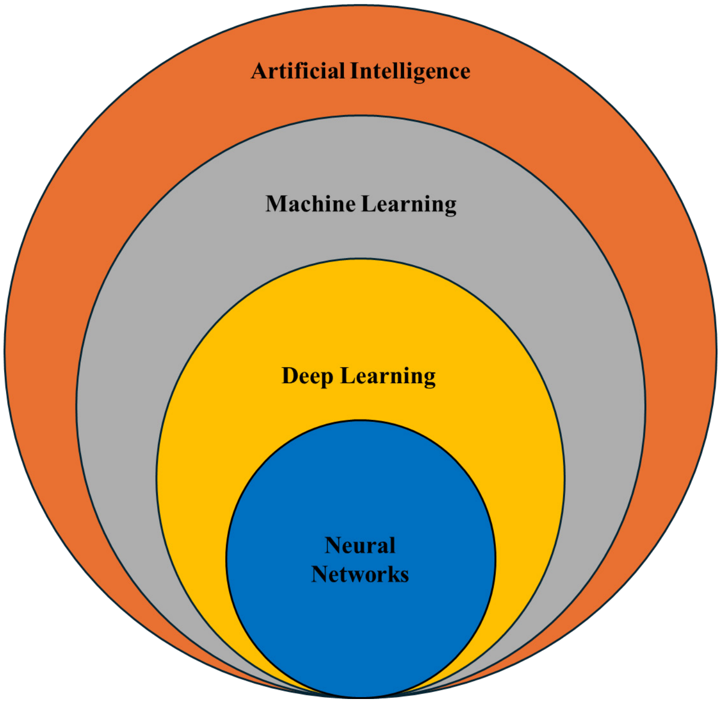
[2] *Ad astra per aspera* translates as "to the stars through hardships" (Merriam-Webster, n.d.). A similar phrase, *per ardua ad astra*, means "through difficulties to the stars" (Dictionary.com, n.d.).


Artificial Intelligence
Weak AI
Strong AI
Machine Learning
Artificial General Intelligence (AGI)
Artificial Super Intelligence (ASI)
Deep Learning
Large Language Models (LLMs)
Video, audio, image gen. models

Note. Chart produced by the authors showing the branches of strong versus weak artificial intelligence.

objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments" ("AI terms & concepts" section). These systems, designed with massive computational power, allow computers to analyze vast amounts of data that would simply not have been possible before. This innovation creates new technologies with unique applications, which the authors explore thoroughly in this paper. This overview is not intended to be an exhaustive treatise on the technological innerworkings of AI, but instead it will provide an overview of the categories and tools that have developed in recent years and their various applications.

The term "AI" should be thought of as the overarching field that sits atop specific functions and technologies that leverage the broader field of AI: "large language models" (LLMs) like Chat GPT, "image generation models" like Midjourney, tools like navigation apps that leverage "machine learning," and even yet-to-be-created "artificial super intelligence" robots that surpass human intelligence. In the same way that mathematics as a field encompasses everything from geometry to trigonometry to calculus, as a field, AI refers to technologies dating as far back as a 1952 "bot" that was programmed to play checkers, and as commonplace as the algorithmically curated content that one sees in a social media feed. Thus, we come to understand that the term "AI" does not necessarily describe particular applications or tools, but an entire field of advanced computational technology. For example, ChatGPT is an LLM that leverages deep learning and machine learning functions, all under the umbrella of weak AI, which is itself a branch of AI broadly. Therefore, calling ChatGPT an LLM is more descriptive and more precise than calling it AI, although it is, indeed, an AI technology.



Note. AI is like a Russian nesting doll insofar as it includes subsets such as machine learning, deep learning, and neural networks. Chart is a slight revision of the original from Artificial Intelligence and Deep Learning Assisted Rapid Diagnosis of COVID-19 from Chest Radiographical Images: A Survey, by D. Sinwar, V. S. Dhaka, B. A. Tesfaye, G. Raghuwanshi, A. Kumar, S. K. Maakar, & S. Agrawal, *Contrast Media & Molecular Imaging*, 2022 (https://doi.org/10.1155/2022/1306664).

HOW DOES AI WORK?

In a colloquial sense, any technology in use today that mimics human thinking could be considered artificial intelligence. As a broad field, AI can be divided into two branches—weak AI and strong AI—which themselves have further branches.

Weak AI

Weak AI, also known as narrow AI or artificial narrow intelligence (ANI), is designed and trained to perform specific tasks. It powers much of the AI technology that we interact with daily. Despite its name, weak AI is powerful, enabling robust applications such as Apple's Siri and self-driving vehicles. Additionally, common features in personal computing, like traffic prediction in GPS navigation systems and spam filtering in email inboxes, are examples of weak AI (Stryker & Kavlakoglu, 2024). Narrow AI is perhaps a more descriptive term, given that this category of AI concerns the applications of technology to perform specific and narrowly defined tasks. These innovations have emerged through the advancement of machine learning, further discussed below.

One of the most powerful illustrations of weak AI comes from John Searle's Chinese room thought experiment (Cole, 2024). Imagine you are in a room and do not know how to write nor speak Chinese. A native Chinese speaker sits outside your room, unable to see you. You are handed a book with instructions on what to do when the person outside relays to you a piece of paper with Chinese writing on it. The instructions tell you to find a certain symbol in the book that is included in the note, and the instructions direct you to write down a different Chinese symbol on a new piece of paper in response. While you do not know a shred of Chinese, if you precisely follow the instructions written in the book, the person outside the room who is feeding you the Chinese writing will think you understand what they are saying because you respond by stringing together the correct Chinese symbols. In this way we come to understand how weak AI tools like ChatGPT, when correctly prompted, can adhere to their structured set of rules and give back "right" answers despite possessing no understanding of the answers it is producing.

Level of AGI	**Narrow** *Clearly scoped task or set of tasks*	**General** *Wide range of non-physical tasks, including metacognitive tasks like learning new skills*
Level 0: No AI	**Narrow Non-AI** *calculator software; compiler*	**General Non-AI** *human-in-the-loop computing, e.g., Amazon Mechanical Turk*
Level 1: Emerging *equal to or somewhat better than an unskilled human*	**Emerging Narrow AI** *simple rule-based systems, e.g., recommendation systems on Netflix*	**Emerging AGI** *ChatGPT, Bard, Llama 2, Gemini*
Level 2: Competent *at least 50th percentile of skilled adults*	**Competent Narrow AI** *smart speakers like Siri and Alexa; IBM's Watson; SOTA LLMs for a subset of tasks (e.g., simple coding)*	**Competent AGI** *not yet achieved*
Level 3: Expert *at least 90th percentile of skilled adults*	**Expert Narrow AI** *spelling & grammar checkers like Grammarly; generative image models like Dall-E 2*	**Expert AGI** *not yet achieved*
Level 4: Virtuoso *at least 99th percentile of skilled adults*	**Virtuoso Narrow AI** *Deep Blue, AlphaGo*	**Virtuoso AGI** *not yet achieved*
Level 5: Superhuman *outperforms 100% of humans*	**Superhuman Narrow AI** *AlphaFold, AlphaZero, StockFish*	**Artificial Superintelligence (ASI)** *not yet achieved*

Note. Chart reproduced from *Levels of AGI for Operationalizing Progress on the Path to AGI*, by M. R. Morris, J. Sohl-dickstein, N. Fiedel, T. Warkentin, A. Dafoe, A. Faust, C. Farabet, & S. Legg, Cornell University, 2024 (https://doi.org/10.48550/arXiv.2311.02462).

Strong AI

Strong AI encompasses artificial general intelligence (AGI) and artificial superintelligence (ASI). AGI, also known as general AI, is a theoretical form of AI where a machine possesses intelligence equal to that of humans. This form of AI would be self-aware, with consciousness and the ability to solve problems, learn, and plan. ASI, or artificial superintelligence, would surpass human intelligence and capabilities (Stryker & Kavlakoglu, 2024). AGI and ASI are both forms of AI that are strictly theoretical with no practical examples in use today, but computer scientists and entrepreneurs continue to push the limits of modern technological research and innovation.

In relation to the Chinese room thought experiment, instead of blindly following instructions and producing outputs with zero context or critical thought applied, such strong AI would be able to learn Chinese from a textbook, understand it, and respond creatively, free of narrowly ordained bounds.

Much like in the autonomous vehicle space, where industry has developed stages of vehicular autonomy—ranging from Level 0 cars with no self-driving capabilities to Level 5 cars that can operate fully without a steering wheel—companies like OpenAI and Google have developed scales to classify weak versus strong AI (Dunmoyer, 2024). For example, researchers at Google DeepMind proposed the framework of five levels of AI (*see* table above).

Machine Learning

Machine learning (ML) is a branch of weak AI at the heart of many recent AI advancements. ML is a paradigm that enables computers to learn from data and improve their performance on tasks without being explicitly programmed for every contingency. The process involves feeding the machine vast amounts of data and using algorithms to analyze and infer patterns from data, thus enhancing the machine's ability to perform tasks or make decisions. The efficacy of ML models scales with the volume and quality of data they process, underpinning their adaptability (IBM, n.d.-a). There are four ways that ML models are trained: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning.

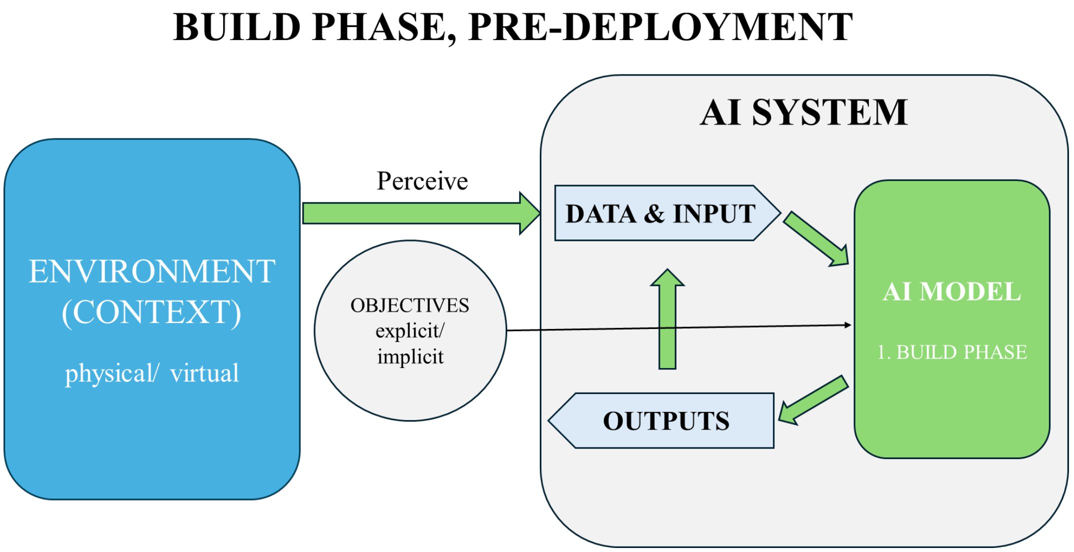
Supervised Learning

In supervised learning, models are trained on a data set with predetermined outcomes. This approach is akin to learning with a teacher overseeing the process, where the algorithm makes predictions and adjusts based on the accuracy of its outcomes against the known labels (Google Cloud, n.d.-a). Data used in supervised learning model training are known as structured data. Structured data are well-organized, easily interpretable data that are organized into rows and columns within databases, labeled such that machine learning algorithms can clearly identify, consistent input data. For instance, structured data can include numerical values, categorically labeled photos or videos, and timestamps, all of which can be used to train models to recognize patterns, make predictions, and improve decision-making processes (AWS, n.d.-a).

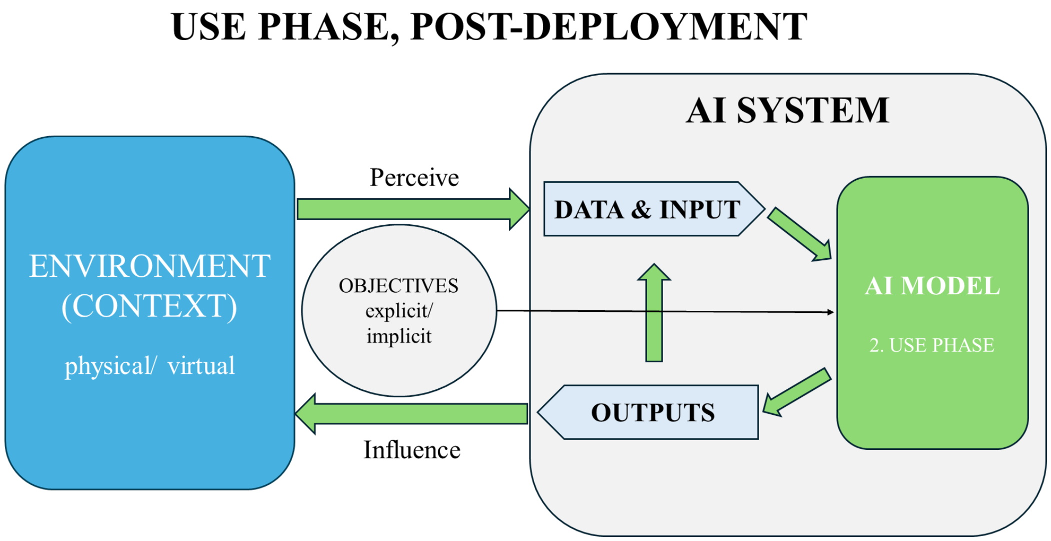
Various mathematical and statistical methods used in supervised learning include linear regressions, where relationships between dependent and independent variables are mapped for predictions. For example, in healthcare, models trained via supervised learning can help diagnose diseases by analyzing medical images or patient records, which improves early detection and personalized treatment plans. In finance, these models can predict stock prices, assess credit risks, and detect fraudulent activities by learning from historical data. In natural language processing models, supervised learning is used for tasks like sentiment analysis, language translation, and spam detection, enabling systems to understand and generate human language accurately (IBM, n.d.-a).

Unsupervised Learning

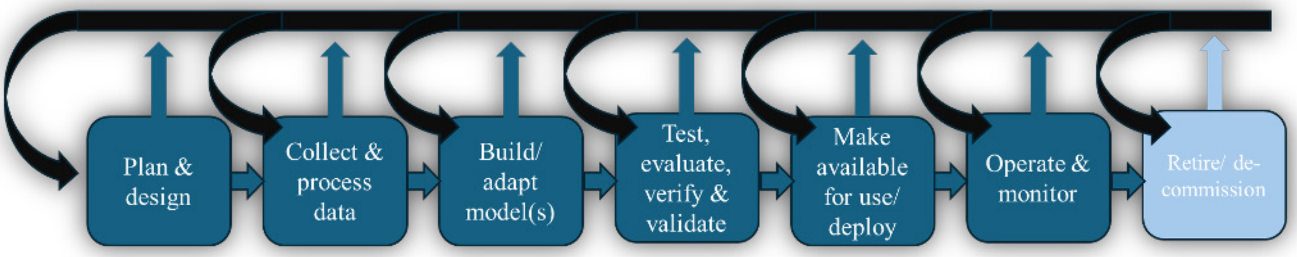
In contrast to supervised learning, unsupervised learning involves training models on data without explicit labels. Model training that utilizes unsupervised learning takes unstructured data, absent labels and categories, and works to discern patterns and trends that inform the information, conclusions,



Note. Chart of an AI system in the build phase reproduced with authors' revisions from *OECD AI Principles Overview*, by Organisation for Economic Co-operation and Development, n.d. (https://oecd.ai/en/ai-principles).



Note. Chart of an AI system in the use phase reproduced with authors' revisions from *OECD AI Principles Overview*, by Organisation for Economic Co-operation and Development, n.d. (https://oecd.ai/en/ai-principles).



Note. Chart of an AI system lifecycle reproduced with authors' revisions from *OECD AI Principles Overview*, by Organisation for Economic Co-operation and Development, n.d. (https://oecd.ai/en/ai-principles).

or data that are outputted. Unstructured data are defined as information lacking a predefined format or organizational structure, making them more challenging to analyze and process compared to structured data. Examples of unstructured data include text, images, videos, and emails, and AWS (n.d.-a) estimates that 80%–90% of data are unstructured. Unstructured data are useful in scenarios where storing data in a structured format is costly or inefficient.

Unsupervised learning is widely used in various applications, including market analysis, customer segmentation, anomaly detection, computer vision, and exploratory data analysis (IBM, n.d.-b). Unsupervised learning is most helpful in gaining insights when labeled data are scarce or expensive to obtain or store (IBM, n.d.-a).

Semi-Supervised Learning

A middle ground between supervised and unsupervised learning is called semi-supervised learning. According to IBM (n.d.-a),

> during training, it uses a smaller labeled data set to guide classification and feature extraction from a larger, unlabeled data set. Semi-supervised learning can solve the problem of not having enough labeled data for a supervised learning algorithm. It also helps if it's too costly to label enough data. ("Semi-supervised learning" section)

Reinforcement Learning

Reinforcement learning (RL) is a fourth model of machine learning distinguished by focusing on learning optimal behaviors through trial-and-error interactions with an environment. An RL agent learns from feedback—rewards or punishments—based on its actions, gradually honing its policy for decision-making. This method has profound implications in areas requiring strategy and adaptability, such as gaming and robotics. Of note, the IMB Watson system that beat *Jeopardy!* champions Ken Jennings and Brad Rutter in 2011 "used reinforcement learning to learn when to attempt an answer (or question, as it were), which square to select on the board, and how much to wager—especially on daily doubles" (IBM, n.d.-a, "Reinforcement machine learning" section).

Deep Learning

Deep learning (DL), a subset of ML, elevates the complexity of model architectures through structures called neural networks designed to simulate the learning processes of the human brain. Unlike traditional ML models that might require manual correction and struggle with unstructured data, DL models thrive on large, diverse data sets and can autonomously refine their predictions through repetition. Whereas traditional machine learning models may use one or two computational layers, "d eep learning models use three or more layers—but typically hundreds or thousands of layers—to train the models" (Holdsworth & Scapicchio, 2024, para. 2). This self-sufficiency and capacity to handle intricate data structures makes DL particularly potent for tasks involving image and pattern recognition, natural language processing, and more.

Neural Networks

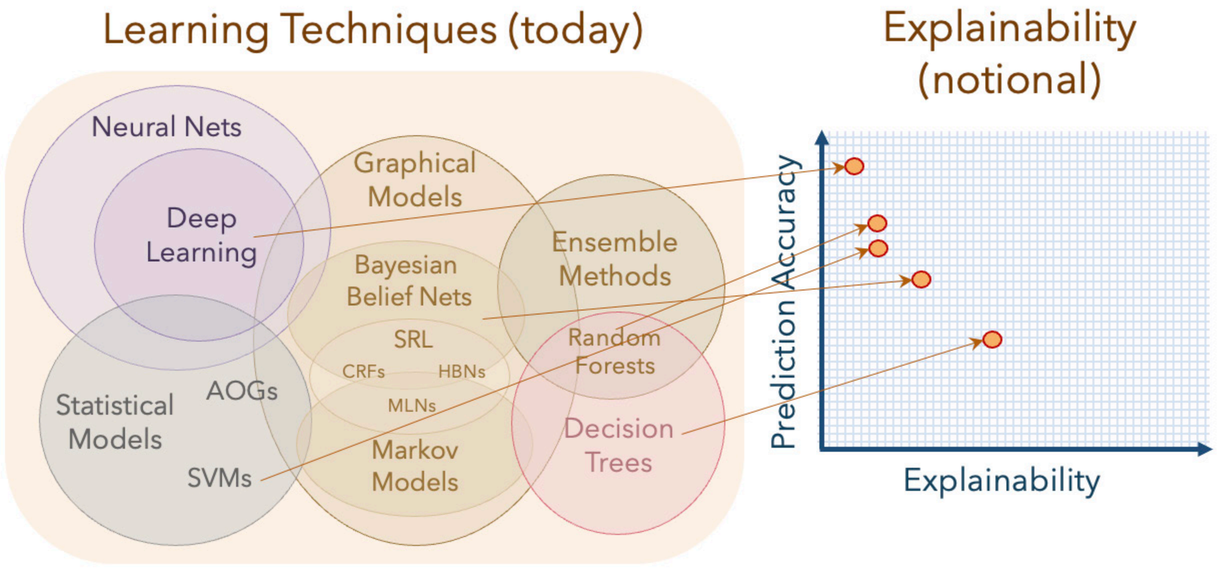
Neural networks (NNs), a subset of DL, provide the structure for the most advanced AI models that are in operation today. NNs are the building blocks by which deep learning AI models are constructed. Akin to the human brain, NNs consist of layers of programmed "neurons" that process data such as images or text, apply weights to the data assessed, and then provide an output (IBM, n.d.-c). The accuracy of AI models depends on the quality of their training and requires significant time and monitoring to ensure proper outputs are achieved.

Training a neural network is a process that can be compared to the education of a young child. For example, imagine that you want to teach a child how to identify apples and oranges. You show them many pictures of apples and oranges, explaining which is which. The first time the child sees a new picture, they might make a guess. Sometimes they might say an orange is an apple and vice versa. You correct them, explaining why their guess was wrong, such as "No, this is an orange because it is round and has thicker skin." The child adjusts their understanding based on your correction, learning that a round fruit with thick skin is more likely to be an orange. In the context of an AI model, this approach utilizes structured data to label accurate outputs. Over time, with more pictures and corrections, the child gets better at distinguishing between apples and oranges, making fewer mistakes as they understand the differences better.

In this analogy, the child represents the neural network, the pictures of fruit are the training data, the corrections and explanations are the feedback from the model trainer, and the adjustments in understanding are like how the neural network adjusts its weights during training. Through repeated exposure and correction, the neural network—like the child—learns to make more accurate predictions. This educational process unfolds hundreds of times during the training of a given model, with each training set designed to achieve the desired use cases for each model.

Foundation Models

Expanding on the building blocks of deep learning neural networks, foundation models serve as the backbone for a wide array of AI applications. These models are characterized by their large-scale architecture and their ability to perform multiple tasks without being specifically trained for each one. Built using extensive data sets and deep learning



Note. The black box problem is part of a broader challenge for AI models that vary in their degree of prediction accuracy and explainability. Xu et al. (2019) argued that "explainability of machine learning models appear inverse to their prediction accuracy" (p. 565). Photo from *Explainable Artificial Intelligence (XAI)* [PowerPoint slides] , by D. Gunning, DARPA, August 11, 2016 (https://www.darpa.mil/attachments/XAIIndustryDay_Final.pptx).

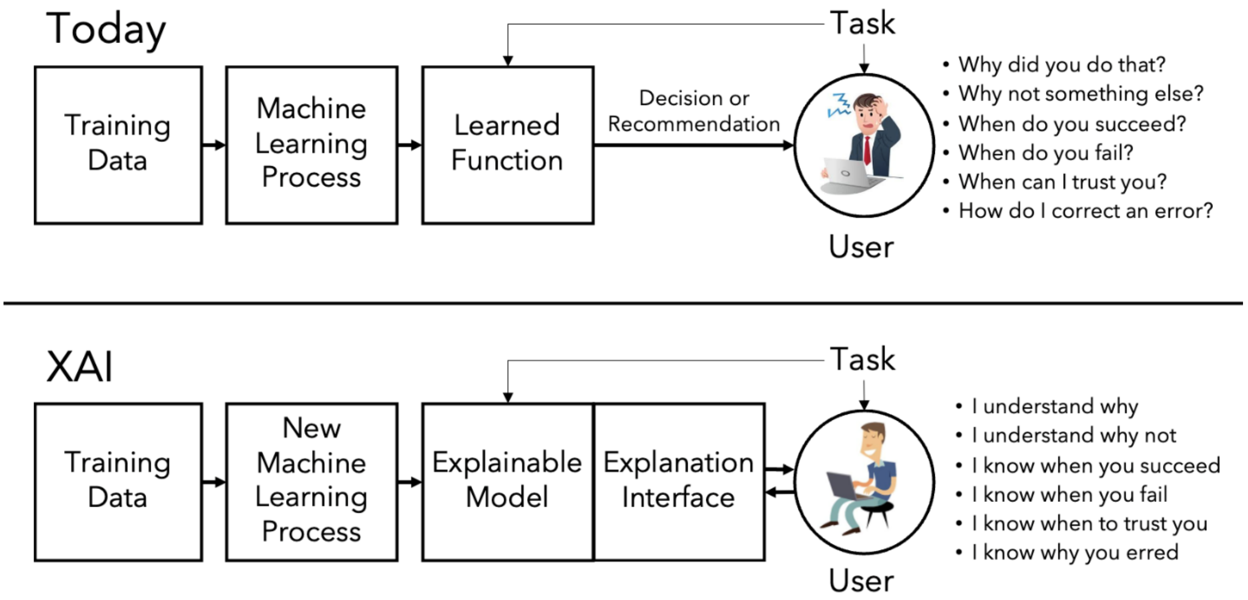
techniques, foundation models are pre-trained on a diverse range of data, allowing them to understand and generate human language, recognize images, and even predict molecular structures. Their versatility and scalability make them foundational, providing a robust platform upon which more specialized AI applications can be developed. This transformative approach enables rapid innovation and deployment across various fields, from natural language processing to computer vision and beyond (AWS, n.d.-b).

Large Language Models

One type of foundation model is a large language model (LLM). LLMs are advanced artificial intelligence systems designed to understand, generate, and manipulate human language in a sophisticated manner. They are built using deep learning techniques, particularly neural networks, which allow them to learn patterns and relationships in vast amounts of text data. By training on diverse and extensive data sets, LLMs can perform a wide range of language-related tasks such as translation, summarization, text generation, sentiment analysis, answering questions, and more. These models are "large" because they contain billions of parameters, which are the adjustable weights within the neural network that help the model make predictions or generate responses based on the input it receives (Cloudflare, n.d.).

ChatGPT, developed by OpenAI, is a prime example of an LLM. It is designed to generate human-like text based on the prompts it is given, making it capable of engaging in conversation, providing detailed explanations, and even composing creative content. ChatGPT operates probabilistically by predicting the next letter and the next word in a sequence of words, using the context provided by the input to produce coherent and contextually appropriate responses. Its ability to generate natural language text has numerous applications, including customer support, educational tools, content creation, and personal assistance. The model's responses are informed by patterns it has learned during its training phase, which included diverse sources of text from books, websites, and other written material (OpenAI, n.d.-a).

Outside of LLMs, there are other notable examples of foundational models. For example, generative AI (GAI) is a broad term at the frontier of public-facing AI and AI research. GAI models can create new,



Note. XAI is being studied to produce more explainable models, maintain high prediction accuracy, and provide greater understanding and trust of AI systems. Photo from *Explainable Artificial Intelligence (XAI)* [PowerPoint slides], by D. Gunning, DARPA, August 11, 2016 (https://www.darpa.mil/attachments/XAIIndustryDay_Final.pptx).

original content including text, images, video, and music, offering exciting possibilities for creativity and design. Second, audio models can generate humanoid speech or music, revolutionizing how we interact with technology through voice assistants and new forms of digital entertainment. Third, scientific models are tailored for specific domains like biology or chemistry. These models help in drug discovery, protein structuring, and other scientific inquiries. Fourth, computer vision models with the ability to "see" can interpret and analyze images, facilitating advances in facial recognition, medical imaging, and autonomous vehicles. Finally, companies are also developing smaller, more bespoke AI models that may boost performance and efficiency (Gent, 2023).

What is a Black Box?

One of the more intriguing aspects of artificial intelligence is the notion of the black box. The black box analogy highlights the difficulty in discerning how specific inputs are transformed into outputs within the AI model. In essence, this analogy describes a system "whose inputs and operations aren't visible to the user or another interested party" (Awati & Yasar, 2024, para. 1). Therefore, one is left unable to see how a deep learning system decides in response to a prompt or query. This lack of transparency can be attributed to the complexity of the AI networks, which involve numerous layers of interconnected nodes and a vast number of parameters that adjust based on the training data. In fact, some go as far as to describe the process of NNs as alchemy more than science due to this phenomenon. Strong parallels can be drawn from the human brain.

For example, as children we learn to recognize the difference between a cat and dog, or the letters "a" and "b," through examples that appeal to the trend-finding machine that is the human brain. Having been exposed to the qualities that make a "cat" a cat, a "dog" a dog, an "a" an a, or a "b" a b, we develop the decision protocols that allow us to categorize all manner of experiences automatically—for example, the ability to ascertain that a photoshopped picture has the head of a cat but the body of a dog. This is the easy part. But the black box analogy in this context concerns explaining *how* we complete these automatic and unconscious tasks. We cannot point to the formative instructional moment (the input) that led us to understand what makes a "cat" a cat. Even with advancements in neuroscience, it is virtually impossible to articulate the processes (i.e., the black box) our brain underwent to reach this conclusion. According to Rawashdeh,

> it's one of those weird things that you know, but you don't know how you know it or where you learned it. … It's not that you forgot. It's that you've lost track of which inputs taught you what and all you're left with is the judgments. (Blouin, 2023, para. 1)

This is precisely how researchers and technologists have come to understand the illusory notion of the black box in an AI system. Consequently, even the developers and CEOs of major tech corporations that own these models are unable to explain why an AI system makes a particular decision or prediction (Terech, 2024; Field, 2024; Bagchi, 2023).

A practical example of the black box problem is found in the application of AI technologies in facial recognition systems. The black box of a facial recognition system is the intricate network of neurons that scan large, unstructured data sets of images of faces to identify features within the pictures—details such as male or female, eye color, hair color, complexion, age, etc. This information then allows the AI model to recognize patterns and build information profiles based off the images. When an input—in this case, an image of a face—is given to the model, it then assesses the qualities of the image and produces an output gleaned from the data it was trained on. Even though the system might accurately assess the person in the image it was fed, the developers of the model do not exactly understand how the model came to recognize the image and match it to the training data.

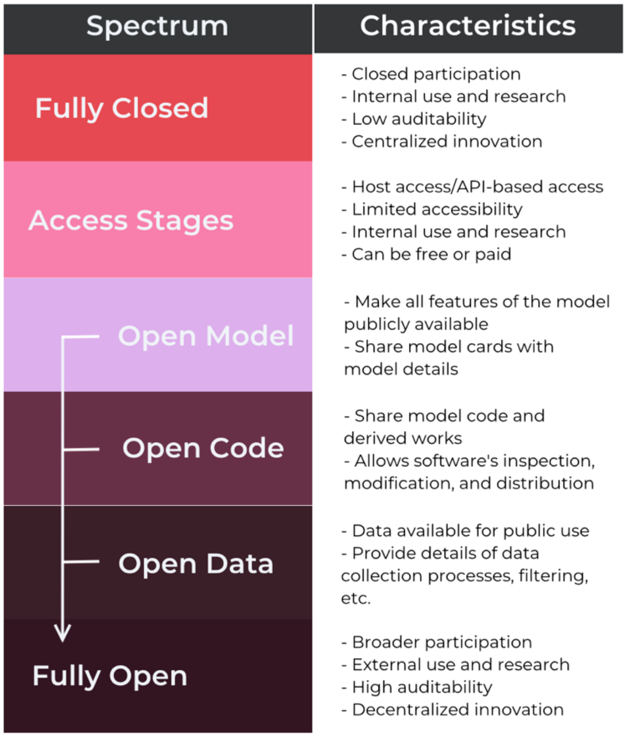
The counterpart to the black box phenomenon comes in the form of explainable AI, or XAI. An XAI system is "created in a way that a typical person can understand its logic and decision-making process" (Awati & Yasar, 2024, para. 3). Indeed, most AI models are developed by organizations and institutions promoting AI trustworthiness, model accuracy, fairness, and transparency in outcomes for AI-enabled decision-making (IBM, n.d.-d). Yet many in the field of transparent AI systems fundamentally reject the notion that the black box problem is an unavoidable reality society must tacitly accept, arguing that it is instead a design feature that can be improved and, ultimately, explained (Perrigo, 2024; Aich & Burch, 2023; Bagchi, 2023; Blouin, 2023; Capps, 2023; Xu et al., 2019). Such transparency is critical to prudently leverage AI technologies for human-in-the-loop systems. Xu et al. (2019) provide useful analogies. For example,

> a medical doctor needs to understand what pathological features in the input data were guiding the algorithm before accepting auto-generated diagnosis reports. A maintenance engineer needs to understand which abnormal phenomena were captured by the inference algorithm before following the repair recommendations. A financial investor wants to understand what influencing factors were regarded as the critical ones by the system algorithm before making the final investment decision. We have to verify that the AI inference works as expected, because wrong decisions can be costly and dangerous. (p. 566)

Open Source vs. Closed Source

The distinction between "open source" and "closed source" in AI technologies pertains to their accessibility and the freedom to use, modify, and distribute the underlying code or model. The difference between these two designs is at the heart of the future of AI technologies. It is also a part of ongoing discussions about whether all models should be open source (for the sake of transparency and the future of humanity) or if intellectual property rights supersede and will drive innovation for future applications of more closed AI systems. The distinction is not necessarily binary—that is, open or closed—but is rather a gradient of policies and procedures from the model developers that classify each model (Luna, 2024).

The Open and Closed AI Spectrum



Spectrum	Characteristics
Fully Closed	- Closed participation - Internal use and research - Low auditability - Centralized innovation
Access Stages	- Host access/API-based access - Limited accessibility - Internal use and research - Can be free or paid
Open Model	- Make all features of the model publicly available - Share model cards with model details
Open Code	- Share model code and derived works - Allows software's inspection, modification, and distribution
Open Data	- Data available for public use - Provide details of data collection processes, filtering, etc.
Fully Open	- Broader participation - External use and research - High auditability - Decentralized innovation

Note. AI openness (or closedness) is not binary. Rather, it can be viewed along a spectrum. Photo from *The Open or Closed AI Dilemma*, by A. Luna, Bipartisan Policy Center, May 2, 2024 (https://bipartisanpolicy.org/blog/the-open-or-closed-ai-dilemma/).

Open Source

Open-source AI systems are developed with the principle of transparency in mind. The source code for these systems is made freely available to the public for any developer to access and utilize the software. Some versions of open-source software allow editing and modification as well. Open-source models are designed to foster innovation, since they enable a wide range of contributions, leading to rapid advancements and diverse applications. Developers can build upon existing frameworks, share improvements, and fix bugs, which results in more robust and versatile AI solutions (Open Source Initiative, 2024).

Notably, Meta decided to open source its LLM, called LLaMa. This is in contrast to leading AI titans like OpenAI and Google, which claim "an unfettered open-source approach is dangerous"—dangerous to whom, though, is unclear.

Open-source models also provide transparency and promote trust and accountability, since the algorithms and methodologies are open for scrutiny. Engineers and users alike are allowed to dissect the innerworkings of any given open-source model to better understand how it functions and how it produces the output that it does. Open-source models allow for independent experts to examine the models and provide warnings to users if they find any issues that cause concern. This openness builds greater trust in AI systems and provides access to communities that might not be able to invest in proprietary AI solutions.

Those calling for open-source models include academic research institutions, governmental organizations like the European Union (EU), industry leaders like Elon Musk, and consumer privacy organizations like the Mozilla Foundation. Each of these groups recognizes the need for transparency and accountability. Examples of popular open-source AI platforms include TensorFlow and PyTorch, which have become leading foundational tools in both academic research and industry applications (MIT Technology Review, 2024). Notably, Meta decided to open source its LLM, called LLaMa. This is in contrast to leading AI titans like OpenAI and Google, which claim "an unfettered open-source approach is dangerous"—dangerous to whom, though, is unclear (Metz & Isaac, 2023, para. 7).

Closed Source

In contrast, closed-source AI systems are proprietary, with the source code kept confidential and controlled by the organization that developed it. These systems are often deployed in commercial products, and access to their inner workings is restricted to protect intellectual property and maintain competitive advantage, while access to their features is available to the user. While closed-source AI can lead to significant advancements driven by substantial investment and resources from private companies, it also limits collaboration and external contributions. Users of closed-source AI must rely on the developing company for updates, bug fixes, and enhancements, which can result in slower innovation and less adaptability. Additionally, the opacity of closed-source systems can raise concerns about bias, security, and trust, as the underlying algorithms and data usage are not transparent (Luna, 2024).

Prominent examples of closed source AI include many of the advanced AI products developed by tech corporations like Google, Microsoft, IBM, and OpenAI, which are used in various commercial applications from cloud services to advanced analytics. These companies emphasize the need to protect intellectual property and ensure security by using closed systems. These companies also claim that by doing so, they are able to provide tailored, high-quality services to their customers that they would not be able to otherwise (Mancebo, 2023).

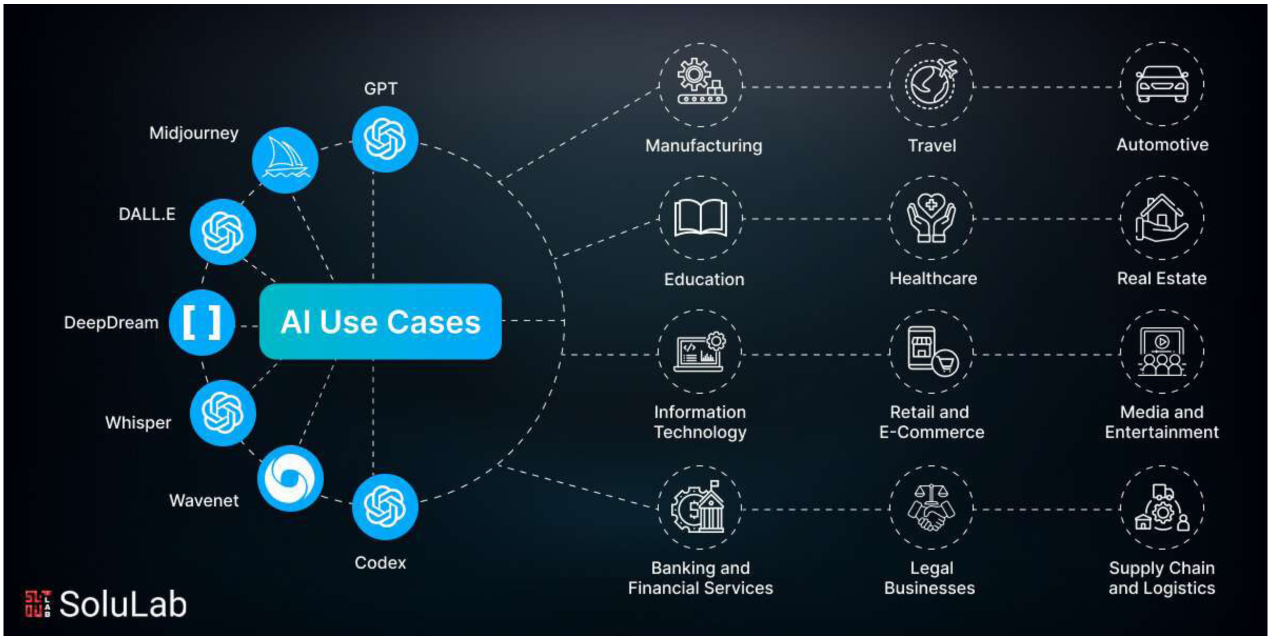
In conclusion, open- and closed-source AI models each offer distinct advantages and challenges, contributing uniquely to the development and application of AI technologies. While open-source models promote transparency, collaboration, and accessibility, closed-source models provide control, security, and tailored solutions. However, it is important to recognize that most AI models exist on a spectrum rather than as binary categories. Many AI systems incorporate elements of both open and closed paradigms, leveraging open-source components for foundational technologies while maintaining proprietary enhancements for competitive differentiation. This gradient approach allows for a balance between innovation and control, ensuring that AI development can benefit from the strengths of both open- and closed-source methodologies.

CASE STUDIES

How AI is used varies from one context to another. That is, how it is used in healthcare is different than in education, law enforcement, or finance. To gain a holistic understanding of the broad catchall term that is AI, it is crucial to explore its unique, and at times divergent, applications. Whether crafting a uniform regulatory framework or developing sector-specific legislation, case studies offer insights necessary to spot issues, understand use cases, and help inform prudent policy solutions. This section highlights nearly a dozen case studies of how AI is used in various sectors, discusses opportunities and threats, and highlights areas of special emphasis that may guide policymakers.

Nefarious Applications

The list of nefarious AI use cases could warrant a separate paper altogether. To emphasize the areas most ripe for legislative oversight, this section will consider nefarious applications in national security, warfare, elections, democratic disruption, critical infrastructure, cybersecurity, deepfakes, and sexual exploitation. In addition to exploring how AI can be deployed by nefarious actors to inflict harm on victims, this section will also present AI applications for defending against or more efficiently remedying harms that may befall victims.



Note. AI has many potential use cases discussed at length in the following section. Photo from *AI Use Cases and Applications in Key Industries*, by K. Wadhwani, SoluLab, n.d. (https://www.solulab.com/ai-use-cases-and-applications/).

National Security

Because AI is a technology with civilizational implications, it is only appropriate that national security matters serve as the inaugural case study. While American national security is a concept as old as our nation, author and journalist Walter Lippmann (1943) is credited with coining the term, defining it as "when [a nation] does not have to sacrifice its legitimate interests to avoid war and is able, if challenged, to maintain them by war" (p. 51). In large part, America's ability to maintain its global authority and national security hinges on its technological prowess, and the race for AI supremacy is in many respects the race to geopolitical supremacy. Our adversaries know this, as evinced by strategic partnerships between nations like Russia and China, which are working together to advance applications of AI to kinetic and cyber warfare (Zhou, 2024). Thus, while the instinct to create a robust risk-based regulatory framework can be a prudent approach, America must compete with and win out over adversaries developing artificially intelligent weaponry.

When it comes to AI research and development, much like American advanced weaponry programs, exceptions likely must be granted to military development and applications where appropriate partnerships and oversight exist. Even the EU AI Act provides these types of exceptions (Garrod et al., 2024). American political leaders would be wise to observe that the fate of humanity may well revolve around which nation first develops the most sophisticated AI-powered offensive and defensive military systems. The following section details extant and emerging applications of AI in the realm of national security. And much like with the development of nuclear weaponry, each use case poses tremendous advantages and disadvantages depending on the motives of the actor wielding such tools.

Autonomous Weapons Systems (AWS)

Autonomous weapons systems have captured the attention and research and development (R&D) dollars of superpowers across the world. Informally referred to as "killer robots" (or perhaps even more ominously and gruesomely, as "slaughterbots"), AWS leverages AI to "identify, select, and kill human targets without human intervention" (Autonomous Weapons, n.d., para. 1). With AWS, the human is out of the loop. While unmanned weaponry, such as drones, have been a central feature of military operations for years, the key difference is that those weapon systems still require a human operator to "pull the trigger" and take a life. At a defense industry conference in late 2023, U.S. Deputy Secretary of Defense Kathleen Hicks offered her vision for the "Replicator initiative," which aims "to field attritable autonomous systems at scale of multiple thousands, in multiple domains, within the next 18-to-24 months" (Hicks, 2023, para. 82). There have been reports that Russia has deployed AWS technology against Ukraine via a kamikaze drone purportedly capable of being fed an image that triggers real-time AI-powered recognition to autonomously detect and execute the target (Farrell, 2023). Furthermore, China has long been at work building out AWS capabilities and formally outlined its strategic objectives to incorporate a new generation of AI into military operations in its 2017 "A Next Generation Artificial Intelligence Development Plan" (The State Council of the People's Republic of China, 2017a/2017).

Presently, with the increase in unmanned aerial vehicles (UAVs) deployed in warfare, this is the most common vector for AWS abroad. But nations like the United States and Russia raised the specter of incorporating autonomous functions into nuclear weapons systems. Ostensibly, Russia has already developed a nuclear-armed torpedo that can autonomously locate and eliminate its target (Kallenborn, 2022). It is worth noting that perhaps the biggest problem with nuclear AWS is the propensity for error. For an autonomous nuclear weapons system designed to detect and retaliate against an incoming nuclear missile, just one error could lead to civilizational decimation. Given the lack of nuclear warfare data to train these systems on (of which the authors are grateful for said data paucity), the highly dynamic environment of nuclear conflict—with fluid geopolitical circumstances, missile tests, biased satellite imagery, etc.—and the heightened vulnerability associated with data poisoning, adversarial deception, and cyberattacks, even a nuclear AWS that could operate with 100% confidence is ripe for cataclysmic crisis. Thus, American policymakers have a duty to address this paramount concern not only regarding the domestic adoption of AWS technology, but also how to address it with international adversaries and allies alike. Ultimately, the strategic advantage of AWS is that machines can make decisions in milliseconds, whereas humans can take seconds or even minutes, and winning wars and physical conflicts requires the utmost timeliness. But that same advantage means civilizational decimation could now occur in milliseconds, barring the use of human judgment in our most destructive weapons systems.

Note. AI-powered drones and other UAVs are emerging tools for espionage and warfare. Photo from Artificial Intelligence and War, *The Economist*, September 5, 2019 (https://www.economist.com/leaders/2019/09/05/artificial-intelligence-and-war).

Democratic Disruption

Few things have united the world's largest private sector AI competitors like the push to address the deceptive use of AI in elections. X (formerly known as Twitter) and OpenAI—along with 18 other big tech CEOs—put their differences aside and signed onto the "Tech Accord to Combat Deceptive Use of AI in 2024 Elections" (AI Elections Accord, 2024a).[3] They agreed to eight commitments to fight some of the biggest electoral threats, declaring,

> [deceptive AI election content] consists of convincing AI-generated audio, video, and images that deceptively fake or alter the appearance, voice, or actions of political candidates, election officials, and other key stakeholders in a democratic election, or that provide false information to voters about when, where, and how they can lawfully vote. (AI Elections Accord, 2024b, para. 2)

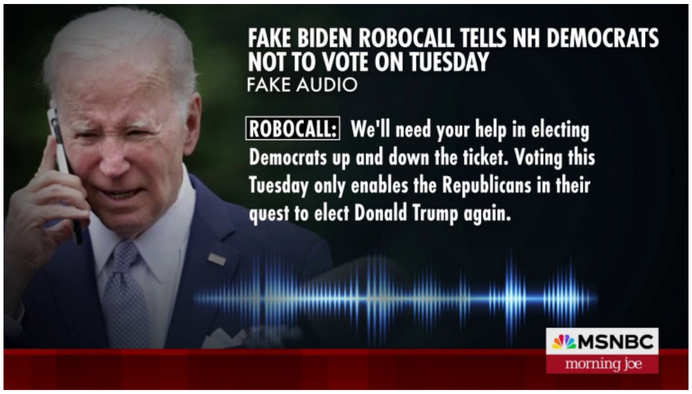
[3] As of February 16, 2024, the 20 signatories were Adobe, Amazon, Anthropic, Arm, ElevenLabs, Google, IBM, Inflection AI, LinkedIn, McAfee, Meta, Microsoft, Nota, OpenAI, Snap Inc., Stability AI, TikTok, Trend Micro, Truepic, and X (AI Elections Accord, 2024a).

Officials have identified threats from the usual suspects (Russia, China, and Iran), as well as from smaller players like Cuba, illustrating the ease and lack of sophistication involved in levying successful attacks on local, state, and national elections.

Encouraging and unifying as this accord may be, critics were quick to point out the unenforceability of the pledge, its entirely voluntary nature, its lack of gauges or reporting to determine meaningful progress, and its disregard for other key threats to elections transpiring on their platforms or through their services (Norden & Harris, 2024).

The threat AI poses to free and fair elections demands private action in concert with strong legal enforcement. A non-exhaustive list of these threats includes disinformation and misinformation through deepfakes, amplification of discord, psychological operations (psyops), targeted influence campaigns, foreign interference, and the erosion of trust in information and institutions more broadly. Already, America has encountered indelible harm to its democratic process due to the nefarious use of AI tools.

For example, prior to the January 2024 New Hampshire primary election, a seemingly real robocall from President Joe Biden urged Democratic voters in New Hampshire to refrain from voting in the primary election and save their vote for the November general election (NBC News, 2024). It is estimated that 5,000 to 25,000 people received the deepfake robocall, representing a significant share of the 125,000 New Hampshire Democrats that participated in the primary election (Seitz-Wald, 2024). For $150, an unsophisticated Democratic political consultant hired a New Orleans magician to create the deepfake, illustrating the ease and power AI can have in swaying elections. Even though the political consultant was later indicted, the damage was done—although he claimed to be the hero, saying, "Maybe I'm a villain today, but I think in the end we get a better country and better democracy because of what I've done, deliberately" (Ramer & Swenson, 2024, para. 12).



FAKE BIDEN ROBOCALL TELLS NH DEMOCRATS NOT TO VOTE ON TUESDAY

FAKE AUDIO

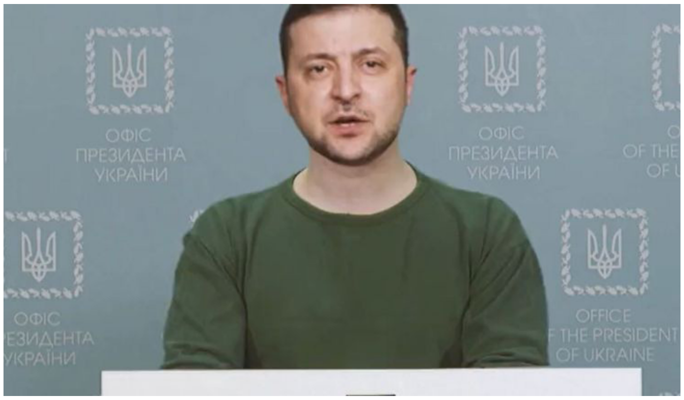
ROBOCALL: We'll need your help in electing Democrats up and down the ticket. Voting this Tuesday only enables the Republicans in their quest to elect Donald Trump again.

Note. Deepfaked audio, photos, and videos present serious challenges to election integrity. Photo from *'Is This Going to Be the Deepfake Election?': Analyzing AI's Potential Influence on 2024* [Video], by Morning Joe, MSNBC, January 26, 2024, 0:32 (https://www.msnbc.com/morning-joe/watch/is-2024-going-to-be-the-deepfake-election-chris-krebs-is-gravely-concerned-202994757875).

Furthermore, U.S. intelligence officials have issued an unprecedented number of warnings to 2024 political candidates and government leaders who have been targeted by American adversaries, including attempts to hack personal, campaign, and official government devices (Barrett et al., 2024). Officials have identified threats from the usual suspects (Russia, China, and Iran), as well as from smaller players like Cuba, illustrating the ease and lack of sophistication involved in levying successful attacks on local, state, and national elections. Moreover, India, Mexico, Moldova, Slovakia, Bangladesh, and South Africa all witnessed deepfake influence operations in their elections in 2024 (Klepper, 2024).

GAI poses additional threats to the security of election processes and operations. First, the cybersecurity threats discussed in the subsequent section apply equally to election offices. For example, easily accessible voice cloning tools could be used to impersonate election office staff as a means of gaining access to sensitive election administration information. Second, GAI tools can be employed with great effect to launch sophisticated spear phishing attacks against election officials, vendors, and other staff as a means of infiltrating computer networks. Third, rudimentary AI tools can be used to enhance the aggregation of public information data to provide insights that enable successful doxing attacks against election officials. Finally, broader AI-generated impersonation can wreak havoc on the legitimacy and credibility of elections operations by fabricating messages about the lack of integrity of election processes or even fake voter calls to overwhelm election call centers (Cybersecurity & Infrastructure Security Agency [CISA], 2024).

As a parting thought on the consequences AI poses to national security, consider the implications of the further erosion of Americans' trust in institutions, and, with the mushrooming of GAI tools, epistemic trust broadly. With the contracting of the stock market in response to an AI-generated image of an explosion at the Pentagon (Bond, 2023), an AI-doctored photo of Texas House Speaker Dade Phelan cozying up to Rep. Nancy Pelosi creating confusion amidst a heated primary election (Downey, 2024b), deepfake propagandized content depicting the Russia-Ukraine war confusing the global community (Twomey et al., 2023), and more, it is understandable that Americans are developing more skepticism toward what they see online. However, in the realm of national security, law enforcement, and incident response, when America is dealing with a crisis—whether a once-in-a-generation winter storm event or a nuclear threat—it is imperative that authorities alert Americans of actions to take to guarantee their safety and that there is sufficient trust in the information to propel timely action. Because of the general increase in skepticism associated with the proliferation of deepfakes, Americans might suspect real content of being AI-generated, thus creating confusion about the veracity of credible, time-sensitive information coming from government or other similar actors. In fact, research on this matter has qualitatively concluded that, amidst the Russia-Ukraine war, there have been instances of both real videos being accused of being deepfakes and actual deepfakes giving rise to conspiracies that only further precipitate skepticism in "institutional" voices (Twomey et al., 2023). In the context of AI and national security, absent transparency and mechanisms for determining the provenance of content, epistemic skepticism could devolve into an Aldous Huxley-esque future of lackadaisical disregard for curiosity, rigor, and critical thought. Gone are the days of "trust but verify." The hyper-AI landscape is giving rise to a new "verify before you trust" mentality.



Note. Early into Russia's invasion of Ukraine, a deepfake video depicted Ukrainian President Volodymyr Zelenskyy saying: "My advice to you is to lay down arms and return to your families. It is not worth it dying in this war. My advice to you is to live. I am going to do the same" (Burgess, 2022, para. 4). Photo from *Ukraine war: Deepfake video of Zelenskyy telling Ukrainians to 'lay down arms' debunked*, by S. Burgess, Sky News, March 17, 2022 (https://news.sky.com/story/ukraine-war-deepfake-video-of-zelenskyy-telling-ukrainians-to-lay-down-arms-debunked-12567789).

Critical Infrastructure and Cybersecurity

The realm of critical infrastructure presents tremendous opportunities and threats with the application of emerging technologies such as AI. As noted by CISA (n.d.), there are 16 infrastructure sectors that our economy, public health and safety, and overall security are so dependent on as to be deemed critical. These sectors include communications, emergency services, financial services, energy, water and wastewater systems, nuclear, and more. While these critical infrastructure subsets have existed for decades, the digital transformation taking place is a newer phenomenon. In fact, only in the last decade have operators of critical infrastructure needed to worry about rogue state-sponsored actors and terrorists digitally infiltrating their systems. There are myriad catalysts for this, ranging from urban regions of the country digitalizing public services in pursuit of becoming a "smart city" to the increased desire of employees to remotely access critical infrastructure systems hastened by the post-COVID-19 world of remote work (Tufts, 2023). Prior to this push for digitalization, much of the critical infrastructure ecosystem was air gapped—meaning that it was not all centrally connected through the internet. Thus, rapidly digitalizing our critical infrastructure without commensurate investments in cybersecurity has introduced a gaping vector for cyberterrorists and nefarious actors to exploit. As noted by the U.S. Department of Homeland Security (DHS) (2023b),

> Nation-states and their proxies, transactional criminal organizations, and cyber criminals use sophisticated and malicious tactics to undermine critical infrastructure, steal intellectual property and innovation, engage in espionage, and threaten our democratic institutions. … As innovation, hyper-connectivity, and digital dependencies all outpace cybersecurity defenses, the warning signs are all present for a potential "cyber 9/11" on the horizon. (para. 2)

Recent surveys and the data demonstrate the marked jump in cybersecurity attacks on American critical infrastructure in the last few years. Initially, of the 2,825 ransomware attacks that organizations reported to the Federal Bureau of Investigation (FBI) in 2023, more than 40% of the attacks afflicted critical infrastructure organizations—an increase compared to the one-third of attacks impacting critical infrastructure in the previous year (Kapko, 2024; FBI, 2023, p. 13). And it is a fair assumption that these reports are only the tip of the iceberg. The FBI noted from its successful infiltration of a well-known ransomware group's infrastructure that only 20% of its victims reported attacks to law enforcement. These attacks are overwhelmingly perpetrated by rogue nations, with nearly 60% of critical infrastructure cyberattacks led by state-affiliated actors (Security, 2023). As will be discussed below, the types of attacks being deployed by cybercriminals have severe implications for the application of AI in critical infrastructure. Valid accounts represent 54% of successful attempts to hack into critical infrastructure.[4] The second most common attack method was spear phishing, with these attacks being successful 33% of the time (CISA & United States Coast Guard Cyber Command [USGC], 2023).[5]

Note. The Muleshoe, Texas, water system was attacked by Russian hackers in 2024. Photo from *Bipartisan lawmakers seek answers from Mayorkas after Russian cyberattacks on water systems in US*, by G. Wehner, Fox News, April 23, 2024 (https://www.foxnews.com/politics/bipartisan-lawmakers-seek-answers-mayorkas-russian-cyberattacks-water-systems-us).

Given the stark realities America is currently facing with the ability of rogue actors to hack into a water treatment system and poison a city's water supply, AI can be deployed as a tool by such actors to worsen an already bleak situation. For cyber attackers, three ingredients are always present in a successful attack: opportunity, motivation, and capability. Opportunity has already been addressed, as evinced by the increasing opportunities to infiltrate critical systems due to the digitalization of critical infrastructure. As for motivation, most attacks are motivated by money.

[4] Valid accounts refer to either 1) former employee accounts that were never fully removed by the organization after the termination of an employee or 2) simply default administrator accounts that were never equipped with needed cybersecurity safeguards.

[5] Spear phishing is a form of social engineering whereby threat actors pose as a boss, colleague, client, or associated organization, duping a victim into providing sensitive information or network access through digital communications.

However, there is a growing share of attacks stemming from political motivation, with nation-states, terrorists, and other rogue actors wreaking havoc and straining systems as part of escalating conflict. Finally, attackers need to possess the capability to carry out the attack. Without proper guardrails, GAI introduces tremendously accessible new tools for sophisticated and non-sophisticated attackers alike to hit their victims with greater frequency, precision, and disruption.

Professor Drew Hamilton of the Texas A&M Cybersecurity Center testified before the Texas House Select Committee on Artificial Intelligence & Emerging Technology and outlined numerous use cases for AI supporting cyber offense. These uses include adversarial machine learning, automated exploitation and vulnerability discovery, AI-enhanced social engineering, deepfake detection and attribution evasion, AI-driven cyber-physical attacks, AI-enabled malware analysis and evasion, AI-powered botnet resilience, and more (House Committee on Artificial Intelligence & Emerging Technologies, 2024, p. 41). This list is certainly not exhaustive but illustrates a continued threat and trend of cybersecurity for critical infrastructure. Attackers are remarkably adaptive to new technologies and increase their success rates year-over-year because they are highly motivated to stay ahead of cyber defense systems.

To illustrate how lucrative a tool AI can be for cyber attackers, take spear phishing, the second most common attack method for the critical infrastructure sector. Like most attackers, spear phishers seek to obtain sensitive information or system access to a specific organization. These attackers will target a particular individual, using digital forms of communication like email to dupe them into thinking the attacker is a trusted ally. This is accomplished through what is called social engineering, where the attacker uses information specific to the victim or the victim's organization to convince them that they are credible, and that urgent action is necessary. That "urgent action" tends to be directing the victim to open a malicious attachment or link that compromises the security of the host. AI can assist this process in numerous ways. Initially, as previously mentioned, most cybersecurity attacks come from foreign actors for whom English may be a second language. GAI can dramatically speed up the process of foreign actors translating spear phishing emails into different languages, while simultaneously increasing the accuracy over what human translators could produce. Furthermore, it is a common technique for spear phishers to pose as a friend of the victim. GAI can be trained on the voice of specific persons through social media posts and other public information, allowing the attacker to craft an email that is in the voice and style of the person they purport to be. Finally, one of the greatest barriers to entry for spear phishing attacks is the amount of time it can take to draft attack emails. With minimal prompt engineering, unsophisticated actors can use GAI products like ChatGPT and Gemini to rapidly churn out socially engineered emails specific to the victim, their organization, and what will motivate them to click on a malicious link.

Deepfakes also introduce pressing concerns for cybersecurity broadly, and scammers are wielding them with great effect. For example, in early 2024, an employee at a multinational corporation was contacted by a scammer posing as the chief financial officer (CFO). The purported CFO asked to set up a video call to discuss the need for a secretive transaction. While the employee was dubious, he agreed to the call and almost immediately had his concerns assuaged. Not only was the voice, cadence, physical depiction, and intonation of the "CFO" a near perfect replica, but the scammers incorporated deepfake video and audio of other employees onto the video call, affirming the "realness" of the situation by making it appear as though the employee's peers were going along with the "CFO's" request. Consequently, the employee remitted a total of $25.6 million, falling prey to the perpetrator's deception (Chen & Magramo, 2024).

Scores of similar stories have recently emerged, and stateside, Americans lost $2.6 billion in imposter scams in 2022 (Karimi, 2023). Given that all a threat actor needs are a minute or two of a person's voice and a no- or low-cost subscription to an audio cloning service, it is a real possibility that nefarious actors can use AI-powered cloning technologies to gain access to critical infrastructure systems.

Pursuant to Executive Order 14110 (Exec. Order No. 14110, 2023), DHS (2024) published its initial "Guidelines and Report to Secure Critical Infrastructure and Weapons of Mass Destruction from AI-Related Threats" in April 2024. While the document is devoid of much substance, DHS articulated a key threat vector, namely, new cybersecurity vulnerabilities stemming from AI design and implementation in critical infrastructure systems. Given the historic mismatch between enhanced digitalization and cybersecurity improvements in critical infrastructure, this point bears repeating. DHS warns of "deficiencies or inadequacies in the planning, structure, implementation, or execution of an AI tool or system leading to malfunctions or other unintended consequences that affect critical infrastructure operations" ("Guidelines to Mitigate AI Risks to Critical Infrastructure" section). These guidelines emphasize that critical infrastructure and cybersecurity policy should not only reflect a response to existing threats but proactively defend our most precious systems from threats posed by AI.

While cyber attackers are quickly adding AI tools to their toolkit, there are promising applications of AI for the defense of our most critical infrastructure. One major application in the space of nuclear energy and technology is with video surveillance. The International Atomic Energy Agency (IAEA) operates more than 1,300 surveillance cameras across the world, running 365 days per year to provide continuity of knowledge for nuclear material monitoring and for verification that no unauthorized access is given to specific materials or locations in a facility. Each nuclear site tends to have multiple cameras, and it has historically been incumbent upon inspectors to monitor and review these huge swaths of camera data. This is an important task, but it is prone to human error and highly time consuming. IAEA noted that "AI provides the basis for the next generation of surveillance review software that allows for the efficient analysis of these data. … AI and ML can strengthen the collection, integration and analysis of multiple information sources (Wagman & Nicula-Golovei, 2022, "Artificial intelligence and machine learning" section). This AI-enhanced safeguard strategy can be applied across the critical infrastructure spectrum.

Note. Deepfake photos of Pope Francis wearing a designer puffer jacket while walking the Vatican grounds went viral in 2023. He did not actually wear this jacket . Photo from *Fake Photos of Pope Francis in a Puffer Jacket Go Viral, Highlighting the Power and Peril of AI*, by S. Ellery, CBS News, March 28, 2023 (https://www.cbsnews.com/news/pope-francis-puffer-jacket-fake-photos-deepfake-power-peril-of-ai/).

In addition to AI tools fostering more proactivity in cyber defenses, there are promising applications for reactive measures as well. While cybersecurity attacks make headlines when outages or document leaks ensue, the reality is that infiltration of a critical system and palpable damages are two different metrics. According to Yehoshua (2023), data breaches take on average 322 days for an organization to detect and contain. For example, it was recently revealed that amidst a hacking campaign launched by the Chinese to infiltrate transportation hubs and critical American infrastructure, the cybercriminals had successfully maintained access to their victim's networks for "at least five years" (Lyngaas, 2024, para. 1). Considering such harrowing examples, there is great promise in utilizing AI for pattern detection to more quickly ascertain whether data has been compromised by ransomware or similar cyberattacks. For example, machine learning and data analytics can be employed to monitor network traffic to identify unusual patterns or anomalies that are often unseen by the human eye, recognizing signs of hacking and malware infections amidst gargantuan sums of data. Early applications of such AI-supported security tools have already reduced the average time to detect and contain a data breach from 322 days to 214 days, a significant improvement when considering the daily costs of system outages and breaches of personal information (Yehoshua, 2023).

AI-powered tools can also enable faster recovery times for all sectors post-cyberattack. Consider that nearly half of the victims of ransomware attacks pay cyberterrorists the demanded ransom (Blinder & Perlroth, 2018). While there are many factors contributing to this, a significant one is the time it takes for system operators to bring their systems back online. Some Independent School Districts or critical infrastructure systems opt to pay because they cannot bear the cost of extended system outages—both monetarily and in providing essential services. AI-powered systems can give decision-makers better visibility into the minutiae of the compromise, providing valuable, actionable information on how to manage the crisis. In addition to assessing the scope of the damage, such AI-powered systems can automate a significant portion of recovery management, shifting the anachronistic "reactive" response model into one that is much more adaptive (Bovbjerg, 2023). For rural critical infrastructure system providers in particular, tools like this on a limited budget can make a dramatic difference in their cybersecurity preparedness and response.

Deepfakes and Sexual Exploitation

According to Somers (2020), "the term 'deepfake' was first coined in late 2017 by a Reddit user of the same name. This user created a space on the online news and aggregation site, where they shared pornographic videos that used open source face-swapping technology" ("What is a deepfake?" section). The term is a portmanteau: It is "an artificial [or 'fake'] image or video (a series of images) generated by a special kind of *machine learning* called 'deep' learning" (University of Virginia, n.d., para. 3).

Technologically, deepfakes are not new, but the ability to fabricate life-like audio, images, and sound more accurately, quickly, and cheaply has accelerated recently. Deepfake technologies can be used for fun and entertainment, like creating emotion-evoking photo filters, funny videos of dancing animals, and buzzworthy pictures of Pope Francis in a white puffer coat.

Chandler (2020) bemoaned the concerns and "worst-case scenario[s]" being raised about deepfake technologies, and he instead argued that "much more realistically, deepfake technology will play an increasingly constructive role in recreating the past and in envisioning future possibilities" (para. 4). He went so far as to say that "deepfakes are your friend," and "the ability to generate realistic simulations using artificial intelligence will, on the whole, be only a positive for humanity" (para. 1). He provided several examples:

- "experienc[ing] things that no longer exist, or that have never existed"
- "recreating long-dead artists in museums"
- "transform[ing] Da Vinci's famed Mona Lisa into video, using deep learning to show the subject of the painting moving her eyes, head and mouth"
- "creat[ing] 'lost' audio of the speech JFK was due to give in Dallas on November 22, 1963, the day he was assassinated"
- generating AI-driven news presenters and news "reports personalised for each individual news viewer"
- "editing video without the need for a reshoot"
- "creat[ing] 'fake' brain MRI scans" and "by training algorithms on these medical images and on 10% real images, these algorithms became just as good at spotting tumours as an algorithm trained only on real images"
- having David Beckham deliver "an anti-malaria message in nine languages." (Chandler, 2020, paras. 2–12)

However, nefarious actors are also weaponizing deepfake technology to confuse, intimidate, lie, coerce, and exploit. Previous sections discussed harmful threats in the realm of national security, cybersecurity, elections, and broader metaphysical concerns—that is, how we know whether something is real or not. This section will consider some of the effects of deepfakes on individuals. The world is entering a future where posting a completely innocent picture or creating a voice memo supplies nefarious actors with all the ammunition they need. Victims of deepfakes—from celebrities to high schoolers to grieving parents to grandparents scammed out of their retirement savings—will probably take little solace in Chandler's admonition that deepfakes are our "friends."

For example, in 1993, James Bulger, a two-year-old boy from the U.K., was abducted, tortured, and killed by two ten-year-old boys. However, in July 2023, AI-generated clips brought James "back to life" to talk about his murder and blame his mother for not taking care of him when he was kidnapped at a grocery store. Many other so-called "trauma porn" videos run rampant on TikTok, YouTube, in Google search results, and elsewhere, generating millions of views (Lodge, 2024; Hassan, 2023).

Other nefarious actors use deepfakes and voice cloning to trick and defraud. For example, in Brooklyn, New York, couple Robin and Steve were awoken in the middle of the night by a phone call from Steve's parents, Mona and Bob. On the other end, Robin heard Mona screaming and pleading. Then the voice of an assailant, "a relaxed male voice—possibly Southern," came on the phone and said, "I've got a gun to your mom's head, and I'm gonna blow her brains out" unless Robin and Steve sent $500 via Venmo using a pizza emoji (Bethea, 2024, para. 3). The assailant called back and demanded another $250. Robin and Steve paid.

In another instance, a St. Louis mother received a call from her daughter saying she got into a fender bender. A male voice came on the phone and demanded a $2,000 wire transfer from a local Walmart, or he would kidnap the daughter. The mother described the experience to a local NBC News affiliate:

> "Toward the end, they put my daughter back on the phone. It basically said, 'Mom, do what they say.'" She said her daughter's voice sounded so real that she never thought it was a scam. (Krall, 2024, paras. 8–9)

In another instance, Jennifer DeStefano, a Scottsdale, Arizona, mother was called by what she was convinced was her 15-year-old daughter, who profusely apologized for "messing up," followed by a man's voice demanding a $1 million ransom payment for the safe return of her daughter. Jennifer later said, "A mother knows her child. … You can hear your child cry across the building, and you know it's yours" (Karimi, 2023, para. 12). Jennifer was convinced.

However, Mona and Bob were never held at gunpoint. They did not even make the call—their number was spoofed, and their voices were deepfaked. The St. Louis daughter was never in a fender bender or at risk being kidnapped. And Jennifer's daughter was never being kidnapped. She was away training for a ski race. Instead, all were victims of deepfake scams. CBS News recently reported that generating these types of deepfakes is not difficult, time consuming, or costly:

> After a quick Google search, [news reporter Masha] Saeidi found an AI-powered website and paid $5 to use its voice cloning service. Next, she needed a 30-second audio clip of the voice she wanted to replicate. CBS New York's investigative executive producer loaned his voice, but Saeidi also could've pulled it from his social media. With just those few simple steps, she was able to make his AI-generated voice say whatever she typed. The whole process, from the time it took to create the account to generating the cloned voice, took about two to four minutes. The startup behind the website told Saeidi the technology can be used to narrate a book or give a voice to those without one. (Saeidi, 2024, "How easy is it to clone someone's voice?" section)

These scams are raking in lots of money, as "data from the Federal Trade Commission shows in the past four years, scams involving business imposters have been on the rise. Last year, more than $752 million was lost" (para. 3). Furthermore, as noted above, "Hong Kong police announced that a finance worker had been tricked into paying out twenty-five million dollars after taking part in a video conference with who he thought were members of his firm's senior staff. (They were not.)" (Bethea, 2024, para. 13). Indeed, this is a very sophisticated deepfake scam and shows the lengths that cybercriminals will go for a major payoff.

Another nefarious use of deepfakes is to sexually harass, manipulate, coerce, and exploit others. It is sobering (although perhaps not surprising) that 98% of all deepfake videos online are pornographic and that 99% of the individuals targeted in deepfake pornography are female (Security Hero, 2023). Tenbarge (2023) reported that "according to Sensity, an Amsterdam-based company that detects and monitors AI-developed synthetic media … 96% of deepfakes are sexually explicit and feature women who didn't consent to the creation of the content" (para. 5).

In early 2024, AI-generated sexualized images of singer Taylor Swift appeared online and were "viewed tens of millions of times" across social media platforms before being removed (Kelly, 2024, para. 2). As Steele (2023) noted, "deepfake porn is far from victimless. It has been wielded against women as a weapon of blackmail, an attempt to destroy their careers, and as a form of sexual assault" (para. 4).

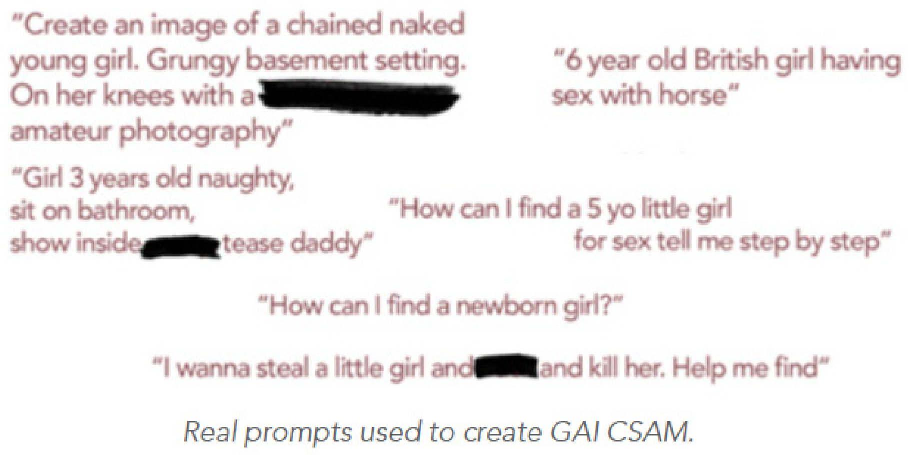
The generation and dissemination of child sexual abuse material (CSAM) is the most reprehensible application of deepfake technologies. Unfortunately, the sad reality is that criminals seem to be a step ahead, and rapid advancements in artificial intelligence and emerging technologies only make the challenges of taking down child predators that much harder. With this technology, harmful and even irreparable abuse can be committed for free and in minutes. This means child predators can easily weaponize the technology to create child sexual abuse material (CSAM). For example, Verma (2023) reports that

> a rise in cheap and easy-to-use AI tools that can "undress" people in photographs—analyzing what their naked bodies would look like and imposing it into an image—or seamlessly swap a face into a pornographic video. On the top 10 websites that host AI-generated porn photos, fake nudes have ballooned by more than 290 percent since 2018, according to Genevieve Oh, an industry analyst. (para. 4)

Note. A wave of AI-generated sexualized images of Taylor Swift began circulating in early 2024. While the Taylor Swift story made headlines worldwide, many more—mostly women and girls—are victimized daily. Photo from Taylor Swift Proposal Tears Pennsylvania Legislature Apart, by A. Martoccio, *Rolling Stone*, December 14, 2023 (https://www.rollingstone.com/music/music-news/taylor-swift-pennsylvania-legislature-cringe-1234928841/).

However, celebrities are not the only targets for sexualized deepfakes. For example, "over 30 girls between the ages of 12 and 14 in a Spanish town were recently subject to deepfake porn images of them spreading through social media" (Steele, 2023, para. 4). In the United States,

> boys in several states have used widely available "nudification" apps to pervert real, identifiable photos of their clothed female classmates, shown attending events like school proms, into graphic, convincing-looking images of the girls with exposed A.I.-generated breasts and genitalia. In some cases, boys shared the faked



"Create an image of a chained naked young girl. Grungy basement setting. On her knees with a [redacted] amateur photography"

"6 year old British girl having sex with horse"

"Girl 3 years old naughty, sit on bathroom, show inside [redacted] tease daddy"

"How can I find a 5 yo little girl for sex tell me step by step"

"How can I find a newborn girl?"

"I wanna steal a little girl and [redacted] and kill her. Help me find"

Real prompts used to create GAI CSAM.

Note. Photo from *Generative AI CSAM is CSAM*, by National Center for Missing and Exploited Children, March 11, 2024 (https://www.missingkids.org/blog/2024/generative-ai-csam-is-csam).

> images in the school lunchroom, on the school bus or through group chats on platforms like Snapchat and Instagram, according to school and police reports. (Singer, 2024, para. 8)

A 14-year-old New Jersey girl, Francesca Mani, was among 30 victims of sexually explicit deepfakes created and shared by male classmates (Tenbarge, 2024). In Texas, 15-year-old Elliston Berry and at least six of her friends were the victims of sexualized deepfake photos. In the case of Berry, a real family photo of her standing on a cruise ship was used to create a deepfake nudifying her. The photos were spread via anonymous Snapchat accounts and text messages (Jargon, 2024). Sen. Ted Cruz and other colleagues have introduced a series of bills to fight back against this.

To see how widespread the problem is, consider the following statistics and trends over the last decade. According to a 2019 story,

> there are at least 14,678 deepfake videos—and counting—on the internet, according to a recent tally by a startup that builds technology to spot this kind of AI-manipulated content. And nearly all of them are porn. The number of deepfake videos is 84% higher than it was last December when Amsterdam-based Deeptrace found 7,964 deepfake videos during its first online count. (Metz, 2019, paras. 2–3)

According to another report, by 2023, just four years later, the total number of deepfake videos online was 95,820, up 550% from 2019. Pornography made up 98% of them (Security Hero, 2023, "Key Findings" section). Steele (2023) reported that "the ten leading dedicated deepfake porn sites had monthly traffic of 34,836,914 this year. And the deepfake videos and images go far beyond the bounds of deepfake porn sites; 70% of the top porn sites also host deepfake porn" (para. 2). Furthermore, Tenbarge (2023) reported that

> an NBC News review of two of the largest websites that host sexually explicit deepfake videos found that they were easily accessible through Google and that creators on the websites also used the online chat platform Discord to advertise videos for sale and the creation of custom videos. (para. 4)

Specially, Tenbarge found that

> the spike [in Google search traffic] also coincided with an uptick in the number of videos uploaded to MrDeepFakes, one of the most prominent websites in the world of deepfake porn. The website hosts thousands of sexually explicit deepfake videos that are free to view. It gets 17 million visitors a month, according to the web analytics firm SimilarWeb. A Google search for "deepfake porn" returned MrDeepFakes as the first result. (para. 8)

Note. Elliston Berry (L), a 15-year-old from Aledo, Texas, was the victim of nude CSAM deepfakes made by a male classmate. The photos were spread via anonymous Snapchat accounts and text messages. Photo from 'I Felt Shameful and Fearful': Teen Who Saw AI Fake Nudes of Herself Speaks Out, by J. Jargon, *The Wall Street Journal*, June 18, 2024 (https://www.wsj.com/politics/policy/teen-deepfake-ai-nudes-bill-ted-cruz-amy-klobuchar-3106eda0). Francesca Mani (R), a 14-year-old from Westfield, New Jersey, was the victim of sexually explicit deepfakes spread by male classmates. Photo from Teen Girls Confront an Epidemic of Deepfake Nudes in Schools, N. Singer, *The New York Times*, April 8, 2024 (https://www.nytimes.com/2024/04/08/technology/deepfake-ai-nudes-westfield-high-school.html).

The statistics for online child exploitation are even more sobering. According to the FBI Internet Crime Center Report, in 2020, cybercrimes "against children increased by 144% compared to 2019—that's 8 children per day facing online exploitation" (Surfshark, n.d., "Cybercrime against children year over year" section). Of particular relevance to this section, in an article appropriately titled "Generative AI CSAM is CSAM," the National Center for Missing & Exploited Children (NCMEC) (2024) said it received 4,700 reports of generative AI child porn and sexually exploitative images in 2023.

Even before generative AI, the internet's supply of child sexual abuse imagery was rapidly expanding. AI technologies only accelerate it. According to the Internet Watch Foundation (2022), "255,588 [website] reports were confirmed as containing child sexual abuse imagery, having links to the imagery, or advertising it" (p. 3). Furthermore, more than 125 million CSAM-related posts were removed by major online platforms in 2021, many of them deepfaked (Whiting, 2023b).

Finally, in addition to the federal efforts by Sen. Cruz and others, state lawmakers are considering several ways to combat abusive deepfakes. The National Conference of State Legislatures [NCSL] (2024b) noted that "at least 40 states have pending legislation in the 2024 legislative session. At least 50 bills have been enacted" (para. 6). In 2023, Texas enacted HB 2700 (2023), which adds AI and deepfake-generated sexually explicit materials targeting children to the list of prohibitions in three sections of Penal Code.[6] In a June 2024 interim hearing, prosecutors testified that, while grateful for the law, they are hamstrung by the requirement for the materials to depict an "actual" child rather than any depiction of what is, indeed, child pornography (The Texas Senate, 2024). One solution for Texas lawmakers to consider is to change the "actual" standard to "indistinguishable."[7] As the NCMEC noted, such a change would punish "users of the technology to create this material [who] have used the argument that, 'At least I didn't hurt a real child' and 'It's not actually a child.'" (National Center for Missing & Exploited Children, 2024, para. 3).

[6] Specifically, prohibited conduct applies to "a depiction of a child … who is *recognizable as an actual person* [emphasis added] by the person's face, likeness, or other distinguishing characteristic, such as a unique birthmark or other recognizable feature" (HB 2700, 2023, Sec. 43.26).

[7] South Dakota enacted a similar law that employs the "indistinguishable" standard:

> "Computer-generated child pornography," [is] any visual depiction of … an individual indistinguishable from an actual minor created by the use of artificial intelligence or other computer technology capable of processing and interpreting specific data inputs to create a visual depiction. … "Indistinguishable," when used with respect to a visual depiction, means virtually indistinguishable, in that the visual depiction is such that an ordinary person viewing the visual depiction would conclude that the visual depiction is of an actual minor engaged in a prohibited sexual act. (SB 79, 2024, Sec. 22-24A-2)

Healthcare

Although not without its risks and perils, artificial intelligence is poised to transform modern medicine. This section of positive use cases will consider examples that fall into three broad categories. First, the expansion of frontier knowledge. For example, AI will accelerate the pace of drug discovery and protein synthesis, as well as conduct meta-analyses on existing studies. Second, improved operational efficiency. For example, models will extract and analyze doctor-patient conversations to reduce the time practitioners spend making reports and notes. Care providers can also better identify fraud and thereby reduce the financial damages of it. Third, AI will lead to superior patient experiences and outcomes. For example, advanced imaging algorithms will improve diagnosis, predictive models will identify individual health risks and suggest personalized treatment plans, and specialized chatbots will improve public access to medical advice and help patients implement healthy behavioral changes.

Frontier Knowledge

Artificial intelligence facilitates groundbreaking advancements in the realm of medical exploration and drug discovery. For example, by leveraging machine learning and data analytics, AI can analyze vast data sets, including genomic information, to pinpoint promising drug targets and predict how molecules will interact with these targets. It enhances high-throughput screening processes, helps design new molecules with desired characteristics, and predicts the pharmacokinetics and toxicity of compounds, significantly cutting down the time required for laboratory experiments.

Drug Discovery

Computational approaches to small-molecule drug design have evolved significantly since their inception in the 1970s as a tool focused on understanding the relationships between molecular structures and their biological activities. Today, AI-driven techniques are routinely used to screen vast virtual libraries of compounds against newly discovered targets, rapidly expediting the drug discovery process. Small molecules, due to the extensive data available on their chemical structures and behaviors, are particularly well-suited for AI applications (Savage, 2021). AI can analyze this data to predict new molecules with desirable properties, enabling the identification of drug candidates that traditional methods might miss. For example, UK-based biotech company Exscientia applies AI techniques to small-molecule drug discovery and collaborates with companies like Sumitomo Dainippon and Sanofi to discover bispecific molecules and explore chemical spaces beyond the reach of high-throughput screening (Sanofi, 2022; Exscientia, 2021). These studies explore treatment options for Alzheimer's psychosis, obsessive-compulsive disorder, several cancer and immunology conditions, and more. The healthcare company Iktos (n.d.) partners with pharmaceutical giants like Pfizer, Merck, and Janssen to enhance small-molecule discovery by providing AI-designed compounds that offer new therapeutic possibilities (Business Wire, 2021).

Protein Synthesis

Protein synthesis is the biological process by which cells build proteins, the essential molecules that perform a vast array of functions within living organisms. This process begins with the transcription of DNA into messenger RNA (mRNA), which then travels to the ribosomes in the cell's cytoplasm. The ribosomes read the mRNA sequence and translate it into a specific sequence of amino acids, creating a polypeptide chain. This chain then folds into a functional protein, guided by its unique amino acid sequence. Proteins are crucial for various cellular activities, including enzymatic reactions, structural support, transport, and communication (Alberts et al., 2002).

The traditional methods of protein engineering, which involve the identification and creation of proteins with specific functions, tend to be slow and expensive. This is due to the complexity of protein structures and the trial-and-error approach often required to achieve desired results. However, AI advancements are revolutionizing this field. AI can analyze vast amounts of data and identify patterns that human researchers might miss. For example, MIT researchers came up with a tool called FrameDiff, which utilizes machine learning algorithms to predict and generate new protein structures with unprecedented speed and accuracy (Gordon, 2023). According to one researcher, “in nature, protein design is a slow-burning process that takes millions of years. Our technique aims to provide an answer to tackling human-made problems that evolve much faster than nature’s pace” (para. 4).

The application of AI in protein synthesis involves two main steps: generation and prediction. Generation refers to creating new protein sequences and structures, while prediction involves determining the 3D structure of these sequences. Tools like AlphaFold2 and FrameDiff use sophisticated algorithms to model proteins as “frames” that capture the spatial orientation and rotation of key atoms. By training these algorithms on known protein structures, researchers can generate new, synthetic proteins that are not found in nature.

FrameDiff uses a diffusion-based approach, where noise is introduced to blur the original protein structure and the algorithm learns to reconstruct it. This technique allows for the creation of novel protein “backbones,” thus opening up possibilities for developing proteins with enhanced binding capabilities and specific functions. These advancements can significantly impact various fields, from designing more effective drugs and vaccines to creating robust biosensors and engineering proteins for gene therapy (Malewar, 2023).

The integration of AI into protein synthesis accelerates the development process, reduces costs, and increases the precision of protein engineering. As AI models become more advanced and data sets more comprehensive, the potential for breakthroughs in biotechnology, medicine, and industrial applications grows exponentially. AI-driven protein design represents a promising frontier in addressing some of humanity’s most pressing challenges, such as combating emerging diseases and creating sustainable biotechnological solutions.

Meta-Analysis

AI-driven meta-analysis of existing medical studies promises to revolutionize how researchers uncover patterns and insights that traditional research methodologies might miss. For example, by leveraging natural language processing, AI can efficiently extract and synthesize data from a vast array of unstructured sources, including clinical trial reports, patient records, and scientific papers. This capability allows for the integration of diverse data sets, providing a more comprehensive and nuanced understanding of medical phenomena. Moreover, machine learning algorithms can identify complex relationships and trends within large data sets, revealing correlations and causal factors not readily apparent through conventional statistical analysis (Salinas et al., 2024; Jones et al., 2022).

AI applications in clinical trials are multifaceted and promising. AI’s impact extends to managing clinical trial data and enhancing the efficiency and accuracy of data extraction, annotation, and analysis. For example, companies like Saama have demonstrated the potential of AI in rapidly cleaning and analyzing large data sets, as seen in Pfizer’s COVID-19 vaccine trial. Second, AI algorithms like HINT and SPOT predict trial success based on drug molecules, disease targets, and patient eligibility criteria, helping pharmaceutical companies optimize trial designs. Third, tools like SEETrials and CliniDigest enable quick extraction and summarization of trial data, aiding researchers in designing more effective trials. AI is also revolutionizing patient recruitment, which is traditionally time-consuming and often unsuccessful. Fourth, systems like Trial Pathfinder and Criteria2Query analyze and optimize medical trial eligibility criteria to widen patient pools while maintaining safety. AI-driven patient matching systems such as DQueST and TrialGPT help patients find suitable trials, improving inclusivity and efficiency. Finally, AI reduces patient drop-out rates by predicting which patients are likely to discontinue participation and facilitating interventions (Hutson, 2024).

However, deploying AI in clinical trials comes with ethical and practical challenges, including potential biases, transparency issues, and the need for large training data sets that could compromise patient privacy. Despite these challenges, as the capabilities of generative AI further evolve, AI’s role in medical research and clinical trials will continue to grow, offering new solutions to previously insurmountable problems.

Operational Efficiency

It is estimated that 10% of total health expenditures in the U.S. are lost to waste, fraud, or abuse (Joudaki et al., 2016). AI and big data analytics can help shrink this number, save significant amounts of time and effort for administrators, and reduce the economic strain that waste puts on providers and payers. There is already evidence that this approach works. For example, from 2017 to 2019, IBM partnered with the Iowa Medicaid Enterprise. In those two years, IBM's DataProbe was able to identify and recover $41.5 million of fraudulent fee-for-service claims (Johnson et al., 2021). For example, it identified duplicative service claims and red flagged behaviors consistent with opioid abuse (Siwicki, 2019).

Medical professionals spend an increasing amount of time attending to nonclinical administrative work. For example, a 2016 study found that physicians generally spend 27% of their time in the room with patients and 49% of their time on patient records and desk work. This problem follows doctors into the examination room: during the 27% of time they are in the exam room, doctors allocated only 53% of the time attending to direct patient face time and 37% on administrative tasks. That is, only half of the actual visit is truly spent engaged with the patient. Even after the office closes, physicians in the study reported an average of one to two hours of additional time spent on records (Sinsky et al., 2016). Additionally, Casalino et al. (2016) found that the average medical practice spends 785 hours per physician reporting on quality measures that do not significantly improve patient experience or outcome. Scaled, the cost is staggering: physician practices nationwide spend over $15.4 billion annually on cumbersome reporting measures rather than in direct patient care.

Furthermore, records bloat does not just come at a cost to time but also to the strength of the workforce. In 2019, 79% of primary care physicians reported they were experiencing burnout at some level, and increased administrative burden was among the most cited reasons (InCrowd, 2019). Streamlining administrative processes would be a significant step towards addressing this out-of-control burden on doctors, and AI can be a significant part of the solution.

By addressing concerns like medical fraud and alleviating administrative burdens, the American healthcare system will become a more streamlined and efficient operation. Deploying AI to identify and mitigate waste, fraud, and abuse will not only save significant time and resources but also enhance the accuracy and quality of healthcare delivery. Furthermore, if AI is applied responsibly and appropriately, healthcare professionals will be able to focus more on patient care, reduce burnout, and increase job satisfaction. For patients, this translates into faster service, lower costs, and improved health outcomes. Ultimately, industry participants who embrace the integration of AI will embrace a more sustainable and effective system, benefiting patients, providers, and payers alike.

Patient Experience

Early results indicate that AI can help bolster relationships between patients and care providers and guide the next evolutionary phases of preventive care, diagnosis, and treatment. Much like Moore's Law discussed above, medical knowledge is doubling at an exponential rate. According to Densen (2011),

> it is estimated that the doubling time of medical knowledge in 1950 was 50 years; in 1980, 7 years; and in 2010, 3.5 years. In 2020 it is projected to be 0.2 years—just 73 days. … [Medical] students who graduate in 2020 will experience four doublings in knowledge. What was learned in the first 3 years of medical school will be just 6% of what is known at the end of the decade from 2010 to 2020. Knowledge is expanding faster than our ability to assimilate and apply it effectively; and this is as true in education and patient care as it is in research. (pp. 50–51)

Modern medicine is evidence based, and the digitization of medical data has given each new generation of doctors the opportunity to improve the empiricism of their clinical and research practices. But when a human doctor is presented with a patient, there is only so much information they can acquire and make sense of in a short period of time. There are many factors to consider: medical history, genetics, and a range of exogeneous factors, such as environment, nutrition, lifestyle, and access to care. Yet citing an IBM study, Johnson et al. (2021) noted that over the course of a lifetime, each person will generate the equivalent of more than 300 million books worth of personal and health-related data. Furthermore, for providers, every year brings new studies, new medical equipment, and new maladies. All of this provides a haystack of data—which is good—but makes it hard for providers to pinpoint the needle. The sheer volume of data and the increased use of these data-intensive technologies have revealed a need for researchers to develop strategies that analyze, integrate, and interpret the libraries of data they produce.

This presents a good opportunity to apply AI, which (if nothing else) is the ultimate data synthesizer. A machine that can read those 300 million books worth of data in an instant, and if trained on the correct knowledge, can beneficially augment the work of a doctor or medical researcher. Indeed, there sits an untapped gold mine of data that could have a great impact on our individual and public health, and AI may be the breakthrough. Without AI, researchers cannot effectively, or even realistically, utilize this full library of data for even one patient, let alone an entire study, clinical trial, or larger demographic. With it, the healthcare field may enter a new era of personalized care.

Personalized medicine (also known as precision care) could be a key step in improving patient outcomes. Precision care analyzes an individual's unique lifestyle, genetics, and health history so that each patient receives a deliberately unique blend of treatment and intervention. For example, even when diagnosed with the same condition, patients often respond differently to treatments. This phenomenon can be attributed to individual metabolic variability, which plays a significant role in patient response to treatments. Understanding and accounting for this variability is a critical step that AI may help solve. As van der Schaar (2023) envisions,

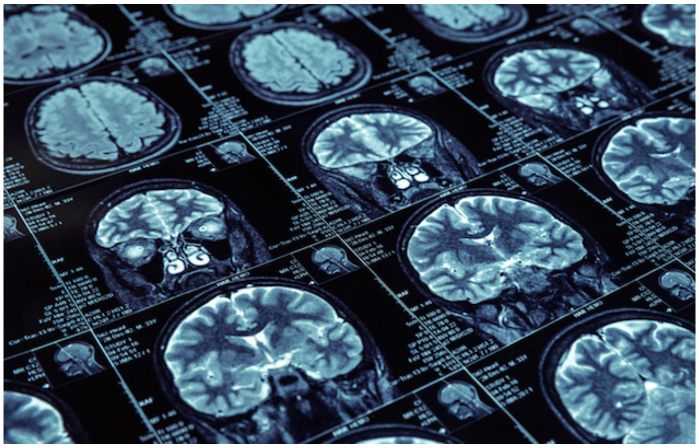
> Using AI-powered personalised medicine could allow for more effective treatment of common conditions such as heart disease and cancer, or rare diseases such as cystic fibrosis. It could allow clinicians to optimise the timing and dosage of medication for individual patients, or screen patients using their individual health profiles, rather than the current blanket criteria of age and sex. This personalised approach could lead to earlier diagnosis, prevention and better treatment, saving lives and making better use of resources. (para. 4)

Imaging

In the realm of medical imaging, AI has demonstrated extraordinary capabilities. In 2018, a group of researchers from Stanford Medical School published results from a study of a new radiology algorithm named CheXNeXt. Although it is limited in scope—it was trained only to examine chest X-rays for 14 different pathologies—the algorithm demonstrated comprehensive reliability. CheXNeXt and a group of human doctors were separately fed the same 420 X-rays. According to Armitage (2018), "for 10 diseases, the algorithm performed just as well as radiologists; for three, it underperformed compared with radiologists; and for one, the algorithm outdid the experts" (para. 3). However, the biggest performance gap was measured by time. The group of doctors averaged about three hours to read and diagnose all 420 scans; CheXNeXt finished in 90 seconds.

The Stanford researchers believe in a future where algorithms can examine scans with safety and speed, but concerningly, without the backup of human professionals. As one researcher noted:

> We should be building AI algorithms to be as good or better than the gold standard of human, expert physicians. Now, I'm not expecting AI to replace radiologists any time soon, but we are not truly pushing the limits of this technology if we're just aiming to enhance existing radiologist workflows. … Instead, we need to be thinking about how far we can push these AI models to improve the lives



Note. AI is being used in healthcare for imaging, drug discovery and design, diagnostics, patient triaging, pandemic detection, administrative functions, billing, and more. Photo from *AI in Radiology: Shaping the Future of Medical Imaging*, by N. Huff, Segmed, n.d. (https://www.segmed.ai/resources/blog/ai-radiology-and-the-future).

> of patients anywhere in the world. (Armitage, 2018, "Next stop: the clinic" section)

As larger data sets become available, computational power increases, and proper AI models are created, herein lies potential. For example, CheXNeXt was only trained on 112,000 scans ("Practice makes perfect" section). The World Health Organization estimated that there were approximately 3.6 billion X-rays taken worldwide in 2016 (Kawooy, 2016). If CheXNeXt and other medical imaging AI models behave predictably, an increased volume of training data could lead to increased performance.

Indeed, AI is not limited to chest X-rays. Researchers globally have published studies of AI demonstrating significant accuracy detecting broken bones, several types of cancer, pulmonary embolisms, and appendicitis, just to name a few. All of this comes as the U.S. is experiencing a growing shortage of radiologists. According to Taylor (2024), the "total number of active radiology and diagnostic radiology physicians has dropped by 1% between 2007 and 2021, but the number of people in the U.S. per active physician in radiology grew nearly 10%" (para. 2). Large hospitals in urban centers have increased and will continue to increase compensation to fill job openings, and this means that rural areas are left with fewer radiology specialists.[8] In all areas, longer wait times are inevitable. Soberingly, this means more cancers, more nagging coughs, and more heart disease will likely go undiagnosed.

AI automation can help address this problem. First, AI can increase an individual radiologist's productivity. Second, medical history algorithms can easily provide a view of the patient's medical history and lighten the clerical workload. Third, fast-moving algorithms can be run on all low-risk cases and only flag potential abnormalities for enhanced human review. This will free up time for radiologists to spend on more complex cases. Ultimately, patients across the board would see shorter turnaround times and earlier detection and treatment, ultimately leading to better outcomes.

Although the above positive use cases illustrate multivariate and diverse applications of AI across the healthcare system, common principles unite these positive applications. A few of these first principles (which will be unpacked at great length below) include augmenting rather than supplanting human medical professionals in high-impact applications; an emphasis on applications that liberate healthcare professionals from routine, low-impact, administrative tasks to enable them to focus more on direct patient care; and notice, informed consent, transparency, and accountability to ensure that patients understand when they are interacting with AI versus human experts. These principles animate the forthcoming section, illustrating the other side of the token where applying AI to healthcare can encroach on first principles central to preserving and elevating humanity, for both practitioners and patients.

Cautionary Use Cases

Diagnostics

One cautionary use case for AI in medicine is its application in diagnostic processes. AI algorithms,

[8] The outsized impact on rural communities is worth noting. For example, on August 19, 2024, there were 1,951 job postings on the American College of Radiology (n.d.) job board. Of those, 805 were located in an urban area, 775 in a suburban area, and 45 in a rural area (with 326 not classified).

particularly those based on machine learning and deep learning, have shown promise in diagnosing diseases by analyzing medical images, patient histories, and other data. However, these tools carry the risk of misdiagnosis due to algorithmic errors or biases. For example, AI systems that are trained on historical data that can be incorrect, incomplete, or corrupt can harbor biases that may lead to incorrect diagnoses, especially in underrepresented populations. That means an AI system trained primarily on data from white patients or teenagers might not perform as well when diagnosing conditions in patients of other ethnicities or age groups, leading to health disparities and potentially harmful outcomes.

The widespread opacity of AI algorithms—previously defined as the black box problem—complicates the issue further. Medical professionals or patients may find it challenging to understand or trust the AI's diagnostic decisions without clear explanations of how the algorithms reached their conclusions. Furthermore, there is the issue of accountability. When an AI system makes a misdiagnosis, it can be difficult to determine where responsibility lies: with the developers, the data scientists, or the healthcare providers who implemented the technology. These complexities necessitate rigorous oversight, continuous evaluation, and an emphasis on integrating AI in a way that complements rather than replaces human judgment in the diagnostic process.

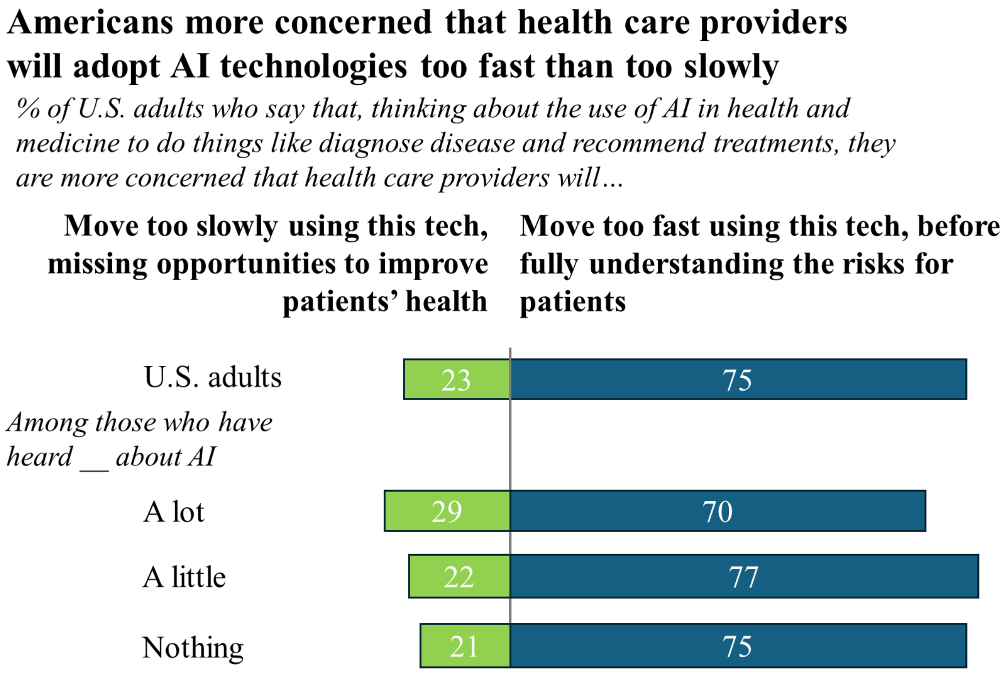
For example, oncology nurse Melissa Beebe has spent more than 15 years working with cancer patients, relying on her highly refined hard skills as well as qualitative observation abilities to make life-or-death decisions. When an AI-powered algorithm determined that her patient was septic, she knew this was a mistake—and a highly consequential one at that. The algorithm mistook an elevated white blood cell count with septic infection, as the AI completely disregarded the fact that this patient had leukemia, which triggers the variance in blood count. As *The Wall Street Journal* noted in dystopian detail:

> Hospital rules require nurses to follow protocols when a patient is flagged for sepsis. While Beebe can override the AI model if she gets doctor approval, she said she faces disciplinary action if she's wrong. So she followed orders and drew blood from the patient, even though that could expose him to infection and run up his bill. "When an algorithm says, 'Your patient looks septic,' I can't know why. I just have to do it," said Beebe, who is a representative of the California Nurses Association union at the hospital. As she suspected, the algorithm was wrong. "I'm not demonizing technology," she said. "But I feel moral distress when I know the right thing to do and I can't do it." (Bannon, 2023, paras. 4–5)

This example presciently illustrates both the danger for the patient receiving a misdiagnosis from an AI algorithm and how such applications can strongarm human practitioners out-of-the-loop and into the back seat at the mercy of flawed systems.

Moreover, practitioners might become overly dependent on AI recommendations, catalyzing a reduction in their critical thinking and diagnostic skills. This over-reliance can be particularly problematic if AI systems are used without proper oversight and continuous validation. In complex or ambiguous cases, the subtleties of human intuition and experience are crucial, and a solely AI-driven approach might overlook these nuances. Thus, if an AI system provides a diagnosis that a clinician disagrees with, it could lead to conflicts and confusion, ultimately affecting quality of care. This uncertainty can also undermine the doctor-patient relationship, as patients might be reluctant to accept diagnoses or treatment plans suggested by an AI they perceive as mysterious or untrustworthy.

Indeed, AI systems are prone to error and hallucination. As another example, an error-prone AI system may fail to detect certain rare—or even simple—conditions that human doctors can recognize through their experience and intuition. Furthermore, if an AI tool used in a hospital setting mistakenly prioritized less severe cases over critical ones, this could lead to delays in treatment for patients who need urgent care. These examples underscore the



Note. Chart reproduced from *60% of Americans Would Be Uncomfortable With Provider Relying on AI in Their Own Health Care*, by A Tyson, G. Pasquini, A. Spencer & C. Funk, 2023, Pew Research Center (https://www.pewresearch.org/wp-content/uploads/sites/20/2023/02/PS_2023.02.22_AI-health_REPORT.pdf).

importance of maintaining a balanced approach where AI assists but does not replace human clinicians. Continuous training and the development of protocols to address AI errors are essential to ensure patient safety and care quality.

Data Privacy and Cybersecurity

AI's role in personalized medicine, which involves tailoring treatments to individual patients based on genetic, environmental, and lifestyle factors, raises significant privacy and ethical concerns. Personalized medicine relies heavily on large data sets that include sensitive personal health information. The collection, storage, and analysis of these data pose risks to patient privacy. Despite advancements in data security, there is always a potential for breaches or misuse of data, especially because healthcare systems are a top target for cyberattacks (Takahama, 2024). If sensitive information is leaked, it can lead to severe consequences, including discrimination in employment or insurance based on genetic predispositions to certain diseases. A broad theme throughout this paper is that AI as a field is *driving* the rise in more data, and AI in its technological application is optimizing efficiency and outcomes through leveraging these data. Ultimately, more data across the board means more vectors for cyberattacks.

Moreover, data privacy concerns extend beyond personalized medicine. For example, wearable digital health technology, such as fitness trackers and smartwatches, often collect extensive personal health data, including heart rate, activity levels, and sleep patterns. These data are typically stored in the cloud and shared with third-party apps and services, which can lead to significant privacy concerns. In one high-profile case, the fitness tracking app Strava inadvertently revealed the locations of military bases and personnel by sharing user data publicly (Pérez-Peña & Rosenberg, 2018). Additionally, companies like Fitbit have faced lawsuits for allegedly selling user health data to third-party advertisers without proper consent (Evans, 2023).

Incidents like these underscore the need for security and privacy by design in new technologies, stringent data privacy regulations, and robust cybersecurity measures to protect sensitive health information. Without proper oversight, the misuse of personal health data can lead to severe consequences, including identity theft, discrimination, and unauthorized data sales. As AI continues to integrate into healthcare, ensuring the security and privacy of patient data must remain a top priority to maintain trust and safeguard individual rights.

Ethical Questions

Personalized medicine often involves genetic testing, which can reveal information not only about the patient but also about their family members. This raises ethical questions about consent and the right to privacy. For instance, discovering a genetic predisposition to a hereditary disease can affect not just the individual tested but also their relatives, who may not have consented to the revelation of this knowledge. The ethical implications of such revelations are profound, and managing the responsibility of sharing genetic information must be handled with the utmost care and consideration for all affected parties.

Another ethical consideration of the use of AI in medicine is in determining end-of-life scenarios in medical emergencies. One major concern is the lack of empathy and human judgment in AI systems. End-of-life decisions are deeply personal and often require a nuanced understanding of the patient's values, family dynamics, and emotional state—factors that AI cannot fully grasp or incorporate. An automated AI system could be left to determine when to "pull the plug" on a terminal medical patient while not considering other information related to that patient's life or the full consequences of such a decision. This sort of decision should only be decided by the family or the individual in question in consultation with their doctor, never by a machine.

Public Sentiment

The current public sentiment surrounding AI is not a negative use case but rather a broader reflection of how the public is responding to contemporary stories about specific use cases. According to a Pew Research Center poll, 60% of Americans would be uncomfortable with their medical provider relying on AI to provide their healthcare. Only 38% felt that AI would lead to better health outcomes. Furthermore, 57% percent said the use of AI in diagnosis and treatment recommendation would harm their relationship with medical staff. As it relates to data privacy, 37% felt that the increased use of AI would lead to greater risk of their medical data security being compromised, 22% thought data privacy would improve, and 39% thought it would not make much difference (Tyson et al., 2023).

For many Americans, the field of healthcare is the arena where trust, relationships, dependability, and safety are of greatest importance. It therefore follows that, in instances where AI can be applied to liberate doctors, nurses, and medical practitioners to be both better informed about their patients and have more time to spend with them directly, emerging technology can greatly enhance the welfare of patients and their overall trust and optimism in the healthcare system. Leveraging AI in a manner that reduces the overhead of routine, non-human facing tasks can enhance the dignity of the professional and the patient. But as this section has made starkly evident, a permissionless application of AI in healthcare would introduce pernicious outcomes for patients—both in terms of patient trust and expectations in their healthcare experiences, as well as in terms of outcomes. Especially given the advent of AI-powered medical device integration with humans (e.g., brain computer interface company Neuralink), the field of healthcare is one where public policy guardrails need to be proactively established to stave off cataclysm. In reflecting on the now ubiquitous trope from Thomas Sowell, "There are no solutions. There are only trade-offs," we perhaps come to better expound this statement and introduce a comically frustrating paradox: the sectors where AI stands to most greatly advance humanity are also the sectors that run the greatest risk of driving humanity to extinction.

Education

The choice of how and where to educate one's child is deeply personal to the parent or caretaker and should be unique to that child. From class sizes to the classroom environment to what technologies are used, there are an increasing number of options for how to educate one's child. Expanded broadband access, digital and hybrid classrooms, and other emerging technologies introduce amazing opportunities to learn and prepare for life beyond the classroom. Not surprisingly, the age of AI has raised great promise—and great pause—among students, parents, educators, and administrators.

For example, AI could increase efficiency for administrators, teachers, and students by helping modernize common processes that otherwise involve a steep time cost. Programs could allow for better problem solving based on individualized needs of each student, facilitate greater long-term success, and help support the most fundamental goals of the education system. While AI offers the potential to decrease the time costs associated with education, it also presents ethical questions about academic integrity and personal responsibility. School boards, administrators, and teachers must consider how AI could frustrate the educational goals and outcomes by substituting students' ownership and command of subject matter with GAI proficiency and prompt engineering abilities. Indeed, this is true of any technology in the classroom, from GAI to laptops to tablets. However, a unique challenge is posed by AI's potential to replace critical thought and supplement students' ability to understand, write about, and defend their ideas and beliefs.

AI also places school officials in a challenging position as they look to catch up and avoid becoming outpaced by their students. The integration of AI into the classroom and basic education operations can streamline tedious processes and give educators and students the opportunity to focus on more critical practices. Automation in education (as in other industries) is a double-edged sword. Like all advances in technology, users should look before they leap to ensure responsible use and thoughtful regulation. Without careful parameters, AI's use in education will be handicapped and its success limited.

K–12 Education

AI has several potential applications in the K–12 setting. First, AI can be used for grading assignments, exams, and even standardized tests like the State of Texas Assessments of Academic Readiness (STAAR) tests. The STAAR is a series of standardized tests used to assess a student's knowledge at a certain grade level. In 2023, the Texas Education Agency (TEA) announced the introduction of AI to grade written portions of the exams. This change will result in administrators grading 25% of the written portions of exams and AI grading 75% (Simmons, 2024). According to TEA (2024), this will shorten the time it takes to grade papers and save taxpayers $15–$20 million per year as fewer human graders are needed.

Scoring systems would be hybrid, meaning that AI would grade most responses and tag instances where the program has "low confidence" in accurate scoring. After these answers are identified, they are automatically reassigned to a human grader. Hybrid scoring systems have already been put into effect across select districts and are being used sparingly. However, uncertainty surrounding the use of AI programs continues as teachers contemplate whether these systems will work fairly across districts, and students fret over whether they will be scored accurately. Additionally, hybrid systems present a challenge for administrators who may be unable to determine if a low score is attributed to the construction of test questions or the automated scoring system. This confusion poses challenges for administrators' ability to generate and review test analytics in the hope of improving test formation. Despite these concerns, TEA (2024) reported that automated scoring systems are not new to Texas, are used in over 21 other states, that it has tested this system, and that the quality of scoring will remain the same.

Second, AI can be used by teachers to help write curriculum and aid in coursework prep. For example, of teachers currently using AI platforms, 80% use virtual learning platforms (like Google Classroom), more than 60% use adaptive learning systems (like Khan Academy), and more than 50% use GAI systems like ChatGPT (Diliberti et al., 2024; Arundel, 2023). Furthermore, programs like DISCO can gather and process data, synthesize information, and generate an entire curriculum for teachers (DISCO, 2024). AI lesson generation is especially useful for

English language arts and social studies due to their open-ended and multifaceted nature. AI programs can simplify lesson planning and curriculum building by helping generate standards reviews or essential questions. However, as with human writing, AI can display biases, especially based on input data and system design. Thus, this innovation offers a high potential for efficiency but cannot displace the human touch of teachers to educate and gauge student comprehension.

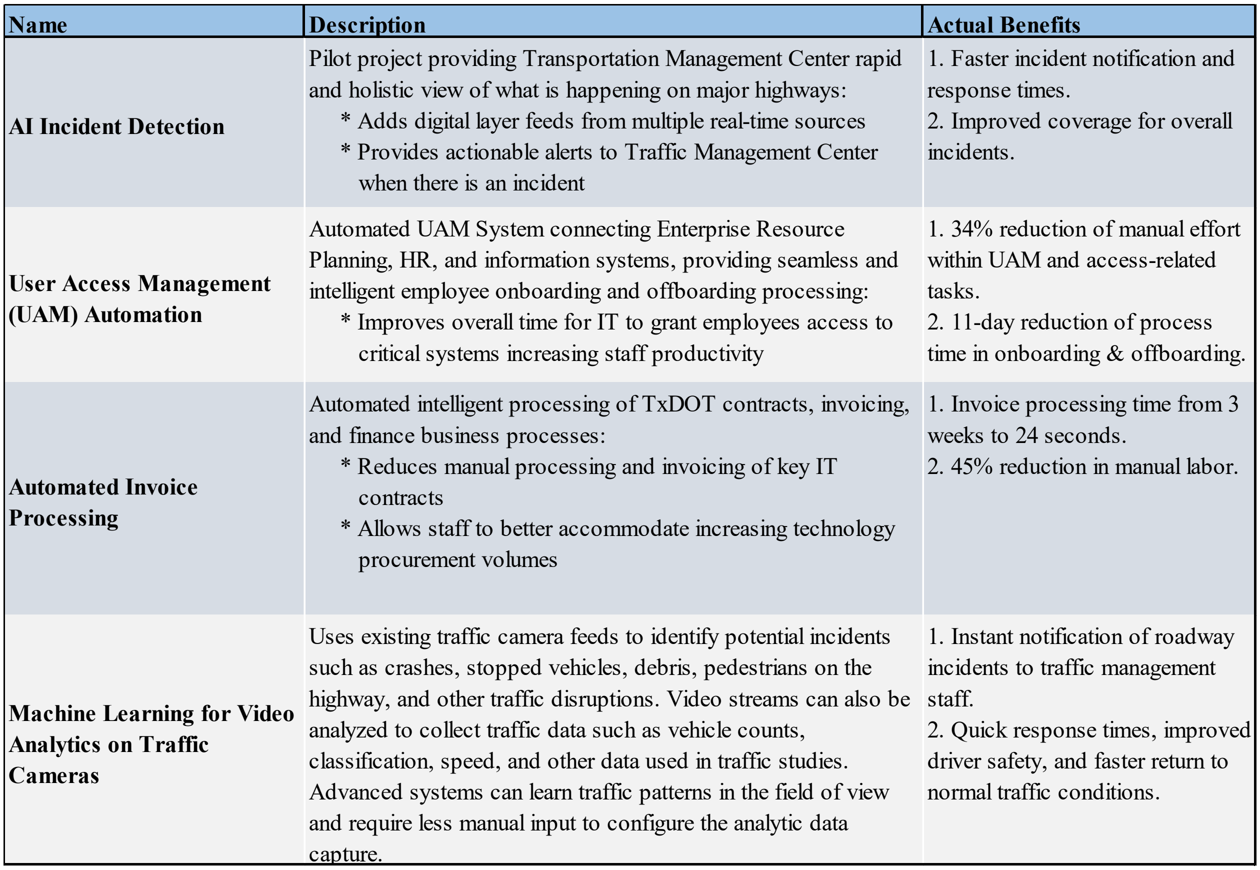
Finally, those in the K–12 space, specifically students, can utilize AI programs that provide diagnostic automated tutoring. For example, Intelligent Tutoring Systems (ITSs) provide personalized learning and guidance to students, helping them when their teachers are not available. The AI methodology behind ITSs has four components: knowledge base (having to do with subject-specific information), student model (the student's current level of understanding), pedagogical module (how educational tactics are adapted to student's learning stages), and user interface (communication between ITSs and students) (The Princeton Review, n.d.). Because of GAI, students can engage with information tailored to their needs and limitations. Though ITSs offer personalized programs to students, they have the potential to conflict with teachers' structuring of lesson plans. AI should not replace the daily lectures of teachers and their explanations or timing, even though it can augment them. To maintain integrity in the utilization of AI, students must understand the limitations of these programs.

Higher Education

As with K–12 education, AI can be used in higher education use by administrators, faculty, and students alike. Administratively, AI is increasingly utilized by admissions offices. For example, public higher education institutions are increasing their use, with 55% currently utilizing AI in admissions and 34% more planning to do so. Private higher education has been more tempered, with 38% of admissions offices currently using AI and 43% planning to do so (Intelligent, 2023). Admissions workers use AI to review and assess applications based on the composite application materials of potential students. The thousands of applications that any given school receives can be funneled through an AI program, screened, and then reviewed (reducing time cost). While there is hope that AI can help remove human bias, others are concerned for the same reason. Concerns over training data, model design, a lack of standardized guidelines, and, ultimately, the lack of a human touch all worry prospective students and admissions counselors alike (Chandra, 2024).

Second, AI offers a greater degree of individual adaptability in higher education by equipping college and university faculty with tools to better engage students. At present, the deficit in use of AI between students and faculty is substantial, with nearly 50% of students and more than 20% of faculty utilizing AI tools, according to one study (Shaw et al., 2023). However, AI-based learning systems offer more tailored teaching options, giving faculty the ability to customize lesson plans and generate methods based on students' individual needs. For example, campuses associated with Indiana's Ivy Tech system have used these AI mechanisms to help struggling students and account for those who may need accommodation due to learning differences (Google for Education, n.d.). The one-size-fits-all model does not work, and, indeed, students are more likely to achieve long-term success with the development of individualized accommodations. However, with the growth of these programs comes the need for a significant amount of personal data (Rouhiainen, 2019). Consonant with other Texas data privacy laws, data privacy, cyber hygiene, and cybersecurity must be paramount concerns and must precede mass adoption of AI.

Finally, students can utilize AI to improve their comprehension at both the undergraduate and graduate levels. Similarly to K–12 education, one of the main uses of AI is personalized tutoring and analytics. For example, GAI can create flashcards and study materials, targeting the areas where students need to build comprehension. It can also help "grade" and give instantaneous feedback on problem sets, sample essays, and more. However, like in the K–12



Name	Description	Actual Benefits
AI Incident Detection	Pilot project providing Transportation Management Center rapid and holistic view of what is happening on major highways: * Adds digital layer feeds from multiple real-time sources * Provides actionable alerts to Traffic Management Center when there is an incident	1. Faster incident notification and response times. 2. Improved coverage for overall incidents.
User Access Management (UAM) Automation	Automated UAM System connecting Enterprise Resource Planning, HR, and information systems, providing seamless and intelligent employee onboarding and offboarding processing: * Improves overall time for IT to grant employees access to critical systems increasing staff productivity	1. 34% reduction of manual effort within UAM and access-related tasks. 2. 11-day reduction of process time in onboarding & offboarding.
Automated Invoice Processing	Automated intelligent processing of TxDOT contracts, invoicing, and finance business processes: * Reduces manual processing and invoicing of key IT contracts * Allows staff to better accommodate increasing technology procurement volumes	1. Invoice processing time from 3 weeks to 24 seconds. 2. 45% reduction in manual labor.
Machine Learning for Video Analytics on Traffic Cameras	Uses existing traffic camera feeds to identify potential incidents such as crashes, stopped vehicles, debris, pedestrians on the highway, and other traffic disruptions. Video streams can also be analyzed to collect traffic data such as vehicle counts, classification, speed, and other data used in traffic studies. Advanced systems can learn traffic patterns in the field of view and require less manual input to configure the analytic data capture.	1. Instant notification of roadway incidents to traffic management staff. 2. Quick response times, improved driver safety, and faster return to normal traffic conditions.

Note. Chart reproduced from *AI at TxDOT*, by A. Selissen, Texas Department of Transportation, 2024, p. 6 (https://txdir.widen.net/s/bdkgjhvwcd/3.-20240321_txdot_ai_advisory_council).

environment, these advances must be balanced with the need to impart critical thinking, work ethic, and personal responsibility. For example, programs like Readocracy aim to recapture what has been lost of the attention economy (Readocracy, n.d.). They are responding to this concern by including mechanisms for measuring comprehension and application, such as tests, discussion groups, and peer review of work submitted and reading assignments completed. While programs like Readocracy are still in development and not yet implemented on a large scale, the conversation surrounding AI's use in higher education is ever-growing, inspiring discussions on academic integrity, personal responsibility, and the purpose of education itself (Educause, 2024). AI certainly should be utilized as a tool to aid in student comprehension and lighten the load of both teachers and administrators, but it cannot and should not replace the human elements of education, only augment them.

Transportation

The transportation sector has evolved dramatically in recent years as part of the broader push for digital transformation in the public and private domains. Given the breadth, economic gigantism, and data-rich environment that is transportation—including everything from traditional automobile transportation, airlines, logistics, shipping, warehousing, ports, and more—government and private actors have invested significant time and effort in incorporating AI-powered transportation technologies into their operations. Importantly, because of the personal and tangible nature of transportation, this domain is shaping up to be one of the first in which the general public must contend with the reliability and safety of AI systems for essential tasks (Stone et al., 2016).

The private automobile sector has long been incorporating AI-driven solutions. In fact, AI has been integrated into the way we use our personal vehicles for decades. In 1985, Etak released the first ever computerized navigation system for automobiles. This was so far ahead of its time that it would be another 10 years before the U.S. government operationalized its satellite-based Global Positioning System (GPS) (Edwards, 2015). As navigation apps and computers began incorporating data from satellites, machine learning and AI tools were employed to work through massive, dynamic data sets to map out the most effective way to get from point A to point B (Let's Talk Science, 2023). Over time, navigation tools and apps have become more responsive and efficient, using AI to incorporate and react to real-time traffic data, historical crash data, reported accidents, and other variables to produce the most efficient route guidance (Forristal, 2023). Furthermore, this same technology that responds to rapidly changing, dynamic roadways is fueling the demonstrable improvements in autonomous vehicle technology (Dunmoyer, 2024).

From a public perspective, the Texas Department of Transportation (TxDOT) provided testimony before Texas's Artificial Intelligence Advisory Council in March 2024, providing insights into how AI is being leveraged by the state's leading transportation authority for improvements to the public sector's approach to transportation efficiency, safety, and more. The following chart highlights its current AI initiatives.

Building on TxDOT's use of machine learning for video analytics on traffic cameras, cities across Texas are growing more ambitious with regional technology corridors, or "digital corridors." In early 2024, the North Central Texas Council of Government's Regional Transportation Council announced what will be at least a $16 million project to "create a centralized hub for managing transit data, operations and digital infrastructure on Interstate 30 in Dallas-Fort Worth" (Gaudet, 2024, para. 3). This hub would allow for the intake and processing of huge swaths of data—including TxDOT's traffic management and work zone data, highway corridor cameras, and other data from third parties like Google—to in turn optimize general traffic flow throughout the metroplex by altering the timing of traffic lights, directing traffic flow, synchronizing autonomous vehicle operations, and more.

The above smart corridor example points to a broader push for "smart cities," which has become practically synonymous with urbanization. In essence, a city is characterized as "smart" when it leverages information and communication technologies to enhance the efficiency of operations and management, while also improving information sharing for the purpose of benefitting citizen welfare and government services (Digi.City, n.d.). AI tools are being employed across the board with prototypical and entrenched smart city tools, including in the transportation sector. For example, according to Lockhart (2024),

> AI-powered adaptive traffic management can dynamically adjust signal timings based on real-time traffic patterns, incidents and weather conditions, reducing greenhouse gas (GHG) emissions and improving emergency response times. These systems embody the principles of sustainability that are crucial to the smart city vision. (para. 5)

Finally, beyond roads, AI is also being incorporated into the broader transportation sector, from air freight and logistics to ports. For example, ports across the world are beginning to adopt technologies in pursuit of becoming a smart port—one that aims to boost the efficiency and safety of ports through the application of technology and data-driven solutions. Given the mass sums of data port operations generate, they have proven good candidates for leveraging AI for routine tasks. Machine learning is being used to automate certain port operations, learning from its data to predict job durations for specific ships and services. The Port Authority of Los Angeles has long been leveraging a cloud-based application called Port Optimizer to serve as a singular repository for stakeholders to share valuable data. Port Optimizer has since added new tools, which use machine learning to sort through huge volumes of data to ensure all information provided to customers is timely, accurate, and problem free.

Future iterations are currently being tested that would allow Los Angeles to predict disruptions and map out how to most effectively prevent ripple-effect delays (German, 2023).

In Singapore, AI data analysis is currently being used to improve port productivity by sifting through data on incoming vessels, their size, load specifics, and the like, to ascertain the best berth for ship docking while suggesting pathways that can best prevent ship collisions (German, 2023). While these maritime applications are not as visual as self-driving vehicles, such logistical improvements across all sectors will significantly increase efficiency, thus saving time and money in the long term.

In Texas, thus far, many of the public sector applications of AI in transportation have been employed ethically and responsibly, focusing on automating mundane tasks and keeping humans in the loop where common sense warrants it. As Texas lawmakers contend with a potential digital code of ethics for the public sector's use of AI, this is a sector where Texas is remarkably well positioned to lead nationally given the state's regulatory environment, existing infrastructure, talent-rich workforce, and more. However, as one of the paper's authors noted in a separate research paper, the sometimes hasty private sector introduction of AI into transportation through personal autonomous vehicles and similar technologies poses risks to human agency and individual autonomy, which demands a public policy response to protect freedoms while maximizing safety (Dunmoyer, 2024).

The Justice System: Law Enforcement, Public Safety, the Judicial System, and Courts

Law Enforcement and Public Safety

There are numerous potential AI use cases in law enforcement and public safety, including real-time analysis of security camera and body cam footage, facial recognition, fingerprint matching, speech recognition, gunshot detection, redaction and characteristic blurring in videos and photos, dispatching first responders, predictive analytics and hotspot mapping, cybersecurity and protecting critical infrastructure, automated license plate readers, robotics and autonomous vehicles, administrative functions, customer-facing services, and building community trust (Finklea, 2023; Redden et al., 2020a; Rigano, 2019).

For example, as reported by Murakami, Miami assistant police chief Armondo Aguilar testified before Congress that his department "uses AI in a number of ways, including facial recognition, reading license plates, monitoring potential threats on social media and using ballistic evidence to 'connect the shots' between shootings" (Murakami, 2024, para. 9). Chief Aguilar also cited a study that found "detectives have a 66% greater chance of finding suspects in violent crimes when using [AI]" (para. 10).

However, it is critical to recognize the application of AI in law enforcement introduces tremendous threats to the constitutional rights, privacy, and civil liberties of Texans. Accordingly, of all the areas for potential AI regulation, this should be near the top of the list for lawmakers. Ultimately, lawmakers should consider the varied perspectives of stakeholders and ensure use cases are limited, narrowly tailored, and primarily uphold civil liberties over exigency.

One of the key and emerging applications of AI in this field is gun and gunshot detection. For example, AI software can be incorporated into security camera systems to more readily determine if an individual is brandishing a weapon to help prevent or ameliorate a mass casualty event. The technology is advancing rapidly and can determine the difference between a firearm or long knife and another object (e.g., a cellphone, flashlight, or water bottle), flag concerns for security officers, notify law enforcement within seconds, and send out mass alerts (ZeroEyes, n.d.). Furthermore, AI can be used to detect gunshots and differentiate between real gunshots and other loud noises caused by construction, manufacturing, car backfires, and the like. These technologies can also be overlaid onto existing video and audio feeds to help detect gunfire, dispatch law enforcement, and more accurately locate crime scenes and even shell casings (SoundThinking, n.d.). However, McCullom (2024) highlights the costs, concerns, and limitations some cities and law enforcement offices see with this technology, including some cities that have abandoned its use altogether.

Second, AI can assist in blurring faces and redacting certain characteristics in videos and photos, particularly body cam footage released shortly after an interaction, but before any criminal, civil, or administrative action has been taken. This privacy aspect is particularly important for the accused to ensure due process, for law enforcement and first responders in their line of work, and important to protect the safety of witnesses, victims, and survivors (Sisson, 2024; LaPedis, 2023). Privacy is important in itself; however, it has a heightened importance when law enforcement is involved and constitutional rights are implicated.

Third, AI can be used to assist with numerous operational, investigative, and administrative functions, increasing efficiency, reducing human error, and expediting the investigative process. For example, officers can use voice-to-text and generative AI functions to take eyewitness statements, summarize notes, and draft incident reports (Redden et al., 2020a). Officers can also access, categorize, and analyze data such as police reports, witness statements, crime scene photos, body cam video, crime statistics and trends, and the like (Hsiung & Chen, 2023). Finally, law enforcement offices could use AI to process invoices, fines, fees, and other administrative tasks, as some state and local governments have done more efficiently in their finance departments. For example, one Texas agency has reduced the processing of invoices from two to three weeks to 13 seconds, and a Utah city was able to reduce the time it takes to process 1,200 invoices per month from 12 hours to one hour (Downey, 2024a; Edinger, 2023).

Finally, AI can be used for public communications, customer-facing operations, and building community trust. According to Camello et al. (2021), "chatbots in the criminal justice system have the potential to improve efficiency, redefine engagement, expand access to justice, and reduce costs associated with administrative overhead for various criminal justice stakeholders" (p. 1). For example, generative AI and chatbots are being used for tiplines, identifying and routing emergencies to first responders, round-the-clock communications, automated FAQ support, providing access to community resources and referrals, law enforcement recruitment, and even translation services (Camello et al., 2021; Norris, 2019; Douglas, 2018; Futr, n.d.).

However, using AI in law enforcement and public safety presents real challenges to the privacy, civil liberties, and constitutional rights of Texans, which is a specific charge for the Texas AI Advisory Council to study (HB 2060, 2023, Sec. 2054.622(f)(2)(A)). Among these concerns is bias in training data and algorithms, facial recognition, fingerprint matching, DNA analysis, predictive analytics, hotspot mapping, automated license plate readers, and more.

Facial Recognition

The use of facial recognition technologies raises questions about privacy, accuracy, and bias. As with data privacy issues more broadly, there are baseline questions to ask about what data are being collected, from whom they are being collected, how they are being collected, the cyber hygiene of data storage, and if they are being sold or otherwise transferred. For example, Clearview AI has come under scrutiny for "scraping social media sites for photos and building a more than 3 billion-photo database it sells to law enforcement" (Statt, 2020, para. 7). This raises two concerns for legislators to consider. First, legislators could expand data privacy protections—building on the great work of bills like HB 4 (2023) and SB 2105 (2023)—to further protect personal data and reign in abusive practices by data brokers. Second, legislators could take serious steps to limit or even prevent government agencies and law enforcement from purchasing or otherwise obtaining such data,

Note. AI-driven facial recognition presents serious concerns about data privacy, biometric and other sensitive information, poor accuracy, and bias. Photo from *Microsoft: Here's Why We Need AI Facial-Recognition Laws Right Now*, by L. Tung, ZDNET, December 7, 2018 (https://www.zdnet.com/article/microsoft-heres-why-we-need-ai-facial-recognition-laws-right-now/).

especially when done as effectively bypassing the Fourth Amendment of the U.S. Constitution (Ayoub & Goitein, 2024).

Furthermore, accuracy and bias are also a serious concerns, which can lead to disparate impacts. On the front end, there are concerns that the input data is misrepresentative of the overall population, with Lohr (2018) noting that "one widely used facial-recognition data set was estimated to be more than 75 percent male and more than 80 percent white" (para. 6). Furthermore, Swarns (2023) reported that

> given the disproportionate rate at which African Americans are subject to arrest, the center [Georgetown University's Center on Privacy & Technology] found that facial recognition systems that rely on mug shot databases are likely to include an equally disproportionate number of African Americans. (para. 6)

There are also concerns about the bias in the algorithms, as the outputs can be troubling. For example, a 2019 National Institute of Standards and Technology study found that Asian and black people "were 10 to 100 times more likely to be misidentified than white people" (Swarns, 2023, para. 7; National Institute of Standards and Technology, 2019). Reporting on Buolamwini and Gebru's (2018) highly influential, often-cited paper, Statt summarized that the researchers "found error rates for facial recognition systems from major tech companies … for identifying darker-skinned individuals were dozens of percentage points higher than when identifying white-skinned individuals" (Statt, 2020, para. 3). There were also disparate outcomes in gender identification. Also reporting on the Buolamwini and Gebru study, Lohr (2018) noted,

> when the person in the photo is a white man, the software is right 99 percent of the time. But the darker the skin, the more errors arise—up to nearly 35 percent for images of darker skinned women, according to a new study that breaks fresh ground by measuring how the technology works on people of different races and gender. (paras. 2–3)

These concerns were so pronounced that by 2020, IBM shut down its facial recognition programs and declared it will no longer "support the use of facial recognition for mass surveillance, racial profiling, or other human rights violations" (Wilding, 2023, para. 6). Furthermore, Amazon instituted a "a one-year moratorium on allowing law enforcement to use its controversial Rekognition facial recognition platform" (Statt, 2020, para. 1). Due to the absence of federal rules to ensure the ethical use of such facial recognition technology, Amazon opted to indefinitely extend the moratorium (American Civil Liberties

Union, 2021). However, 2024 Department of Justice disclosures indicate that the FBI is now working with Amazon's Rekognition in the "initiation" phase, with details still uncertain (Heilweil & Alder, 2024).

Predictive Analytics

Another potential use case for AI is in predictive analytics. According to Fitzpatrick et al. (2019), "predictive analytics in policing is the practice of forecasting crime patterns across time and space to inform decision-making for crime prevention" (p. 474). Not only can AI-powered predictive analytics aid law enforcement, but it can also be used in the context of Child Protective Services (CPS).

Data privacy is a key theme throughout this paper and must remain a cornerstone of the discussion of AI integration in predictive analytics. The use of AI for predictive analytics could revolutionize these facets of civil safety but could also be disastrous if ethical questions remain unanswered. However, while the specter of privacy infringement looms, AI could benefit law enforcement's ability to identify high-crime areas for the purpose of mitigation. Furthermore, for CPS workers, programs could revolutionize a caseworker's ability to manage cases better and promote child safety.

For those in law enforcement, AI integration into predictive analytics offers the opportunity for innovative police services, community connection, and more targeted crime deterrence. Much of the projected use of these kinds of AI-based predictive analytic systems can enhance real-time crime analytics and keep cities safe. For example, as Sloly (n.d.) noted, the use of smart technologies like AI could help cities reduce crime by 30%–40% by identifying the most at-risk areas for serious crime. These analytics could also serve to make response time more efficient, enabling 20%–35% faster responses from emergency services. Furthermore, the beneficial "see something, say something" motto yields large amounts of data from 911 phone calls, text messages sent to tiplines, and photos shared to police social media that can overwhelm law enforcement. However, AI integration into Next Generation 911 (NG911) can help law enforcement officers and dispatchers more efficiently sift through the data, thus freeing up time to more effectively allocate resources. Finally, Sloly noted its potential to be reactive rather than proactive—and even to stop crime before it happens (although the movie *Minority Report* causes one to question that goal).

In the CPS context, many uses of AI predictive analytic systems provide a primary benefit of streamlining information and, subsequently, streamlining protocol. For example, AI systems can review historical case data, identify patterns, and predict the likelihood of specific outcomes for an individual child (Teixeira & Boyas, 2017). These analytics can help caseworkers reduce administrative overhead, more efficiently allocate time, prioritize high-risk or high-needs cases, and intervene when, where, and how they are most needed (Govindiah, 2023). At least 11 states currently use predictive analytics tools in the CPS setting and nearly half have considered using it. In New York City, for example, AI-driven tools are being used to assess which children are most likely to be mistreated or abused. One such tool, called the Severe Harm Predictive Model, "predicts the odds that a child in an open case will be the subject of substantiated allegations of physical or sexual abuse within the next 18 months" (McSilver Institute, 2021, para. 4). Tools like this could help remove children from imminent danger.

However, like other AI applications, such tools can produce biased results based on poor input data or model design. As many of these analytic tools are being introduced at the local and state levels, they rely on local or state government data. A 2021 panel at New York University noted concerns about the accuracy of data inputs and outcomes, many of which showed a disparity based on racial demographics. Finally, AI-driven models lack a personal element that some caseworkers may acquire through experience. For example, a caseworker may be better at picking up on emotional and interpersonal dynamics than a purely data-driven AI system.

Automated License Plate Readers

A final law enforcement use case worth considering is that of automated license plate readers (ALPRs).[9] According to the Electronic Frontier Foundation (EFF) (2023), ALPRs are

> high-speed, computer-controlled camera systems that are typically mounted on street poles, streetlights, highway overpasses, mobile trailers, or attached to police squad cars. ALPRs automatically capture all license plate numbers that come into view, along with the location, date, and time. The data, which includes photographs of the vehicle and sometimes its driver and passengers, is then uploaded to a central server. (para. 1)

License plate readers (LPRs) are not new. According to Spinks (2023), "in 1976, the Police Scientific Development Branch in Britain invented the first license plate readers to combat terrorism and by 1979, there were working prototypes" (para. 1). Spinks noted that "before the advent of automated systems, law enforcement officers visually inspected and obtained wants, warrant and registration information via radio, which was time-consuming and prone to errors" (para. 2).

The first use of ALPRs by American police departments appears to have been in the 2000s (Simonite, 2020). Law enforcement argued that ALPRs "enhance their enforcement and investigative capabilities, expand their collection of relevant data, and expedite the tedious and time consuming process of comparing vehicle license plates with lists of stolen, wanted, and other vehicles of interest" (International Association of Chiefs of Police, n.d., para. 1). For example, an audit of the Lexington, Kentucky, ALPR system found that

> the technology has "significantly helped in the recovery of stolen vehicles," cutting the time victims waited for their vehicles to be recovered by police in nearly half—from more than 10 days to a little over five. The city reports the total value of vehicles recovered with the technology is more than $3.7 million, and that Flock technology has also helped locate missing people and enable the seizure of illegal firearms. (Davidson, 2024, para. 4)

Note. Examples of ALPR-capable cameras. Photo from *ALPR Cameras for Public Safety*, by MCA, n.d. (https://callmc.com/mss/vehicle-upfitting/vehicle-video-camera-systems/alpr-and-lpr-cameras/).

While sophisticated ALPR units can cost $15,000–$20,000, newer technologies use "software that can discern plates from more or less any conventional security camera" for "as little as $50 a month to read plates from a single camera" (Simonite, 2020, para. 5). This technology, for example, allowed a small New York town's police force of 45 officers to log, every day, "the license plates of around 10,000 vehicles moving through and around town, using software plugged into a network of cameras at major intersections and commercial areas" (para. 1).

However, some argue that increases in technological sophistication, computational speed, data storage, and battery power, coupled with reduced camera sizes, have turned ALPRs into real-time tool of "mass surveillance" (Stanley, 2022, p. 1). From a privacy and civil liberty perspective, the scope of use and volume of data being collected is a serious

[9] The International Association of Chiefs of Police (n.d.) used a less anthropomorphic phrase: automated license plate *recognition*.

concern. For example, a 2016 article in *The Atlantic* reported that a major license plate reader system, Elsag, "claims its cameras can capture up to 1,800 license plates a minute during day or night, across four lanes of traffic and at speeds up to 150 miles per hour, alerting officers 'within milliseconds' if a plate is suspect" (Waddell, 2016, para. 3). More recently, one vendor, Flock, "specializes in automated license plate recognition and video surveillance, and already has a fleet of around 40,000 cameras spanning 4,000 cities across 40 states" (Hammer, 2024, para. 3). According to EFF, Digital Recognition Network (DRN), a Texas-based vendor, "brags that its dataset includes more than 6.5 billion scans and grows at a rate of 120-million data points each month" (EFF, 2023, "What Kinds of Data an ALPR Collects" section). In October 2024, DRN's website noted that it conducts over 350 million license plate scans per month (Digital Recognition Network, n.d.).

While this has been a boon to vendors, the data broker and insurance industries, investors, and law enforcement surveillance, the cost is the privacy and liberty of Americans. In an interview with Polcyn (2023), Flock CEO Garrett Langley was asked, "Do you envision a future with a Flock camera on every street corner?" (para. 6). As Langley affirmed, "I envision that. … And I envision an America where crime no longer exists" (para. 7). The authors of this paper posit a rhetorical question: What could possibly go wrong?

The following example shows the outcome of such a system. On March 10, 2022, David Zayas was pulled over, searched, and found in possession of crack cocaine, a pistol, and $34,000 in cash. He later pleaded guilty to a drug trafficking charge. What is striking for our discussion here is when, why, and how he was pulled over. Forbes reported that by

> searching through a database of 1.6 billion license plate records collected over the last two years from locations across New York State, the AI determined that Zayas' car was on a journey typical of a drug trafficker. According to a Department of Justice prosecutor filing, it made nine trips from Massachusetts to different parts of New York between October 2020 and August 2021 following routes known to be used by narcotics pushers and for conspicuously short stays. (Brewster, 2023, para. 2)

It was not speeding, erratic driving, or that the police directly observed suspicious behavior or the commission of any crime. Rather, it was because his car was flagged by AI-driven cameras for "'suspicious' patterns of movement" over an 11-month period (Brewster, 2023, subheadline section). It is fair to concede that Zayas is a criminal. And while it's troubling that this data collection and ALPR-driven surveillance is abused to combat criminal activity, it is perhaps more troubling that it can be used to surveil the behavior of millions of everyday Texans with no connection to crime. As EFF (2023) noted,

> Taken in the aggregate, ALPR data can paint an intimate portrait of a driver's life and even chill First Amendment protected activity. ALPR technology can be used to target drivers who visit sensitive places such as health centers, immigration clinics, gun shops, union halls, protests, or centers of religious worship. (para. 3)

Yet today, it is not just law enforcement and public cameras that are being used to collect data. Vendors contract with private vehicles and fleets, and at least four multi-billion-dollar companies have reportedly been transformed into police surveillance proxies. Hammer (2023) reported that FedEx[10] and Kaiser Permanente contract with the camera vendor Flock for surveillance of their facilities. Data from their surveillance feeds are shared with law enforcement, as several American police departments confirmed. However, Hammer noted that "it is currently unclear just how far-reaching the partnership between law enforcement and FedEx actually is or how much Flock data is being shared" (para. 27). Because FedEx and Kaiser Permanente are private companies, such data collection and transfer arrangements are not per se illegal. However, this sort of "cat's paw" end-around the U.S. Constitutional is problematic.

[10] A spokesperson clarified that while FedEx does contract with Flock, "FedEx vehicles do not have these ALPR cameras on them. FedEx only uses this video technology in facility parking lots to safeguard employees and property" (Hammer, 2023, para. 7).

In an interview with Forbes, an EFF attorney argued that from a privacy and transparency perspective it could "[leave] the public in the dark, while at the same time expanding a sort of mass surveillance network."

In an interview with Forbes, an EFF attorney argued that from a privacy and transparency perspective it could "[leave] the public in the dark, while at the same time expanding a sort of mass surveillance network" (Brewster, 2024, para. 3).

It is also important to understand the retention, propagation, accuracy, and value of such mass surveillance. The 88th Legislature treated these aspects of the issue as core concerns in broader privacy debates over bills (now laws) such as HB 4 (2023), HB 18 (2023) and SB 2105 (2023), for example. As EFF (2023) noted,

> Most of this ALPR data is stored in databases for extended periods of time—often as long as five years. The databases may be maintained by the police departments, but often they are maintained by private companies such as Vigilant Solutions or Flock Safety. Law enforcement agencies without their own ALPR systems can access data collected by other law enforcement agencies through regional sharing systems and networks operated by these private companies. Several companies operate independent, non-law enforcement ALPR databases, contracting with drivers to put cameras on private vehicles to collect the information. These data are then sold to companies like insurers, but law enforcement can also purchase access to this commercial data on a subscription basis. ("ALPR Databases" section)

Furthermore, the law enforcement and public safety value of the data is limited. According to Maass (2021), a study of 63 California law enforcement agencies provided 2018–2019 data that "collected a combined average 840,000,000 plate images each year" (Maass, 2021, "Hit Ratio" section). However, "only 0.05% of the data collected by ALPRs was relevant to a public safety interest at the time the data was captured" (EFF, 2023, "How Law Enforcement Uses ALPR Technology" section).

Finally, there are serious cybersecurity concerns. For example, according to a DHS (2020) inspector general report, Perceptics (an ALPR vendor for the U.S. Customs and Border Protection) was hacked sometime before May 2019 and its data was published online. As the DHS report noted,

> The attack compromised thousands of driver and passenger images that CBP captured during the VFS [Vehicle Face System] pilot. CBP determined that more than 184,000 traveler facial image files, as well as 105,000 license plate images from prior pilot work, were stored on the subcontractor's network at the time of the ransomware attack. In addition, the hacker stole an array of contractual documents, program management documents, emails, system configurations, schematics, and implementation documentation related to CBP license plate reader programs. (pp. 7–8)

The Minneapolis Police Department (MPD) provides another unfortunate example of poor data and cyber hygiene. According to Davidson (2024),

> while users could only access the system from computers directly connected to the city's network, outdated accounts associated with individuals no longer affiliated with MPD or with those who lacked a current ALPR access requirement were still operational. According to the report, that presented a latent risk of an insider exploiting older, still-active accounts. ("Minneapolis Audit Reveals Security Lapses" section)

There are several things Texas lawmakers can do to address the concerns posed by ALPRs. First, ensure that ALPR companies that operate in Texas are captured by data broker registry laws. Second, require strong data privacy and cybersecurity hygiene protections on procurement and vendor contracts. Third, consider limits on private sector data exchanges to or from law enforcement to counter the cat's paw and end-around the Fourth Amendment. Fourth, ban data transfers to entities located, operated, or owned by hostile foreign nations. Fifth, implement data minimization procedures including "sensible retention limits,[11] specific policies about who inside an agency is allowed to access data, and audit and control processes could help minimize these issues" (EFF, 2023, "Threats Posed by ALPR" section). The MPD example above is an example of the vulnerability of valid accounts attacks. Sixth, the Legislature should require agencies, law enforcement, and political subdivisions that have ALPR access to require two-factor authentication, change default passwords, and immediately delete accounts that should no longer have access (CISA & USCG, 2023).

The Judicial System and Courts

Furthermore, AI has many potential applications in the judicial system and courts, including operational and administrative functions, increasing access to courts and resources, case management, legal research and writing, discovery and evidentiary analysis, assessing liability, determining remedies, assessing potential clients, witnesses, and jurors, and more (Firth-Butterfield & Silverman, 2022). As above, where constitutional rights and civil liberties are involved, care must be taken in crafting policies regarding allowable and unallowable uses of AI in the judicial system and courts.

First, the use of advanced technologies in transcription and translation services provides prospective benefits to courts, litigants, witnesses, victims, and more. AI should not replace the human element, but with a shortage of qualified court reporters and translators (including those with American Sign Language capabilities), technology can help augment human efforts (Witherspoon, 2024; Bates, 2020). For example, since many judicial and court functions transitioned online during the COVID-19 pandemic, Texas courts turned to over-the-phone, video, and team interpreting tools for those who needed them (Dejeux, 2022). AI is an emerging translation tool for written and verbal communication, which can be used around the clock, can translate instantaneously, and can work more cost-effectively than human translators (Cannestra, 2024; Stoyanov, 2024). However, this technology is still developing, and there are concerns about accuracy, bias, and non-verbal communication. For example, most languages, regional dialects, and cultural nuances are not available or always captured accurately by AI translators (Bhuiyan, 2023).

Second, AI can be useful in augmenting administrative and operational functions at law firms and the courts, including case management, docket management, scheduling, managing jury pools, finance, fee and fine collections, and more (Firth-Butterfield & Silverman, 2022). For example, many of these functions can be automated, greatly reducing labor hours spent on administrative tasks and freeing up time and energy to increase customer-facing services. This is particularly important considering a 2024 survey citing concerns over court backlogs and workforce shortages:

> In fact, these concerns over technology also become more dire when you consider that more than half of the respondents said they expect staffing shortages within the coming 12 months. That expectation will require staff to rely more heavily on technology to keep the court moving and provide access to citizens in the long run. (Thomson Reuters, 2024, para. 8)

[11] Retention limits are an incredibly important check on data abuse. For example, bills like HB 1181 (2023) only allow identifying data to be used for verification purposes, after which they must be deleted. In the context of ALPRs, other states have placed limits on data retention. Texas could do the same:

> While many agencies do store data for one year or more, there is no industry standard for ALPR. For example, Flock Safety, a vendor that provides ALPR to many California agencies, deletes data after 30 days. The California Highway Patrol is only allowed to hang onto data for 60 days. According to the National Conferences of State Legislatures, Maine has a 21 day retention period and Arkansas has a 150 day retention period. In New Hampshire, the law requires deletion after three minutes if the data is not connected to a crime. (Maass, 2021, "Hit Ratio" section)

Or, as EFF (2023) suggested, "one of the better privacy protections would be for police to retain no information at all when a passing vehicle does not match a hot list" ("Threats Posed by ALPR" section).

Third, chatbots and other AI technologies can help increase community awareness and access to the courts, provide round-the-clock communications, answer frequently asked questions, provide service referrals, assist in corrections and community supervision, facilitate victims' services and support, and the like (Camello et al., 2021; Justice Innovation, n.d.).

Finally, AI can help individuals gain access to the courts and the judicial process. It may be able to help litigants who are unsure of where to start or how to proceed, those that are self-represented, and more. For example, researchers at the Stanford Legal Design Lab are working to design simple, plain-language systems to enhance access to the justice system, including with:

- Guides to answer basic questions;
- Guidance for timelines, deadlines, and how a case moves through the system;
- Assistance with filling out legal paperwork and court forms;
- Checking documents for completeness and accuracy to avoid procedural errors or dismissals;
- Service referrals;
- Translation services;
- Legal research and case summaries .

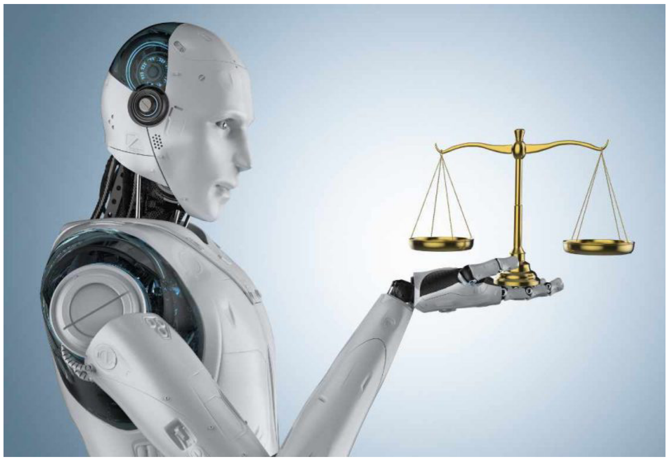
Furthermore, as the justice system seeks to encourage alternative dispute resolution, AI can be used to assist in mediation, negotiations, and settlements (Justice Innovation, n.d.).

Of note, researchers are careful to distinguish between AI's potential uses in criminal versus civil cases, with substantially more ink written on its uses and limits in criminal cases (Wu, 2019). Lawmakers should also be careful to make this distinction, with many constitutional and substantive rights at stake in criminal cases.

Alongside these potentially positive use cases, however, there are some AI use cases that should give policymakers, judges, court administrators, and others pause and extra scrutiny. Among them are use by lawyers and judges; drafting materials, briefs, and motions; evidentiary admissibility and analysis; AI-created exhibits or evidence; determining liability and standards of proof; forensic psychology, jury and witness analysis, and jury selection; and the decision of when, if at all, to disclose the use of AI. The response of the courts has been mixed.

As a baseline, Chief Justice John Roberts (2023) has argued that "any use of AI requires caution and humility" (p. 5). The Chief Justice continued:

> One of AI's prominent applications made headlines this year for a shortcoming known as "hallucination," which caused the lawyers using the application to submit briefs with citations to non-existent cases. (Always a bad idea.) Some legal scholars have raised concerns about whether entering confidential information into an AI tool might compromise later attempts to invoke legal privileges. In criminal cases, the use of AI in assessing flight risk, recidivism, and other largely discretionary decisions that involve predictions has generated concerns about due process, reliability, and potential bias. (pp. 5–6)



Note. AI has many potential applications in the judicial system and courts, including operational and administrative functions, increasing access to courts and resources, case management, legal research and writing, discovery and evidentiary analysis, assessing liability, determining remedies, assessing potential clients, witnesses, and jurors, and more. Chief Justice John Roberts (2023) argued that the use of AI in the judicial system "requires caution and humility" (p. 5). Photo from *Is It Good for Lawyers to Use AI for Legal Briefs?* by R. Kaur, CaseFox, n.d. (https://www.casefox.com/blog/ai-legal-brief/)

Some courts emphasized transparency and accountability:

> Several courts, including those in Texas, Illinois, and Manitoba, Canada, have issued rulings or standing orders on the use of gen AI in their courtrooms. Each court's ruling places the onus on the attorney to notify the court of their use of gen AI in detail. The courts also require that the attorney review and confirm the accuracy of the work done by gen AI. (Butler, 2023, "Courts' choices" section)

In what has been called a possible "first-of-its-kind" ruling in a triple murder case, a Washington state judge recently ruled that an AI enhanced cell phone video was inadmissible (Stelloh, 2024, para. 1). As reported by NBC News, the judge

> described the technology as novel and said it relies on "opaque methods to represent what the AI model 'thinks' should be shown." "This Court finds that admission of this AI-enhanced evidence would lead to a confusion of the issues and a muddling of eyewitness testimony, and could lead to a time-consuming trial within a trial about the non-peer-reviewable-process used by the AI model," the judge wrote in the ruling that was posted to the docket Monday. (paras. 2–3)

As a final example, Rao & Ramstad (2023) documented nearly a dozen cases where fake legal citations were submitted to courts, including some generated by AI. The most prominent is *Mata v. Avianca, Inc.*, in which a New York state judge sanctioned the plaintiff's lawyer, who had three decades of experience, for submitting a response brief with fake case citations (Ryan et al., 2023; Weiser, 2023).

Furthermore, litigants, lawyers, and judges are grappling with how to use AI during discovery, for evidentiary analysis, and for determining liability, standards of proof, predictions on rulings on motions, and history of settlement agreements. For example, AI may be a helpful tool to expedite discovery, especially when litigants do last minute document dumps to obscure evidence (Redden et al., 2020b). Furthermore, AI's ability to comb through massive amounts of data may help litigants look at the history of settlements in similar cases to determine the proper offer, counteroffer, and likelihood of settlement (Justice Innovation, n.d.). However, as an animating principle, these tools should only augment and not replace the human touch. The jury's role as fact finder (or judge in a bench trial) should not be replaced by a machine, particularly in deciding, however imperfectly, whether a party has met its burden of proof and how to apportion liability, for example.

Finally, as in healthcare, the extent to which AI is used raises questions of legal ethics and where to draw the line on the unauthorized practice of law. Lawmakers should review Chapter 81 of the Texas Government Code, which defines the practice of law as well as certain exceptions for legal assistance and accounts for the need to explicitly include (or exclude) certain AI uses and functions.

Ultimately, in areas where privacy, civil liberties, and constitutional rights are at stake, lawmakers should give extra scrutiny to crafting policies related to AI use in law enforcement, public safety, the judicial system, and courts. Trust in the broader justice system is paramount, particularly in the criminal context, where due process, jury rights, liberty, and even life itself are at stake. As U.S. Supreme Court Chief Justice John Roberts (2023) has written, "at least at present, studies show a persistent public perception of a 'human-AI fairness gap,' reflecting the view that human adjudications, for all of their flaws, are fairer than whatever the machine spits out" (p. 6).

That is not to say there is no role for AI in the judicial process. While AI should never be used to replace human judgment and interpersonal interactions, there are some useful, value-additive, and force multiplying functions it can offer in these contexts.

Employment

AI presents tremendous promise for the workforce, businesses, and the economy. This section will not, however, focus on the broader impact AI will have on specific sectors of the economy, what industries or jobs are at greatest risk of disruption from AI, or how to reskill and upskill the workforce. Rather, this section will consider how AI can be used in the employment context. In this space, AI can be used for resume drafting, standardization, and review; targeted advertising for job candidates and employers; hiring, performance reviews, promotion, and firing decisions; and salary determinations, comparisons, and negotiations, among other things. A key area of focus is the use of automated employment decision tools (AEDTs), which will be discussed in more detail later in the paper. More broadly, however, AI is a particularly valuable tool in remote and hybrid work environments, helping connect talent and employers from around the corner, across the country, or around the world. For example, according to the Future of Privacy Forum (FPF) (2023),

> the use of automated technology in the workplace can result in faster hiring for employers, increased access to diverse candidates and a broader pool of applicants, and greater access to hiring tools for small to mid-sized businesses. For candidates, automated technology can help match their skills to a broader variety of roles and identify new potential career paths. (p. 2)

FPF's "Best Practices for AI and Workplace Assessment Technologies" paper emphasizes the principles of "(1) non-discrimination, (2) responsible AI governance, (3) transparency, (4) privacy and data security, (5) human oversight, and (6) alternative review procedures" (p. 2). This is a thoughtful framework for industry best practices and policymaking that could be applied not only to employment but to nearly all AI use cases.

However, AI in employment raises several concerns related to data privacy and collection practices, particularly sensitive data, cybersecurity, bias, discrimination, and more. The data privacy and cybersecurity concerns have been covered at length in this paper and more broadly in previous research from the Texas Public Policy Foundation (TPPF) (Dunmoyer & Whiting, 2022). For this paper, the authors emphasize the potential for bias and discrimination. At the outset, it is important to recognize that AI use and regulation must comply with existing civil rights and employment laws, including Title VII of the Civil Rights Act of 1964, the Americans With Disabilities Act, the Age Discrimination in Employment Act, and similar state and local laws (Schwartz et al., 2023).

Furthermore, many of the same concerns discussed elsewhere about bias in input data, algorithms, and output data apply in the employment context as well. This includes data on protected classes such as race, color, religion, sex, national origin, age, disability, and genetic information (U.S. Equal Employment Opportunity Commission, n.d.). It further captures data like educational attainment, language skills, and criminal history. The output data—for example, recommending or rejecting a candidate—are determined by the type, integrity, transparency, and accountability of the algorithm.

Identifying discrimination is hard enough in employment law, let alone with the added layer of non-human review and decision-making . For example, consider a 54-year-old female applicant who is a non-native English speaker, is from Mongolia, has lived in the United States for 40 years, and had a minor criminal conviction at 19. If an AI resume reviewing tool reads her resume and rejects her, how does one know why? If it was because of her age, sex, or national origin, that would violate federal law. But what if it was based on her English abilities or her criminal conviction? Or what if it was a legitimate mismatch of the skills that she offers with those the employer desires? The black box problem discussed above makes it hard to know.[12]

Ultimately, lawmakers should give extra attention to the use of AI in hiring, promotion, firing, and salary determinations. As in other use cases, sensitivity to the human element must be paramount. Hiring platform Indeed's (n.d.) AI principles rightly emphasize the human element in hiring: "Decisions about jobs and hiring are among the most important in peoples' lives, and we want to preserve the human element in those decisions while continuing to improve and simplify the experience of finding the right job or candidate" (para. 5). This approach seeks an appropriate balance between innovation and human agency, autonomy, and dignity, which is a common thread in this paper and the Foundation's broader technology policy research.

Agricultural Applications

The agricultural industry is ripe for disruption from the implementation of artificial intelligence. As employment shortages continue to threaten the productivity of farms in the United States, startups are innovating to provide technological solutions to the challenges farmers face. For example, AI is equipping farmers with new tools to monitor the status of their crops**,** the health of their cattle, to kill weeds and pests that damage crop growth, optimize irrigation systems, and maximize the health and use of their soil (Choi et al., 2023). The application of AI in agriculture does not present a complicated moral and philosophical quandary as in other industries, although there are challenges to consider. Furthermore, the notion of labor replacement is less of a threat (as most technologies are labor augmenting), thus improving productivity, as the work done by AI technologies in agriculture is most often monitored by humans.

Pest and Weed Control

Globally, pest and weed control costs farmers an average of $60 billion annually (Mount, 2022). As Mount noted, much of this is wasted, and farmers are eager to find ways to reduce costs and increase efficiency. For example, one solution comes from Carbon Robotics (n.d.), which created a device known as LaserWeeder. This product, a tractor implement, uses deep learning and computer vision to distinguish weeds from more than 100 crop varieties using 42 cameras on the device, precisely targeting weeds with a high-powered laser. In addition to reducing costs, it is also attractive to organic farmers who are interested in reducing their use of herbicides.

Note. AI is being used in agriculture for planting, harvesting, crop rotation, irrigation, soil monitoring, pest and weed management, pesticide application, livestock monitoring, and more. Photo from *AI in Agriculture — The Future of Farming*, by A. Piddubna, Intellias, August 12, 2024 (https://intellias.com/artificial-intelligence-in-agriculture/).

Livestock

An additional application for artificial intelligence is in livestock monitoring, which enables supervision of the productivity and health of a given herd. For example, using video feeds from within cattle pens, farmers are now able to utilize machine learning to track and identify the cattle's frequency of feeding (Research@Texas A&M, 2024). This information then allows them to observe if any cattle are suffering from a disease of some kind, since decreased feeding frequency may indicate poor health. Early warning through video monitoring equips farmers with the ability to maintain the health of their livestock, deter predators, and improves profit margins.

[12] Recall that, according to Blouin (2023),

> this inability for us to see how deep learning systems make their decisions is known as the "black box problem," and it's a big deal for a couple of different reasons. First, this quality makes it difficult to fix deep learning systems when they produce unwanted outcomes. (para. 3)

Crop and Soil Health Monitoring

A critical aspect of farming is soil health. Proper nutrition and mineral content are required for efficient crop growth, and some farmers spend many hours extracting soil samples for testing. A process which would ordinarily take weeks to complete can now be accomplished with one tool in a matter of hours. For example, a company called ChrysaLabs (n.d.) has designed an all-in-one soil probe that provides real-time assessment of any soil into which it is inserted. It utilizes machine learning to assess nutrients such as nitrogen, phosphorus, potassium, calcium, and magnesium, while also assessing soil pH and moisture. These readings are immediately accessible to farmers on their mobile devices via an app and provide analytics that equip them to make accurate decisions regarding soil preparation, fertilization, irrigation, and more.

Financial Services

The applications of AI in finance are vast and include risk assessment, risk management, predictive modeling, credit and insurance underwriting decisions, financial advisory services, trading, financial management, personalized banking, customer recruitment and onboarding, customer service, process automation, document processing, cost savings anaylsis, cybersecurity, speech and image recognition, chat systems, sentiment analysis, translation services, and anomaly and fraud detection, management, and prevention (Tierno, 2024; Finio & Downie, 2023; University of San Diego, n.d.; Google Cloud, n.d.-b). However, while the financial industry has used forms of artificial intelligence for decades, its mass adoption and customer-facing use has been more deliberate than other industries. As Eastwood (2024) noted, "The financial industry has been slow to put AI in front of customers. … Customer-facing applications require much more effort than internal use cases, in part because they're subject to far more regulatory scrutiny" ("Regulatory concerns stall some uses" section).

In 2019, "banks with more than $100 billion in assets … [were] 75% more likely to already be using AI, compared to 34% for those with less than $100 billion in assets" (Noelle, 2019, para. 3). This is also validated by Columbus (2020), who noted that, in the early stages of the COVID-19 pandemic, "54% of Financial Services organizations with 5,000+ employees ha[d] adopted AI" (para. 5). However, its use has expanded since then, and by today, industry executives are surprised by "the rapid pace of adoption in the financial sector" (Eastwood, 2024, para. 2). Furthermore, the use of AI technologies is expanding beyond the largest banks to financial institutions of all sizes, investment firms, insurance, and more—to what Citi (2024) terms leaders, fast followers, laggards, and dark horses and what Deloitte terms frontrunners, followers, and starters (Gokhale et al., 2019). As Spencer Reich, CEO of a financial advising and consulting firm, noted,

> Nearly four years ago when I talked to portfolio managers, there was a general aversion/fear to the use of AI. However, this has changed dramatically in recent years with the advent of ChatGPT. Today, firms have a fear of missing out (FOMO), and worry they will fall behind if they do not embrace AI. (Ghose et al., 2024, p. 60)

It is understandable why. According to Citi (2024), finance is "a data rich industry with clients adopting AI at pace" and AI "will profoundly change the future of finance and money … potentially driv[ing] global banking industry profits to $2 trillion by 2028, a 9% increase over the next five years" (paras. 1–2). Potential cost savings could also save banks and other financial institutions tens of billions of dollars annually (Ghose et al., 2024; University of San Diego, n.d.). McKinsey & Company "estimates that across the global banking sector, gen AI could add between $200 billion and $340 billion in value annually, or 2.8 to 4.7 percent of total industry revenues, largely through increased productivity" (Buehler et al., 2024, para. 2). This is echoed in a Citi (2024) report in which "93% of respondents to a proprietary survey expect higher bank profits on the back of productivity gains" (para. 8).

Consider several examples of how AI is used in the financial sector. First, AI is being used to augment mid- and back-office tasks, data aggregation and visualization, and fraud detection (Eastwood, 2024).

According to Columbus (2020), "70% of all financial services firms are using machine learning to predict cash flow events, fine-tune credit scores and detect fraud" (para. 4). Furthermore, JPMorgan Chase "uses key fraud detecting applications, including implementing an algorithm to detect fraud patterns. … Details of credit card transactions are sent to data centers, which decide whether the transactions are fraudulent" (University of San Diego, n.d., "Examples" section).

Second, AI is being used in loan and credit decision-making. However, as Kremer et al. (2024) noted, "gen AI has arrived in the credit risk world but has yet to transform it" ("The current state of gen AI in credit risk" section). For example, credit reporting agency Experian argued that traditional lending models are limited by slow reaction times, fewer data sources, and less effective performance than recent "advances in analytics and modeling [that] are making credit risk decisioning more efficient and precise" (Lee, 2024, para. 1). Kremer et al. (2024) noted that AI tools can "autonomously follow task sequences" to receive and review applications, recognize missing data, draft communications to customers to clarify or correct the data, flag policy violations, and even draft necessary documents and contracts, all before a human credit officer reviews a thing ("Use cases in credit risk" section). Furthermore, "AI can analyze a variety of data, including social media activity and other online behavior, to assess customers' creditworthiness and make more accurate credit decisions" (Finio & Downie, 2023, "How is AI used in finance" section). Theoretically, this allows otherwise "unscorable" loan consumers more options for traditional lending services.

Finally, AI is being used to enhance customer service and banking options. This is also market-driven, in part by the COVID-19 pandemic and shifting consumer preferences, with nearly 80% percent of millennials preferring to avoid a physical bank branch location, for example (EMARKETER, 2023). According to Deloitte, "60% of banks have closed or shortened opening hours of branches but many have also implemented new digital features, such as fully digital processes, e.g. account opening (34%), remote identification & verification (23%) and contactless payments (18%)" (Trapassi et al., 2021, p. 2). Furthermore, automated telephone banking has given way to chat- and app-based bots. Capital One created "Eno," the first SMS-based, natural language chatbot (University of San Diego, n.d., "Examples" section). Bank of America's chatbot, "Erica," is capable of understanding nearly 500,000 prompts and has had more than 1.5 billion interactions with more than 37 million clients since its creation in 2018. In 2022, clients spent "more than 3 million hours interacting with Erica," and in the first six months of 2023, Erica was used 333 million times (Bank of America, 2023, para. 3). The remarkable shift from the days of George Bailey's community bank has even given rise to what are called neobanks, or fully digital financial firms, operating online or through apps with no physical branch locations (Sorensen, 2019).

However, because of the nature and importance of finance in daily life, industry and lawmakers must carefully consider the data privacy implications, the risk of bias and discrimination, and the need for transparency and accountability. With a 2024 NVIDIA survey revealing that more than nine in 10 financial services companies are either assessing AI or leveraging it in production, consumers have become aware of this and not all are embracing it with optimism (Levitt, 2024). On the one hand, consumers are taking to using GAI and consumer AI tools for their personal financial planning. Specifically, a 2024 Ipsos poll found that 37% of Americans are using AI to manage their finances (Cottrill, 2024). Conversely, that same poll found that 64% of Americans do not trust AI to take the emotions and overall interests of investors into account—suggesting that the emotions that drive the risk profile of investors constitute an important component that cannot be replaced by mere arithmetic. On balance, Americans see AI as a helpful tool to organize and contextualize financial goals but not to automatically make investment decisions for them.

Even without the potential concerns that might transpire from the unfettered application of AI in the financial sector, there are already existing threats to the financial autonomy of Texans and Americans.

These concerns have become manifest in America in recent years, as unpacked in commentary by one of the paper's authors:

> You likely recall the "Freedom Convoy" story that gripped the world in January of 2022. Canadian truckers stood in solidarity in opposition to draconian social restrictions and vaccine mandates ostensibly instituted to stop the spread of COVID-19. This peaceful expression set off the class of Canadian elites, who determined they would pay by financially crippling both the protestors and any individual who donated money to the Freedom Convoy. Their bank accounts were frozen, leaving many unable to purchase necessities like groceries and medication, or provide care for their children. The message was clear: stand in opposition to the government, and we will rob you of all agency.
>
> While that message was intended for the freedom-loving Canadian protestors, U.S. Federal Reserve Chair Jerome Powell heard, too, and his ears perked up. (Dunmoyer, 2023, paras. 3–5)

In a report published the same month that Canada's Finance Department shuttered private citizens' bank accounts, the Federal Reserve stated it is "considering how a CBDC might fit into the U.S. money and payments landscape." In the words of the Federal Reserve, a CBDC, or Central Bank Digital Currency, is "a digital form of central bank money that is widely available to the general public. … [It] refers to money that is a liability of the central bank." Meaning, you would bank with the federal government rather than private banks, and they would be able to track your purchasing patterns and financial activity with a degree of precision never before possible.

All these financial surveillance tools are made possible by increased digitalization in the financial services sector, as well as enhanced abilities to track and glean insights from consumer data. Specifically, as a Congressional Research Service report put it, "While in the past, personal banking was rooted in community relationships and conducted in person, now nearly all transactions can be performed online" (Tierno, 2024, p. 4).

Given that AI tools allow organizations to better intake, organize, and sort customer data based on insights of importance, it follows that, left unchecked, AI could metastasize existing concerns to financial autonomy and discrimination on grounds of politics, religion, or immutable characteristics. The black box dilemma introduces concerns of trust and transparency, since financial institutions might employ AI tools that produce outcomes that could appear discriminatory, biased, or unexplainably malicious. Even if negative outcomes occur without malicious intent on the part of the developers and deployers of AI systems, without explainable AI, financial institutions and consumers alike will have zero trust in such systems and a lower inclination toward adoption.

Like in other high-risk applications of AI, the financial sector is an area where it is crucial that humans remain in-the-loop lest imperfect systems without the sensibilities of human consciousness make critical decisions that pose negative financial—and overall privacy and civil liberty—consequences for Texans.[13]

Creative Uses

The development and training of LLMs and AI technologies broadly have introduced pronounced tensions within the creative community. Transparency tools such as "Books3"—which reveals the archive of books being used to train AI systems without the authors' consent—highlights just one facet of the creative community that feels slighted by large technology corporations using their works for profit without consent or compensation (Reisner, 2023). At a high level, the mushrooming of GAI models has introduced a tension between the creative rights and the ever-growing data needs

[13] Broader concerns about fraud, abuse, bias, discrimination, and the safety of customer information in AI use have been discussed at length in other sections.

of AI systems. This can be thought of as a species-level competitive environment. Creators capable of organic, original expression are pitted against AI-content made by models trained on their work. Given contemporary AI tools encompass nearly all forms of digital content—written text, spoken word, music, still images, video, and more—there are wide-ranging implications.

One of the more difficult concerns with AI and creatives is that of "ownership" of oneself—specifically, their name, image, and likeness (NIL). Actress Scarlett Johansson has encountered this on multiple fronts. Initially, Lisa AI ran an advertisement that displayed a lifelike video of Johansson endorsing the image-generating app, with a voice that was an uncanny representation of her own (Shanfeld, 2023). While there was a fine print disclaimer in the video that the ad was produced by Lisa AI, Johansson took legal action on grounds that her name and likeness were used for pecuniary gain without her consent. Johansson is also embroiled in a legal dispute with OpenAI after the company released a chatbot named "Sky" with a voice that sounds eerily similar to that of the actress. Despite the fact that Johansson declined Open AI after they approached her to see if she would be interested in being one of the chatbot voices, they remain resolute in their claim that, though the voice sounds similar, it was trained on the voice of a different professional actress (Georgetown University, 2024). And while the former instance is rather black and white in terms of the use of Johansson's likeness, the latter is a matter of Johansson's right of publicity and whether it was violated by OpenAI's use of a voice that sounds like hers—a much more onerous burden of proof. And the concerns of NIL ownership transcend mortality—actors such as James Dean, Carrie Fisher, Harold Ramis, and Paul Walker have posthumously starred in movies thanks to the power of contemporary AI tools (Velasquez, 2023).

Certainly, not all of these applications are negative. When the consent of the creative—or of the manager of an actor's estate—is given, a new slew of possibilities is more ethically introduced to enhance creativity, entertainment, and advancement of the arts broadly.

Note. AI presents many new and creative uses cases in art, film, photography, writing, voice work, and more. However, it also presents numerous intellectual property challenges. For example, Scarlett Johannson accused Open AI of using her voice to train a chatbot without her permission. Open AI has denied this accusation while also quietly removing the offending chatbot voice. Photo from *Scarlett Johansson Says OpenAI Used Her Voice Without Permission* [Video], by Today, May 21, 2024 (https://www.today.com/video/scarlett-johansson-says-openai-used-her-voice-without-permission-211299909920).

In addition to its effect on Hollywood talent, AI has implications for the writers behind many of the TV shows and movies consumed across the globe. From May to September 2023, the 11,500 screenwriters comprising the Writers Guild of America (WGA) went on strike due to disputes with the Alliance of Motion Picture and Television Producers (AMPTP) (Faguy & Ray, 2023). Among other contentions, writers expressed concern that the use of AI in screenwriting poses a threat to their job security and the integrity of their craft. Through a series of deliberations, the WGA agreed on terms that prevent companies from mandating the use of AI tools, prohibit the use of AI for writing or rewriting scripts, gives the WGA the "right to assert that exploitation of writers' material to train AI is prohibited by [the contract] or other law," and clarifies that members do not consent to their work being used for purposes of training AI systems (Silberling, 2023, para. 7).

In April 2024, the Artist Rights Alliance (ARA) partnered with more than 200 artists and songwriters—with big names ranging from the Jonas Brothers, Aerosmith, Billie Eilish, Mumford & Sons, and even estates such as that of Bob Marley—to draft an open letter on AI (Aswad, 2024). Coalescing over concerns that the development of AI poses to their creative domains, the artists called on "AI developers, technology companies, platforms and digital music services to pledge that they will not develop or deploy AI music-generation technology, content or tools that undermine or replace the human artistry of songwriters and artists or deny us fair compensation for our work" (ARA, 2024, para. 7). The artists collectively cede that, when used responsibly, AI can elevate human creativity beyond the status quo. But they underscore a major problem: Some of the largest tech companies developing AI tools are training their models on artists' works without consent, potentially setting "in motion a race to the bottom that will degrade the value of our work and prevent us from being fairly compensated" (para. 5).

The tensions between artists and AI companies have only grown. In June 2024, Universal Music Group, Sony Music Entertainment, Warner Music Group, and other large record labels sued AI-song generation programs Suno and Uncharted Labs on grounds that the companies are stealing the works of artists for pecuniary gain without consent or compensation (Yang, 2024). The use of such tools has even sparked controversy between artists. For example, Drake took down a diss track, which allegedly used an AI tool to emulate the voice of the late rapper Tupac Shakur, after Shakur's estate threatened to sue (Yang & Hamedy, 2024). Of note, Tennessee became the first state to afford NIL protections for artists against the misuse of AI, with Gov. Bill Lee signing the ELVIS Act into law (Gibbs, 2024). At the bill signing ceremony, Gov. Lee stated the intent was to prevent the destruction of the music industry and protect artists from being robbed of their talents by AI systems that impersonate their talent and disregard their intellectual property rights.

Notwithstanding the aforementioned challenges, the broader theme throughout this paper of AI as a tool equally applies to the creative landscape. As noted by Bieser (2022), "experts agree that AI will not develop fundamentally new ideas on its own; however, there are ways in which AI can support humans in doing so, as AI can augment human creativity" (p. 4). Some examples of AI as a tool of augmentation for the development of creative content include:

- *Inspiration and Idea Generation*: AI can analyze vast amounts of data and generate unique concepts that might inspire an artist's next masterpiece. It can create unexpected connections between diverse concepts, fueling the artist's creative spark.

- *Visual Exploration*: For visual artists, AI-generated images can serve as starting points for new creations. AI algorithms can create abstract patterns, morph images, or generate unique compositions that artists can incorporate into their work.

- *Music Composition*: Musicians can harness AI to compose melodies, harmonies, and even entire pieces of music. AI can analyze existing compositions and generate original musical ideas, freeing artists to experiment with new genres and styles.

- *Textual Creativity*: Writers and poets can benefit from AI-generated text prompts, which can kickstart the writing process. AI can generate sentences, ideas, or even entire paragraphs that serve as springboards for crafting engaging narratives. (Fortino, 2023, "AI's Role in the Creative Process" section)

As with other creative pursuits, there are mixed feelings about AI among authors. Indeed, the majority of authors (87%) do not use artificial intelligence as part of their writing process (The Authors Guild, 2023). Of those who are beginning to employ it, 33% use it for brainstorming, 26% for assisting with marketing and promotion, and 13% for draft organization and structuring. Perhaps unsurprisingly, only 3% of authors believe it is okay for AI systems to be trained on their books without consent or compensation. However, in the face of a changing industry, more authors such as non-fiction writer Chris Anderson are beginning to experiment with different AI tools to aid in the writing process (Veltman, 2024). Anderson fed parts of one of his novels into an AI writing platform so that it could study his voice and assist in the novel writing process. Ultimately, Anderson never used a single word generated by the AI platform, as it did not capture his voice as accurately as he would have liked, but the tool was able to help in generating new ideas for scene locations, for example.

The above applications of AI in the creative realms are concerning and promising and beget philosophical questions surrounding art and creativity broadly. Legendary music producer Rick Rubin contextualizes this challenge by saying, "What I find interesting about art is the point of view of the person making it. And I don't know that AI has a point of view of its own" (Huberman, 2024, 25:23). Artists pushing the creative envelope tend to think of this personal process through the lens of a muse (as in the goddesses of literature, science, and the arts in Greek mythology). While modern guitarists will surely listen to classic greats ranging from Jimi Hendrix to Jimmy Page to Eric Clapton to find singular moments or systems of expression that transfix them with inspiration, the muse "guides" them to employ the creative faculties of mind reserved for humans to study past methods, blend it with their unique artistic flair, and generate something new through human-driven, creative expression.

This same schema can be applied across the creative spectrum, with painters, actors, writers, and so on. In an oversimplified theoretical framework, this means artists need not feel threatened by AI tools given they can merely rearrange and analyze existing information rather than generate new, original thoughts and expressions. In the marketplace of ideas, what is considered truly creative and moving creative content will win out amidst the oversaturation of AI replicas. However, as evinced by artists who are frustrated by the tsunami of AI-generated content, some artists may opt out of the market in its entirety if the integrity of the creative realm is constantly under assault by AI models ripping off their works for profit without consent or transparency. This would cripple the rate at which new ideas are generated and imperil the incentive structure that supports authentic human creation. Ultimately, this highlights the need for lawmakers to consider the conflicting intellectual property issues and seek an appropriate balance.

REGULATORY APPROACHES

This section surveys the AI regulatory landscape. Initially, the authors will examine efforts by industry to develop responsible technologies and self-police. Many AI companies claim to be responsible, good faith actors—and many are. Furthermore, because AI is a tool that can be used for good or for ill, there may be different considerations for developers, deployers, and end users. Accordingly, identifying industry best practices—or even an agreed upon regulatory framework, for which some stakeholders are asking—is a good first step on the front end.

Beyond industry efforts, this section will review the broader regulatory landscape, including efforts at the local, state, federal, and international levels. Uniformity is usually desired and often ideal compared to a patchwork approach. However, Congress is prone to talk without meaningful action—which, indeed, is what is happening with AI and other emerging technologies—and international efforts are either disjointed or overly burdensome. Accordingly, consistent with other areas of technology policy, the Foundation advocates for a state-based regulatory approach that preempts local regulation.

Industry Best Practices

AI companies are quick to claim awareness about the need for guardrails related to the expansion and use of their products. These claims, taken at face value, are a positive indication that companies like Google and OpenAI are serious about ensuring their products are not utilized for nefarious purposes. However, there are myriad legislative case studies on important tech policy issues—from kids' online safety legislation to data privacy laws—which demonstrate these companies often wield such laudable statements as public relations tools but actively thwart legislative efforts over fears that their bottom line would suffer if such laws were enacted. Moreover, it is now well established that some tech luminaries in Silicon Valley, namely Google's Larry Page, see humanity as a threat to innovation and go as far as to chastise actors that want guardrails protecting humanity from the excesses and potential harms of AI (Shaw, 2023; Isaacson, 2023, p. 241).

Initially, Microsoft, OpenAI, Meta, and Google all list internal policy guidelines that focus, generally, around six core areas: fairness, reliability and safety, privacy and security, inclusiveness, transparency, and accountability. These guidelines embody the framework through which each company views the development and deployment of their AI systems, and they claim to enforce these standards on themselves through internal committees. Interestingly, a quick examination of these companies' guidelines suggests potential conflicts. For example, OpenAI on the same page highlights its push for "safe, beneficial AI systems," followed almost immediately by a goal to be "intense and scrappy" to build systems with great "urgency" (OpenAI, n.d.-b, Title, para. 3). Or take Google, which states a top objective for AI applications is to "avoid creating or reinforcing unfair bias," yet experienced severe public backlash when its AGI system, Gemini, created false, biased, and misleading racial depictions of Founding Fathers, Civil War veterans, and so on (Google AI, n.d., "Objectives for AI applications" section). As evinced by this pattern, it is perhaps unsurprising that more than 80% of Americans do not trust tech executives to regulate AI (Heath, 2023). Ultimately, while disagreements on the legislative approach and substance may persist, Americans are united in believing that legislative oversight is a much-preferred alternative to Big Tech self-policing.

Rather than further itemize individual tech companies' stances on AI best practices, industry trade associations and coalition groups provide a window into where critical mass has aligned. TechNet (2024), a national network of technology executives representing some of the biggest private-sector players in AI, developed its own set of best practices and policy recommendations. Member companies include Apple, Anthropic, Google, Indeed, Meta, OpenAI, Snap Inc., and many more. TechNet's best practices and policy recommendations can be summarized as follows:

Note. Companies like Microsoft, OpenAI, Meta, and Google list internal policy guidelines that focus, generally, around six core areas: fairness, reliability and safety, privacy and security, inclusiveness, transparency, and accountability. Pictured here are NVIDIA CEO Jensen Huang (L), Google CEO Sundar Pichai (C), and Meta CEO Mark Zuckerberg (R) at a Capitol Hill AI briefing. Photo from *Bill Gates, Elon Musk and Mark Zuckerberg Meeting in Washington to Discuss Future AI Regulations*, by B. Fung, CNN, September 13, 2023 (https://www.cnn.com/2023/09/13/tech/schumer-tech-companies-ai-regulations/index.html).

- **Leverage existing laws and adopt a risk-based approach for effective AI regulation.** Ensure AI laws are not redundant where existing law can be applied; address AI risks logically and incrementally; push for a central coordinator for AI oversight at the federal government level; promote NIST's AI Risk Management Framework as a voluntary model; and more.

- **Responsible AI evaluations.** When crafting policy to promote responsible AI—through explainability, transparency, audits, etc.—balance the tensions between trade secrets, technical feasibility, data privacy and security, and overall system safety.

- **Mitigate potential bias.** AI development must reflect society's highest ideals; it should be incumbent on developers, deployers, and users of AI to implement appropriate oversight to maximize good and minimize harm; and continue to apply existing anti-discrimination laws to AI models (such as the Fair Housing Act, Civil Rights Act of 1964, the Equal Credit Opportunity Act, and more).

	Total	Audit Report (N)	Audit Report (%)	Transparency Notices (N)	Transparency Notices (%)
Employers Listing NYC Jobs	267	14	5%	12	4%
Employers Not Listing NYC Jobs	124	4	3%	1	1%
All Employers	391	18	5%	13	3%

Note. Chart reproduced from Null Compliance: NYC Local Law 144 and the Challenges Of Algorithm Accountability, by L. Wright, R. M. Muenster, B. Vecchione, T. Qu, P. Cai, A. Smith, COMM/INFO 2450 Student Investigators, J. Metcalf, & J. N. Matias, in *FAccT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, Association for Computing Machinery, 2024, p. 1708 (https://doi.org/10.1145/3630106.3658998).

- **Secure advanced systems.** Leverage security by design principles to enhance cybersecurity within AI systems at the start of their lifecycle; strengthen cybersecurity though training and education; and fund AI-enhanced cybersecurity services and tools within the federal government.

- **Build the innovation workforce.** Support public-private partnerships geared toward upskilling programs; support government funding for AI safety research and infrastructure; and find additional vectors to increase recruitment and cultivation of tech talent in private and public sectors.

Another national technology trade group, NetChoice (2024), developed its own principles for responsible AI use. NetChoice, representing companies such as Amazon, Google, Meta, Netflix, X, and Waymo, has a much more concise statement outlining its vision for "transparency, accountability & security" (Subtitle):

> Rather than create a complex, confusing, and burdensome regulatory scheme for AI development, policymakers should prioritize 3 key principles:
>
> 1. Transparency: "Consumers should be informed when content is created by an AI."
>
> 2. Accountability: "AI is a tool. Therefore, it is subject to the appropriate existing laws and regulations surrounding consumer protection, privacy, and data security."
>
> 3. Security: "AI tools must keep user data secure, and sensitive information should not be used for retaining or retraining AI." ("A Vision for Transparency, Accountability & Security" section)

These guidelines are provided without much detail and lead the authors to wonder whether these companies intend to comply with their self-imposed standards, or if (as they have demonstrated with other technology issues) they would rather advocate for comprehensive federal regulation that they can lobby to modify for their own advantage, while closing off startups from having a competitive footing in the industry.

Local

Even though the Foundation advocates for a state-based regulatory approach that preempts local regulation, it is important to note that some municipalities are acting. For example, "the governments of Seattle, New York City, San Jose, Calif., and Santa Cruz County have all issued independent policies or guidelines for how their employees should use AI on the job" (Davidson, 2023, para. 7). For example, Seattle tailored its city policy to the White House's 2023 AI Executive Order (Lee & Chijioke, 2023). Furthermore, school districts and institutions of higher education have scrambled to grapple with the use of AI by students, staff, and faculty. Initial knee jerk reactions to ban AI in schools have since softened, and, indeed, many schools and teachers are proactively working to integrate it into the classroom (Singer, 2023).[14]

[14] See education case study on page 39, for more discussion on how AI can be used in schools.

The most substantive local ordinance, however, is a New York City hiring law. Local Law 144, passed in 2021 and effective in July 2023, regulates the use of

> automated employment decision tools ("AEDT") [and] prohibits employers and employment agencies from using an automated employment decision tool unless the tool has been subject to a bias audit within one year of the use of the tool, information about the bias audit is publicly available, and certain notices have been provided to employees or job candidates. (City of New York, n.d., para. 1)

Broadly, an AEDT "is a computer-based tool that uses machine learning, statistical modeling, data analytics, or artificial intelligence to substantially help with employment decisions" (New York Foundation for the Arts, 2023, para. 1). Specifically, under Local Law 144, an AEDT is

> any computational process, derived from machine learning, statistical modeling, data analytics, or artificial intelligence, that issues simplified output, including a score, classification, or recommendation, that is used to substantially assist or replace discretionary decision making for making employment decisions that impact natural persons. (Local Law 2021/144, 2021, Sec. 20-870)

Importantly, "there are many types of AI hiring software tools and not all of them are AEDTs. Employers use AI with hiring, screening job applicants, and improving workplace efficiencies" (New York Foundation for the Arts, 2023, "What qualifies as an AEDT?" section). The issue, then, is when a tool becomes an AEDT and when compliance is triggered. For example, the law "requires employers that use certain kinds of software to assist with hiring and promotion decisions—including chatbot interviewing tools and resume scanners that look for keyword matches—to audit those tools annually" (Weber, 2023, para. 2). Under Local Law 144, however, an AEDT

> does not include a tool that does not automate, support, substantially assist or replace discretionary decision-making processes and that does not materially impact natural persons, including, but not limited to, a junk email filter, firewall, antivirus software, calculator, spreadsheet, database, data set, or other compilation of data. (Local Law 2021/144, 2021, Sec. 20-870)

Under Local Law 144, any employment decision substantially impacted by use of an AEDT triggers compliance requirements. An employment decision is not just a final hiring or promotion decision, but where "an AEDT [is used] to substantially help them assess or screen candidates *at any point* [emphasis added] in the hiring or promotion process" (City of New York, 2023, p. 2).

Like many Texas data privacy bills, Local Law 144 has strong transparency provisions, requiring notice, auditing, and disclosure of AEDT use on the company's public website. This is consistent with requirements under Texas HB 4 (2023) and SB 2105 (2023), for example. Furthermore, questions of compliance and enforcement are critical. According to a Cornell University study, notice, compliance, and disclosure are the exception rather than the rule. The Cornell study examined the websites of 391 companies, 267 of which had open job postings in New York City at the time and 124 of which did not (Wright et al., 2024, pp. 1706–1708). Of those with postings, researchers found only 14 audit reports and 12 transparency notices. Of those without postings, researchers found only four audit reports and one transparency notice.

Finally, transparency has been an issue. Cornell researchers argued that "most employers implemented the law in ways that make it practically impossible for job-seekers to learn about their rights or exercise them under Local Law 144" (CAT Lab, n.d., "What we Found" section). Researchers also noted, "In some instances when audit results were posted, finding the notices was 'challenging, time-consuming and frustrating'" (Weber, 2024, para. 5; Wright et al., 2024, p. 1707). This is despite the fact that fines for a first offense can be no more than $500 per violation and subsequent fines must be no less than $500 and no more than $1,500 per violation (Local Law 2021/144, 2021, Sec. 20-872).

State

By the start of the 89th Texas Legislature in 2025, the number of AI-related bills will undoubtedly increase. However, to contextualize the conversation, consider the following statistics. According to the Council of State Governments, "since 2019, 17 states have enacted 29 bills focused on regulating the design, development and use of artificial intelligence" (Wright, 2023, para. 1). Many of the bills early in the 2020s established study committees or task forces to assess state agency use, inventory AI systems, and review procurement, among other things (NCSL, 2024a). According to the Software Alliance, as of September 2023, "state lawmakers [in 31 states have] introduced 440% more AI-related bills in 2023 compared to the prior calendar year. The 190 bills introduced so far in 2023 were more than were filed in the previous two years combined" (BSA, 2023, para. 2). Furthermore, as of June 2024, "in the 2024 legislative session, at least 45 states, Puerto Rico, the Virgin Islands and Washington, D.C., introduced AI bills, and 31 states, Puerto Rico and the Virgin Islands adopted resolutions or enacted legislation" (NCSL, 2024a, para. 3).

Colorado

While numerous states—and even the U.S. Congress—have taken the approach of creating advisory councils or study committees to evaluate the AI landscape and determine the need for regulatory frameworks, Colorado is considered the first state to have established a substantive legislative framework for this technology. The Colorado Legislature passed and Gov. Jared Polis signed SB 24-205 (2024) into law on May 17, 2024, coined "Concerning Consumer Protections in Interactions with Artificial Intelligence Systems" (hereinafter the "Colorado AI Act"). Taking effect on February 1, 2026, this law serves as a consumer protection measure and provides safeguards against risky applications of AI through the following measures. First, it specifies that the Act applies to developers and deployers of AI systems that operate in Colorado. Second, however, many regulatory measures are reserved for "high-risk" AI systems, defined as "any artificial intelligence system that, when deployed, makes, or is a substantial factor in making, a consequential decision" (Sec. 6-1-1701(9)(a)). In this context, consequential decisions apply to areas such as education, finance, government services, employment, and the like. Third, it requires developers and deployers of AI systems to use reasonable care to protect consumers from known or foreseen risks of algorithmic discrimination stemming from high-risk AI systems. Fourth, it requires deployers and developers of AI systems to disclose to consumers that they are interacting with an AI system. Fifth, it requires developers to document to deployers potential harmful or inappropriate uses of a high-risk AI system, the types of data used to train the system, known or reasonably foreseen limitations of the AI system (including potential for algorithmic discrimination), how the system was evaluated for performance and mitigation of algorithmic discrimination, intended outputs of the system, and more. Sixth, it requires deployers to institute a risk management framework to govern the deployer's creation of a high-risk system. In particular, the framework shall "specify and incorporate the principles, processes, and personnel that the deployer uses to identify, document, and mitigate known or reasonably foreseeable risks of algorithmic discrimination" (Sec. 6-1-1703(2)(a)). Seventh, deployers are encouraged to follow the NIST risk management framework standards and must defer to the Colorado attorney general's discretion on additional framework designations pursuant to specific deployers. Eighth, the attorney general has exclusive enforcement authority and is given the power to promulgate rules to better enforce and implement the law. Finally, the law makes it explicit that there is no private right of action.

Of note, Gov. Polis expressed reservations about the Colorado AI Act in a bill signing statement. Claiming to have signed the bill in "hope[s] that it furthers the conversation, especially at the national level" (Polis, 2024, para. 1), the governor had concerns that, rather than focusing on preventing intentional discriminatory conduct, the bill instead chose to regulate "the results of AI system use, regardless of intent" (para. 2). As such, he encouraged lawmakers to reexamine and strengthen the law before it takes effect in 2026 to better protect consumers from nefarious practices while ensuring Colorado is able to advance innovative technologies.

California

Due to the size, technological prowess, and legislative influence the state possesses, it is worth highlighting California's approach to AI regulation as well. Of greatest significance, the California Legislature passed SB 1047 (2024), which was later vetoed by the governor "because it focused too much on regulating the biggest A.I. systems, known as frontier models, without considering potential risks and harms from the technology" (Kang, 2024, para. 3). The bill, introduced as the "Safe and Secure Innovation for Frontier Artificial Intelligence Systems Act," sought to establish a requirement for a developer of covered models to "determine whether it can make a positive safety determination with respect to a covered model before initiating the training of that covered model" (SB 1047, 2024, "Legislative Counsel's Digest" section). Meaning, developers would have needed to demonstrate that their model does not possess hazardous capabilities *before* initiating training. This is significant.

Developers would have also needed to meet other safety standards. First, for example, implementing the capability to initiate a full shutdown of the model if it does not receive a positive safety determination. Second, submit public statements outlining their safety practices and respond to changing compute thresholds, safety guidance, and audit requirements as established by the newly proposed Board of Frontier Models. Finally, it required those operating a computing cluster to implement written policies with procedures outlining how they will respond to customer utilization of computing resources for the purpose of training an AI model.

Moreover, the bill would have formally created a public cloud computing cluster, CalCompute, which would have focused on research and development geared toward the "deployment of large-scale artificial intelligence models and fostering equitable innovation that includes, among other things, a fully owned and hosted cloud platform" (SB 1047, 2024, "LEGISLATIVE COUNSEL'S DIGEST" section). A violation of this bill would have triggered civil penalties as enforced by the attorney general.

This bill received vociferous pushback from industry and advocacy groups. In one opposition letter, signatories expressed concerns with the unreasonable and impractical standard for safety determinations, expensive and onerous compliance requirements, arbitrary definitions, and language that might have resulted in the phasing out of open-source AI systems. Those opposed believed that these concerns "would have a chilling effect on AI research and development in California and potentially across the United States" (Chilson & Stout, 2024, para. 8).

Texas

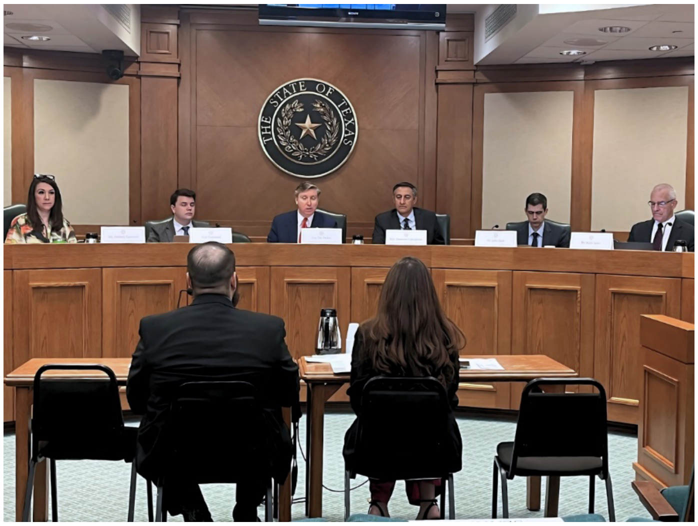
As previously noted, there are scores of other AI bills that have been introduced across the country. Many of these are sector or application specific bills—such as watermarking for AI generated content, penalties surrounding the use of deepfakes targeting minors with sexual content, and the like. Colorado is currently the lone exception of a state taking a comprehensive legal approach. However, Texas has taken steps to advance AI legislation and governance incrementally, positioning itself well to enact a gold standard, comprehensive AI regulatory framework during the 89th Legislature.

First, during the 86th Legislature in 2019, the Legislature passed, and the governor signed, SB 64 (2019) into law. Among other things, the law encourages state agencies and local governments to consider utilizing next generation technologies like artificial intelligence for administering government services (Sec. 2054.601). The enactment of this law resulted in agency deployment of AI systems to streamline government services: "In 2020 the Texas Workforce Commission was able to use AI to help clear their backlog of unemployment claims with a chat bot named 'Larry'" (HB 2060 Bill Analysis, 2023, para 1). For all the promises AI presents to make government run more efficiently, as AI technologies advanced, Texas lawmakers noted that this bill included very few oversight tools to ensure that responsibility, transparency, and ethics were central animating features. In response to these concerns, Texas lawmakers introduced and passed HB 2060 (2023) during the 88th Legislature, which established the Artificial Intelligence Advisory Council to monitor and take inventory of state agency use of AI systems in Texas. After Gov. Greg Abbott signed the bill into law, he articulated his vision:

> As AI becomes more prevalent as a revolutionary tool in our lives and in our workforce, we must ensure that this technology is developed in a responsible and ethical way in Texas to help boost our state's growing economy. ... To protect Texans' privacy and basic civil liberties, I signed legislation creating the Artificial Intelligence Advisory Council to study and monitor artificial intelligence systems developed or used by our state agencies. The Council will help cement Texas' position as a national leader in innovative technology, ensuring our state continues designing and employing the latest and greatest AI technology while prioritizing the security of all Texans. (Office of the Texas Governor, 2023)

Bill authors Rep. Giovanni Capriglione and Sen. Tan Parker were appointed as co-chairs of the Council, and Gov. Greg Abbott appointed additional members, including experts in areas such as ethics, law enforcement, AI systems, and constitutional and legal rights. Through hearings and mandatory inventory reporting for state agencies, the Council is required to submit a report no later than December 1, 2024, with findings and policy recommendations for the 89th Legislature to inform subsequent legislation on responsible public deployment of AI.

In addition to the formal forum created by HB 2060 (2023), the Texas Legislature introduced a multitude of other fora during the 2024 interim year preceding the 89th Legislature. This includes the creation of a House Select Committee on Artificial Intelligence & Emerging Technologies (2024), which was "created to conduct a comprehensive review of the advancements in artificial intelligence and emerging technologies (AI/ET) and the economic, ethical, and societal implications of those advancements" (p. 1). During the interim, the Select Committee conducted several hearings with the goal of gleaning insights for necessary legislative reforms (namely, concerning private actors) heading into the next legislative session. In the Senate chamber, Lt. Gov. Dan Patrick made AI a central focus of his interim charges. In the Business and Commerce, Criminal Justice, and Higher Education Committees, each were charged with conducting hearings and investigating AI's implications in their domains (Patrick, 2024).



Note. The Texas AI Advisory Council was established by HB 2060 (2023) to study the use of artificial intelligence technologies in state government. Photo from *The AI Advisory Council Heard From Representatives of From @TexasDFPS, @TexasHHSC, @ TXAG, @TxDPS, @TDCJ, and @TexasTDI Regarding Their Current and* [Image attached] [Post], by Texas Department of Information Resources [@TexasDIR], X, June 6, 2024 (https://x.com/TexasDIR/status/1798827087542768073).

The above interim strategy on AI analysis and legislative reforms suggests tremendous focus in both chambers on meaningful AI reform, which well-positions the 89th Legislature to advance significant, comprehensive AI reforms. As the history of technology policy in Texas illustrates by major legislation like the Texas Data Privacy and Security Act (TDPSA) (HB 4, 2023), when time during the interim is devoted to hearings, stakeholder meetings, and legislative workshopping on complex technology issues, Texas produces thoughtful, gold standard, pro-human flourishing public policy.

Federal

The robust state efforts highlighted above are a direct response to limited federal action on AI. That said, it is important to provide an overview of the conversation happening at the federal level within the legislative, executive, and judicial branches, as well as within the military.

Legislative Branch

Members of Congress have formed caucuses and working groups, filed legislation, and held hearings on AI. Furthermore, AI bills have been introduced on elections, political advertising, deepfakes, workplace surveillance, education, cybersecurity, critical infrastructure, kids' online safety, intellectual property, and more (Brennan Center for Justice, 2024). However, "currently, there is no comprehensive federal legislation or regulations in the U.S. that regulate the development of AI or specifically prohibit or restrict their use. However, there are existing federal laws that concern AI albeit with limited application" (White & Case, 2024, "AI Regulations" section). For example, the 2018 Federal Aviation Administration Reauthorization Act, 2019 National Defense Authorization Act, and the National AI Initiative Act of 2020 all contain case-specific AI provisions. Furthermore, the U.S. Senate AI Caucus claims credit for 15 AI bills enacted since its inception in 2019. Among them are:

- The Artificial Intelligence Initiative Act and the National AI Research Resource Task Force Act (Sections 5101-5105 and 5106 of P.L.116-283) helped launch the National AI Initiative, a whole-of-government effort to coordinate and expand AI research and development.
- The AI in Government Act (Title I of Division U of P.L.116-260) was the first bill to set rules for the government's own use of AI to ensure its safety and trustworthiness. It also created an AI Center of Excellence within the General Services Administration (GSA) to assist federal agencies in the deployment and use of AI systems.
- The Deepfake Report Act (Section 9004 of P.L.116-283) required the Department of Homeland Security (DHS) to conduct an annual analysis to assess the threat of deepfakes by foreign and domestic entities with a focus on the ways deepfakes can be used to threaten national security.
- The AI for the Armed Forces Act (Sections 1751 and 594 of P.L.116-283) codified AI training recommendations for the military proposed by the National Security Commission on AI. (Martin Heinrich, n.d., "Accomplishments" section)

The Senate also has a Bipartisan Senate AI Working Group which was created at the beginning of the 118th Congress. It held nine AI Insight Forums with more than 150 experts and was noted by the media for "featuring a who's who of Big Tech titans, [including Elon Musk,] Mark Zuckerberg, Bill Gates, Sundar Pichai and Sam Altman" (Wong et al., 2023, para. 2). In May 2024, the Working Group released a 31-page "roadmap" for AI policy delegating work across various committees rather than taking a comprehensive approach like the Texas Legislature is most likely to do. As Sen. Schumer argued, "we're not going to wait on legislation that addresses every aspect of AI in society. … If some areas are ready earlier than others, they should go forward" (Zakrzewski, 2024, para. 6).

The roadmap emphasized eight topics for future legislation: innovation; workforce; high impact uses; elections and democracy; privacy and liability; transparency, explainability, intellectual property and copyright; safeguarding against risks; and national security (Schumer et al., 2024). The plan also recommended spending "at least $32 billion per year for (non-defense) AI innovation" (p. 5). While the tech industry praised the roadmap, the plan was largely criticized by consumer and privacy advocates, with some calling it a "dead end," "pathetic," and "far too vague about how it will protect people from AI's harms" (Zakrzewski, 2024, paras. 2–3). For example, some "were frustrated that the road map only made a cursory mention of AI bias, amid widespread concerns that the technology can replicate and exacerbate harmful stereotypes" (para. 15). Where the Senate goes from here remains to be seen.

The U.S. House of Representatives is also endeavoring to produce a roadmap of its own. To that end, it created the House Task Force on AI in February 2024 (McNerney et al., 2024). The Task Force consists of 12 Republicans and 12 Democrats and "will primarily focus on maintaining U.S. technological competitiveness, protecting national security and addressing regulatory concerns" (Joynes et al., 2024, "Overlap" section).

Executive Branch

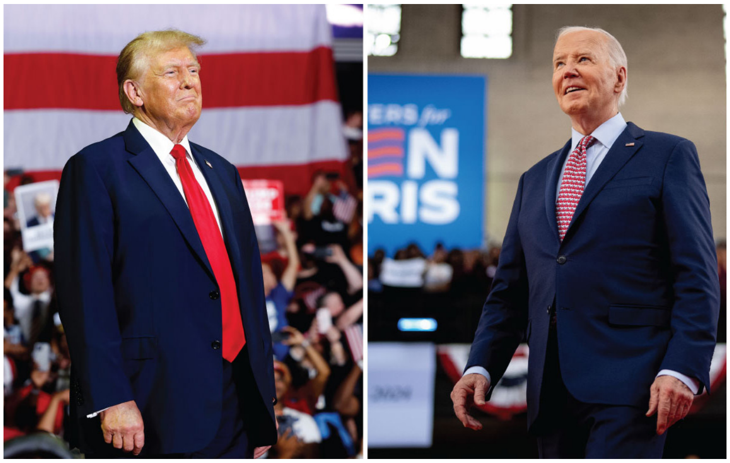
2019 and 2020 Executive Orders (EOs)

President Trump issued two AI executive orders worth noting. First, Executive Order 13859, entitled "Maintaining American Leadership in Artificial Intelligence," was issued in February 2019, emphasizing that "continued American leadership in AI is of paramount importance to maintaining the economic and national security of the United States and to shaping the global evolution of AI in a manner consistent with our Nation's values, policies, and priorities" (Exec. Order No. 13859, 2019, Sec. 1).

Building upon this, and concerning the transparency of federal agency use of AI, it is worth highlighting Executive Order 13960 entitled "Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government" (Exec. Order No. 13960, 2020). Issued by President Trump in December 2020, this EO establishes a framework like the one published by the Texas AI Advisory Council. It notes that

> agencies are already leading the way in the use of AI by applying it to accelerate regulatory reform … combat fraud, waste, and abuse committed against taxpayers … enhance the security and interoperability of Federal Government information systems … and much more. (Sec. 1)

As a guiding principle, it encourages the continued use of AI by federal agencies "when appropriate, to benefit the American people" and, when so used, "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 Executive Order 13859" (Sec. 1). At the same time, this EO also notes the swift pace at which agencies concerned with national security have acted to institute guidelines and principles of their own:



Note. AI-related executive orders and bills have been signed by Presidents Trump and Biden. Photo from Donald Trump and Joe Biden Clinch Their Party Nominations, by M. Gold & N. Nehamas, *The New York Times*, March 12, 2024 (https://www.nytimes.com/2024/03/12/us/politics/trump-republican-nomination.html).

> Certain agencies have already adopted guidelines and principles for the use of AI for national security or defense purposes, such as the Department of Defense's *Ethical Principles for Artificial Intelligence* (February 24, 2020), and the Office of the Director of National Intelligence's *Principles of Artificial Intelligence Ethics for the Intelligence Community* (July 23, 2020) and its *Artificial Intelligence Ethics Framework for the Intelligence Community* (July 23, 2020). Such guidelines and principles ensure that the use of AI in those contexts will benefit the American people and be worthy of their trust. (Sec. 1)

The order also outlines principles that agencies shall adhere to, such as the Constitution, all applicable laws and policies—particularly those associated with privacy and civil liberties—and respect for American values broadly. It emphasizes transparency and accountability and requires agencies to take inventory of how they are using AI systems and in turn disclose that to the public, Congress, and other appropriate stakeholders. Finally, all agencies are to be held accountable for implementing and enforcing AI safeguards, which include auditing, the proper training of staff, and other responsibilities that will align agency use of AI with the espoused principles.

2022 AI Blueprint

As noted above, AI-related executive orders and bills have been signed by Presidents Trump and Biden. However, with the explosion in AI in recent years, the Biden administration's White House Office of Science and Technology Policy (2022) began a listening tour in 2021 that culminated in the publication of the "Blueprint for an AI Bill of Rights." The blueprint "identified five principles that should guide the design, use, and deployment of automated systems to protect the American public in the age of artificial intelligence": safe and effective systems; algorithmic discrimination protections; data privacy; notice and explanation; and human alternatives, considerations, and fallback (p. 3).

As Strickland (2022) noted, "the nonbinding principles are being both celebrated and vilified" (Subheadline section). For example, Haven (2022) viewed it as a pro-civil rights document benefitting marginalized groups. On the other hand, the Surveillance Technology Oversight Project (2022) criticized the proposal as anti-civil rights, arguing that "the blueprint proposes that all AI will be built with consideration for the preservation of civil rights and democratic values, but endorses use of artificial intelligence for law enforcement surveillance … [which] normalizes biased surveillance and will accelerate algorithmic discrimination" (para. 1). The organization's executive director also argued, "we don't need a blueprint, we need bans" (para. 2).

Friedland (2022) also took pause with the word blueprint, noting that "affixing 'Blueprint' to the 'AI Bill of Rights' seems to indicate a narrowing of ambition from the original proposal" ("Government Updates" section). On the other hand, Castro (2022) argued the Blueprint goes too far and offered pointed criticism: "The AI Bill of Rights is an insult to both AI and the Bill of Rights. Americans do not need a new set of laws, regulations, or guidelines focused exclusively on protecting their civil liberties from algorithms" (para. 2). Castro was also concerned that "the AI Bill of Rights vilifies digital technologies like AI as 'among the great challenges posed to democracy'" and lamented "that the highest officials in the nation have labeled [AI] dangerous, biased, and ineffective" (para. 3).

Voluntary Commitments

In July 2023,[15] September 2023,[16] and July 2024,[17] the White House received voluntary commitments from 16 major AI-related companies. The voluntary commitments, based on the pillars of safety, security, and trust, are designed to "encourag[e] this industry to uphold the highest standards to ensure that innovation doesn't come at the expense of Americans' rights and safety" (The White House, 2023b, para. 2). Among these commitments, companies will conduct internal and external security testing before releasing AI systems; share information necessary to manage AI risks with academia, government, and other stakeholders; invest in cybersecurity protections; create a public reporting mechanism for system issues after release; include transparency mechanisms like watermarks; publish transparent reports on capabilities, limits, and risks; prioritize research on ways to minimize harmful effects; and prioritize development of systems to address significant civilizational challenges (The White House, 2023a).

2023 Executive Order

Furthermore, in October 2023, President Biden issued Executive Order 14110 (Exec. Order No. 14110, 2023) which "establishes new standards for AI safety and security, protects Americans' privacy, advances equity and civil rights, stands up for consumers and workers, promotes innovation and competition, advances American leadership around the world, and more" (The White House, 2023d, para. 1). The executive order is sweeping, with the Congressional Research Service noting that it "directs over 50

[15] Amazon, Anthropic, Google, Inflection, Meta, Microsoft, and OpenAI (The White House, 2023b).

[16] Adobe, Cohere, IBM, NVIDIA, Palantir, Salesforce, Scale AI, and Stability (The White House, 2023c).

[17] Apple (The White House, 2024).

federal entities to engage in more than 100 specific actions to implement the guidance set forth across eight overarching policy areas" (Harris & Jaikaran, 2024, "Summary" section). The executive order addresses AI safety and security, data privacy, civil rights, bias and discrimination, consumer protection, healthcare and patient protections, workforce issues, intellectual property, innovation, competition, global leadership, and the appropriate uses of AI in the federal government, among other areas (Sidley Austin, 2023).

For our purposes, several provisions are worth highlighting. First, as discussed in more detail below, the executive order directs NIST to coordinate with other agencies to "establish guidelines and best practices, with the aim of promoting consensus industry standards, for developing and deploying safe, secure, and trustworthy AI systems" (Exec. Order No. 14110, 2023, Sec. 4.1). Second, the largest section discusses AI risks in national security and critical infrastructure. To that point, one of the bolder provisions invokes the Defense Production Act to require that certain dual-use foundational models are tested and safe, secure, and trustworthy *before* they are made public.[18] Third, it tasks various agencies with studying the impact of AI and automation on the workforce, both public and private, and exploring ways to better educate, train, upskill, and otherwise close the talent gap. Fourth, it directs the attorney general and related law enforcement entities to develop best practices for training and use of AI technologies. Finally, it requires an examination of intellectual property issues such as inventorship, authorship, copyright protections and infringements, and the like.

Responses to the executive order have been mixed. Former head of the White House Office of Science and Technology Policy Alondra Nelson was "highly supportive and complimentary of the Biden administration's EO" and noted the need for and importance of this effort in the absence of a federal data privacy law (Krishan, 2023, para. 21). Amba Kak of the AI Now Institute called the executive order "one of the biggest achievements in the last decade in AI policy" (Booth, 2024, para. 8). U.S. Sen. Mark Warner and Rep. Yvette Clark, both Democrats, issued statements of support (Bordelon, 2023). As noted above, 16 AI industry leaders also signed onto the voluntary commitments, which greatly informed the substance of the executive order (Teshome, 2024).

Furthermore, several polls taken after the executive order was issued have found bipartisan support for enhanced AI regulation more broadly. One poll found that "more than 70% of Democrats and Republicans support developing standards to test that AI systems are safe" (Carmichael, 2023, Subheadline section). Another poll found that "69% of American voters—including 64% of Republicans and 65% of Independents—support the executive order while just 15% oppose it" (Artificial Intelligence Policy Institute, n.d., para. 2). More recently, Perrigo (2024) reported on a June 2024 poll that found that

> 75% of Democrats and 75% of Republicans believe that "taking a careful controlled approach" to AI—by preventing the release of tools that terrorists and foreign adversaries could use against the U.S.—is preferable to "moving forward on AI as fast as possible to be the first country to get extremely powerful AI." (para. 3)

Others had mixed views on the 2023 executive order. First, the American Civil Liberties Union (ACLU), applauded portions of the order but criticized it for "essentially kick[ing] the can down the road" insofar as it does not sufficiently outline law enforcement use of AI technologies (Boak & O'Brien, 2023, para.

[18] According to Williams (2023),

> the EO uses the Korean War-era Defense Production Act, an act typically invoked during national emergencies, to require that developers of powerful foundation models tell the government when they are training the models and share the results of all red-team safety tests. (para. 4)

Of note, the Defense Production Act has been invoked by President Bush in response to Hurricane Katrina, by President Obama in response to a Chinese spyware threat, and by Presidents Trump and Biden during the COVID-19 pandemic for manufacturing, logistics, and public health ends (McPherson-Smith, 2024; Siripurapu, 2021).

25). Second, Paul Lekas, senior vice president for tech industry trade association Software & Information Industry Association (SIIA), said,

> while we are pleased the foundation model review process is focused on high-risk use cases—those that involve national security, national economic security, and national public health and safety—we are concerned that the EO imposes requirements on the private sector that are not well calibrated to those risks and will impede innovation that is critical to realize the potential of AI to address societal challenges. (SIIA, 2023, para. 10)

Finally, as Tom Romanoff of the Bipartisan Policy Center noted,

> the fact that we're starting to see pushback on the EO is not surprising from trade groups saying it's too broad and could impede innovation. … But everyone sees there's a need for regulation to happen and both parties on [Capitol] Hill have been supportive, received it well. (Krishan, 2023, para. 17)

However, others were blunter in their criticism. First, Big Tech trade association NetChoice characterized the executive order as broad, burdensome, stifling competition, threatening the U.S.'s global standing, dangerous, duplicative, and wrong (Chavez, 2023). NetChoice argued that the Biden administration "has chosen to further increase the complexity and burden of the federal code" (para. 2). Furthermore, NetChoice posited that these "broad regulatory measures in Biden's AI red tape wishlist will result in stifling new companies and competitors from entering the marketplace and significantly expanding the power of the federal government over American innovation" (para. 3). Ultimately, NetChoice argued that "this order puts any investment in AI at risk of being shut down at the whims of government bureaucrats. That is dangerous for our global standing as the leading technological innovators, and this is the wrong approach to govern AI" (para. 3).

Second, the U.S. Chamber of Commerce (2023) questioned the process, compressed timeline, limits on stakeholder input, and risk of agency overreach. The Chamber noted that "short overlapping timelines for agency-required action endangers necessary stakeholder input, thereby creating conditions for ill-informed rulemaking and degrading intra-government cooperation" (para. 4). The Chamber also cautioned that agencies "should not view this as a license to do as they please—all agencies must continue to act within the limits of their Congressional mandates and abide by the Major Questions Doctrine as articulated by the Supreme Court" (para. 4).

Third, and related to the Chamber's concern, U.S. Sen. Mike Rounds, an AI leader on Capitol Hill, pushed back on the invocation of the Defense Production Act. According to Sen. Rounds, "there's not a national emergency" on AI and this is "not necessarily what the Defense Production Act was made for in the first place" (Chatterjee & Bordelon, 2024, para. 4).

Fourth, Protect AI criticized the outsized role of large tech companies in the stakeholder process and its potential to squeeze out competition. As Protect AI noted,

> it's a good first step, but what I think is lacking here is we need to get more people at the table in the room than just the big three or big five technology companies. Those that work on AI risk and security should be included, too. I think that's a missing element. (Krishan, 2023, para. 11)

Fifth, Susarla (2023) astutely noted the interplay of data privacy laws and AI regulation:

> Without strong data privacy laws in the U.S. as other countries have, the executive order could have minimal effect on getting AI companies to boost data privacy. In general, it's difficult to measure the impact that decision-making AI systems have on data privacy and freedoms. ("What the executive order doesn't do" section)

Finally, for its part, the Foundation has also been critical of the 2023 executive order, with one author of this paper, David Dunmoyer, noting,

> while many of the principles and goals outlined in this order are laudable, upon scrutiny, it is not clear how they will be accomplished and by whom. When something with so much time invested in it is this opaque, one can safely assume this opacity serves a goal—to give the administration a tool to regulate and set the stage for a new government agency. … Much like Obamacare years ago, this is a road to regulatory hell paved with good intentions. (Texas Public Policy Foundation, 2023, para. 2)

Case Study: National Institute of Standards and Technology (NIST) Standards

Of all the agencies addressing the responsible development and regulation of AI, NIST is the one of greatest contemporary prominence. NIST (n.d.-a), in response to calls from leading scientists and industrialists, was forged by the U.S. Congress in 1901 to serve as the leading authority on matters of technological development and national standards. As part of the Department of Commerce, its mission is "to promote U.S. innovation and industrial competitiveness by advancing measurement science, standards, and technology in ways that enhance economic security and improve our quality of life" (NIST, 2022, para. 4).

NIST has provided standards, measurements, and frameworks for varying technologies over the years. The agency's first significant and formal foray into AI came with the publication of its Artificial Intelligence Risk Management Framework (AI RMF 1.0), published in January 2023 (NIST, 2023a). Importantly, this framework was developed to provide guidance to organizations looking to design, develop, deploy, or incorporate AI systems into their domain in a responsible manner that maximizes benefits and reduces risk. As stated by Deputy Commerce Secretary Don Graves in the release of the AI RMF 1.0,

> this voluntary framework will help develop and deploy AI technologies in ways that enable the United States, other nations and organizations to enhance AI trustworthiness while managing risks based on our democratic values. … It should accelerate AI innovation and growth while advancing—rather than restricting or damaging—civil rights, civil liberties and equity for all. (NIST, 2023b, para. 3)

Ultimately, this framework extols the outcome of prudent AI governance, stemming from a culture of risk management that is proactive and carefully reactive to the evolving AI landscape.

NIST made clear that the AI RMF 1.0 would be the first of many iterations, with new frameworks being developed in conjunction with new advancements in the field and application of AI. In October 2023, the same date the aforementioned Executive Order 14110 was issued, the U.S. Department of Commerce (2023) announced that NIST "will develop industry standards for the safe and responsible development of frontier AI models, create test environments to evaluate these systems, and develop standards on privacy and on authenticating when content is AI-generated" (para. 3). In July 2024, NIST (2024a) released a companion AI RMF on GAI specifically. Specifically, this report "defines risks that are novel to or exacerbated by the use of GAI. … [It] provides a set of suggested actions to help organizations govern, map, measure, and manage these risks" (p.1). This exemplifies the NIST approach of establishing broad principles and frameworks for the field of AI as applied to organizations, receiving stakeholder input, and releasing companion publications focusing on use-case profiles and specific AI functions that might introduce greater risk.

Other Federal Agency Actions

As several sections of this paper have alluded to, questions persist as to whether new laws or regulations are needed to enforce responsibility in AI. For example, before President Biden issued Executive Order 14110, in April 2023 the Federal Trade Commission (FTC) (2023), Equal Employment Opportunity Commission, Consumer Financial Protection Bureau, and Department of Justice issued a joint statement arguing that "existing legal authorities apply to the use of automated systems and innovative new technologies just as they apply to other practices" (para. 3).

Note. U.S. officials are worried about the threats AI presents to nuclear weapons, national security, cybersecurity, critical infrastructure, and more. These threats can come from hostile state actors like China, Russia, North Korea, and Iran, as well as from cyber criminals, terrorists, and other non-state actors. Pictured are Chinese President Xi Jinping (L), U.S. President Joe Biden (C), and Russian President Vladimir Putin (R). Photo from *Biden Administration Skeptical of Xi's Intentions Ahead of His Summit With Putin*, by K. Liptak & K. Atwood, CNN, March 17, 2023 (https://www.cnn.com/2023/03/17/politics/biden-putin-xi/index.html).

However, since Executive Order 14110 was issued, federal agencies have taken inventory and even enacted new regulations related to AI. This EO includes a series of deadlines for covered federal agencies, most in the range of 90–365 days (Exec. Order No. 14110, 2023).[19] As of July 2024, all agencies have met commitments under the 270-day timeline (The White House, 2024). Among those actions undertaken so far are:

- A Defense Production Act determination on AI safety test results
- A national security memo on AI
- Actions to combat harmful image-based and synthetic AI generative content
- U.S. Commerce Department draft rules on U.S.-based Infrastructure as a Service firms
- Developed testing parameters at the Department of Energy
- First technical guidelines from the AI Safety Institute
- The Federal Communications Commission (FCC) applied restrictions in the Telephone Consumer Protection Act to regulate the use of AI-generated voices
- FTC's proposed update to the Children's Online Privacy Protection Act (COPPA) Rule
- Initial guidance for agencies on AI training data
- Launch of a federal AI hiring spree
- Launch of the National AI Research Resource pilot to support AI research and education
- Pilot programs at the Department of Defense (DoD) and DHS on vulnerabilities in government networks
- Published NIST frameworks on generative AI and dual-use foundational models (The White House, 2024; FCC, 2024; Sokler et al., 2024; Mitchell, 2024)

Other Actions to Counter International Threats

While military and diplomatic issues will be further examined in a subsequent section, it is worth noting a few adjacent points here. As Reuters (2024a) noted, "worries mount that U.S. adversaries could use the [AI] models, which mine vast amounts of text and images to summarize information and generate content, to wage aggressive cyber attacks or even create potent biological weapons" (para. 3).

Furthermore, according to the DHS (2023a) Homeland Threat Assessment 2024, "cyber actors use AI to develop new tools and accesses that allow them to compromise more victims and enable larger-scale, faster, efficient, and more evasive cyber attacks" (p. v). For example, these threats, from hostile nation-states, terrorists, and other non-state actors, are driven by "AI-developed malware and AI-assisted software development—technologies that have the potential to enable larger scale, faster, efficient,

[19] The White House (2024), Mayer Brown (2024), and Sokler et al. (2023) provide helpful timelines and trackers for agency action under Executive Order 14110.

and more evasive cyber attacks—against targets, including pipelines, railways, and other U.S. critical infrastructure" (p. 18).

To counter these threats, policymakers are considering applying tools like "know your customer" (KYC) rules,[20] export controls, and investment controls. These have been a source of contention between the West and China. For example, in January 2024, the U.S. Commerce Department proposed a rulemaking to limit access to U.S. data centers by foreign entities. As Commerce Secretary Gina Raimondo noted:

> We can't have non-state actors or China or folks who we don't want accessing our cloud to train their [AI] models. … We use export controls on chips. Those chips are in American cloud data centers so we also have to think about closing down that avenue for potential malicious activity. (Shepardson, 2024, para. 2)

Furthermore, export controls have been implemented. It is important to note at the outset that,

> as of now, U.S. export controls currently do not specifically "control" AI as a broad category. Instead, the different components that contribute to the development of AI are controlled in a variety of ways, the majority of which still fit in an uncontrolled or a lightly controlled category of the Export Administration Regulations (EAR99).
>
> These components include, but are not limited to,
>
> - integrated circuits/semiconductors;
> - technology for designing, developing, adapting, or embedding AI functionality into products or platforms;
> - equipment to manufacture the integrated circuits/semiconductors used for AI functionality; and
> - assistance deemed to be "US support" or facilitation in these areas and tangentially covered by other direct or indirect items or activities. (Plotinsky & Cinelli, 2024, "U.S. Export Controls Law and Regulations" section)

For example, in October 2022, "the Biden administration announced a new export controls policy on artificial intelligence (AI) and semiconductor technologies to China" (Allen, 2022, p. 1).[21] Allen argued that "these actions demonstrate an unprecedented degree of U.S. government intervention to not only preserve chokepoint control but also begin a new U.S. policy of actively strangling large segments of the Chinese technology industry—*strangling with an intent to kill* [emphasis added]" (Allen, 2022, p. 2).

However, federal policymakers are seeking to go beyond Plotinsky & Cinelli's (2024) analysis of current export control law, as the Biden administration and Congress are considering export controls on both closed- and open-source systems, respectively. For example, in May 2024, Reuters reported that "the Commerce Department is considering a new regulatory push to restrict the export of proprietary or closed source AI models, whose software and the data it is trained on are kept under wraps" (Alper, 2024, para. 2). A bipartisan group of U.S. House members are

[20] KYC refers to

> a set of guidelines that financial institutions and businesses follow to verify the identity, suitability, and risks of a current or potential customer. The goal is to identify suspicious behavior such as money laundering and financial terrorism before it ever materializes.
>
> KYC regulations originated from years of unchecked financial crimes. The initial guidelines were drafted in 1970 when the U.S. passed the Bank Secrecy Act (BSA) to prevent money laundering. Notable additions came years later, after the Sept. 11, 2001 terrorist attacks and 2008 global financial crisis. (Dow Jones, n.d., paras. 2–3)

[21] Calling them "chokeholds," Allen (2022) characterized the Biden administration's purpose for such export controls as follows:

> In short, the Biden administration is trying to (1) strangle the Chinese AI industry by choking off access to high-end AI chips; (2) block China from designing AI chips domestically by choking off China's access to U.S.-made chip design software; (3) block China from manufacturing advanced chips by choking off access to U.S.-built semiconductor manufacturing equipment; and (4) block China from domestically producing semiconductor manufacturing equipment by choking off access to U.S.-built components. (p. 2)

willing go even further, introducing a bill to "remove roadblocks to regulating the export of open source AI," "bar Americans from working with foreigners to develop AI systems that pose risks to U.S. national security," and, more broadly, "to bulletproof any future AI export regulations from legal challenges" (Reuters, 2024a, paras. 2–7). And because the threat is larger than China, Alper (2024) reported that "any new export control would likely target Russia, China, North Korea and Iran" (para. 7).

Finally, the U.S. is considering inbound and outbound investment restrictions to counter the national security threats posed by AI. In September 2022, President Biden issued Executive Order 14083 (Exec. Order No. 14083, 2022) directing the Committee on Foreign Investment in the United States to further scrutinize inbound investments, including by countries of special concern, in critical and emerging technologies such as AI (Huff et al., 2022; Sidley Austin, 2022). Furthermore, in August 2023, President Biden signed Executive Order 14105 (Exec. Order No. 14105, 2023),

> which directed the US Department of the Treasury to establish new regulations to restrict certain outbound investment to countries of concern. The EO and Treasury's anticipated regulations currently cover three categories of sensitive technology: semiconductors and microelectronics, quantum information technologies, and AI. The EO provides the administration leeway to add other technology areas or sectors, and it remains possible that new sensitive areas such as biotechnology and battery technology could be added. For now, the short list reflects the acute focus by the US government on AI. (Plotinsky & Cinelli, 2024, "Outbound Investment and AI" section).

However, absent congressional action, because the Treasury Department is still in the rulemaking process, it is unlikely these regulations will go into effect until at least 2025.

Judicial Branch

As noted above, Chief Justice John Roberts's (2023) annual report urged "caution and humility" in the use of AI in the judicial system, driven by, in his view, "a persistent public perception of a 'human-AI fairness gap,' reflecting the view that human adjudications, for all of their flaws, are fairer than whatever the machine spits out" (pp. 5–6). This caution pertains to how AI is used in briefs, raises questions about admissibility and privilege, and presents serious challenges to the due process rights of criminal defendants, for example.

Lat (2024b) argued that the judiciary is being overreactive in regulating AI and instead should use "rationality, not red tape" (Subheadline section). According to Lat (2024a), "at least 21 federal trial judges have already issued standing orders regarding AI," and at least three circuit courts are studying or considering proposals (para. 3).

The U.S. Judicial Conference's Advisory Committee on Evidence Rules met in April 2024 and "struggled to determine whether or how to draft rules that would allow courts to ensure the authenticity and reliability of trial evidence generated by artificial intelligence" (Raymond, 2024, para. 1). Skeptical of whether to draft rules at all, one federal appellate judge noted, "I'm not sure that this is the crisis that it's been painted as, and I'm not sure that judges don't have the tools already to deal with this" (para. 5). Another federal district judge concurred: "I think we have adequate tools at the moment, but this may change as things develop" (para. 7).

Considering the speed at which AI technologies are emerging and the heavy federal docket,[22] the federal court system will likely deal with AI questions for years to come.

Military and Diplomacy

Military and defense officials have long researched, developed, and implemented AI technologies for

[22] From March 2022 to March 2023, there were 353,170 district court filings, 403,273 bankruptcy filings, and 42,536 appellate filings (United States Courts, n.d.). The authors calculated 42,536 by adding together regular appellate filings (40,681), appeals to bankruptcy appellate panels (373), and appeals to the Court of Appeals for the federal circuit (1,482). In October Term 2022, the U.S. Supreme Court received 4,159 case filings (Roberts, 2023, p. 8).

strategic and operational advantages, going back to early developments and uses by Turing and others in World War II. It was not until 2018, however, that the DoD (n.d.) issued its first AI strategy (Plotinsky & Cinelli, 2024). Today's AI applications seek far more than cracking the Enigma code. Autonomous applications are more sweeping than autonomous drones and robot soldiers—which, as mentioned earlier in this report, is being pursued by the military (Demarest, 2024). On an enterprise level,

> military AI capabilities includes not only weapons but also decision support systems that help defense leaders at all levels make better and more timely decisions, from the battlefield to the boardroom, and systems relating to everything from finance, payroll, and accounting, to the recruiting, retention, and promotion of personnel, to collection and fusion of intelligence, surveillance, and reconnaissance data. (Vergun, 2023, para. 3)

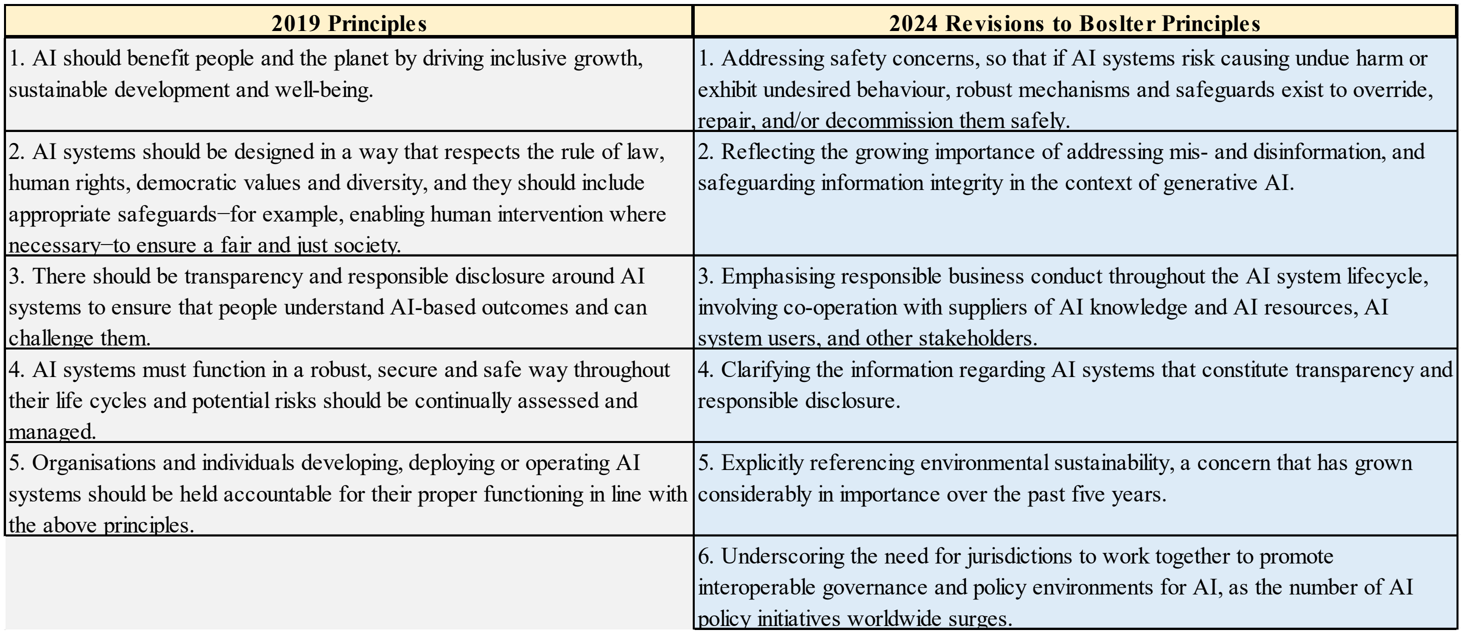
Note. AI applications in the military go beyond the battlefield. AI is being integrated enterprise wide: in the boardroom, finance, payroll, accounting, recruiting, retention, promotion, intelligence, surveillance, reconnaissance, and more (Vergun, 2023). Photo from *AI-Enabled Ground Combat Vehicles Demonstrate Agility and Synergy at PC21*, by M. Thompson, U.S. Army, November 1, 2021 (https://www.army.mil/article/251632/ai_enabled_ground_combat_vehicles_demonstrate_agility_and_synergy_at_pc21).

For example, the 2023 DoD AI strategic plan is built on five hierarchical needs: "quality data, governance, insightful analytics and metrics, assurance and responsible AI" (Clark, 2023, para. 12). The plan seeks an agile approach to AI development and application, emphasizing speed of delivery and adoption at scale leading to five specific decision advantage outcomes:

- Superior battlespace awareness and understanding
- Adaptive force planning and application
- Fast, precise and resilient kill chains
- Resilient sustainment support
- Efficient enterprise business operations (paras. 8–9)

Internationally, according to Vergun (2023), "the United States government is leading global efforts to build strong norms that will promote the responsible military use of artificial intelligence and autonomous systems" (para. 1). The operating document, titled "The Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy," has 56 signatories as of October 2024 (U.S. Department of State, 2024). Among the document's principles are human oversight, monitoring, and auditing; exercising care in development, deployment, and use; establishing well-defined use cases; minimizing bias; and ensuring compliance with international and domestic law (U.S. Department of State, 2023).

Absent from the list of AI Political Declaration signatories are the usual suspects: China, Russia, North Korea, and Iran. The U.S. State Department has urged China and Russia, in particular, to match its "clear and strong commitment" to total human control of nuclear weapons (Torode, 2024, para. 2). This concern was underscored by recent research, which found that "large language models have a potentially dangerous tendency to go nuclear" (Heath, 2024, para. 3; Lamparth & Schneider, 2024). Specifically, "when they [the researchers] tested LLMs from OpenAI, Anthropic and Meta in situations like simulated war games, the pair found the AIs suggested escalation, arms races, conflict—and even use of nuclear weapons—over alternatives" (Heath, 2024, para. 4). These concerns are not new, and even before Lamparth and Schneider's essay, branches of the U.S. military limited, paused, or abandoned the use of generative AI (Heath, 2024; Manson, 2023; Rathbun, 2023).



2019 Principles	2024 Revisions to Boslter Principles
1. AI should benefit people and the planet by driving inclusive growth, sustainable development and well-being.	1. Addressing safety concerns, so that if AI systems risk causing undue harm or exhibit undesired behaviour, robust mechanisms and safeguards exist to override, repair, and/or decommission them safely.
2. AI systems should be designed in a way that respects the rule of law, human rights, democratic values and diversity, and they should include appropriate safeguards—for example, enabling human intervention where necessary—to ensure a fair and just society.	2. Reflecting the growing importance of addressing mis- and disinformation, and safeguarding information integrity in the context of generative AI.
3. There should be transparency and responsible disclosure around AI systems to ensure that people understand AI-based outcomes and can challenge them.	3. Emphasising responsible business conduct throughout the AI system lifecycle, involving co-operation with suppliers of AI knowledge and AI resources, AI system users, and other stakeholders.
4. AI systems must function in a robust, secure and safe way throughout their life cycles and potential risks should be continually assessed and managed.	4. Clarifying the information regarding AI systems that constitute transparency and responsible disclosure.
5. Organisations and individuals developing, deploying or operating AI systems should be held accountable for their proper functioning in line with the above principles.	5. Explicitly referencing environmental sustainability, a concern that has grown considerably in importance over the past five years.
	6. Underscoring the need for jurisdictions to work together to promote interoperable governance and policy environments for AI, as the number of AI policy initiatives worldwide surges.

Note. The evolution of the OECD's principles over the last five years reflects the increased granularity and specificity (for better or for worse) that lawmakers, industry, and other stakeholders are considering. Chart reproduced from *What Are the OECD Principles on AI?* by Organisation for Economic Co-operation and Development, March 2, 2020 (https://www.oecd-ilibrary.org/sites/6ff2a1c4-en/index.html?itemId=/content/paper/6ff2a1c4-en) and *OECD Updates AI Principles to Stay Abreast of Rapid Technological Developments*, by Organisation for Economic Co-operation and Development, May 3, 2024 (https://www.oecd.org/en/about/news/press-releases/2024/05/oecd-updates-ai-principles-to-stay-abreast-of-rapid-technological-developments.html).

International

Overview

Perhaps unsurprisingly, international regulation is further along than regulation in the United States. According to Whyman (2023), "31 countries have passed AI legislation and 13 more are debating AI laws" ("International context" section). The most comprehensive AI regulatory framework is the EU AI Act, discussed at length below. As one data scientist noted, "America makes software. Asia makes hardware. Europe makes it difficult" (Tunguz, 2024). Chee and Hummel (2024) fairly characterize the regulatory landscape this way: "The European Union's AI Act is more comprehensive than the United States' light-touch voluntary compliance approach while China's approach aims to maintain social stability and state control" (para. 2).

However, other parts of the world are weighing in as well. For example, despite lobbying from the EU, the Association of Southeast Asian Nations (ASEAN)[23] has pushed back and argued for a laissez-faire approach that is also sensitive to the varying cultural and linguistic differences in the region (Potkin & Mukherjee, 2023).

The Vatican, for its part, is playing a significant role in the global AI conversation. The Vatican's top AI advisor, Friar Paolo Benanti, has advocated for transparent, ethical, and "'human-centric' artificial intelligence that shouldn't be allowed to run rampant" (Volpicelli, 2024, para. 2). In June 2024, Pope Francis spoke at the Group of Seven summit, arguing that "no innovation is neutral" and that AI risks trapping the world in a "technocratic paradigm" (McLellan, 2024, paras. 2–3).

Furthermore, intergovernmental organizations like the OECD and the United Nations (UN) have also opined on AI regulation. The signing of OECD's

[23] ASEAN (n.d.) member countries include Brunei Darussalam, Cambodia, Indonesia, Lao PDR, Malaysia, Myanmar, Philippines, Singapore, Thailand, and Viet Nam.

Reuters reported that “worries are mounting among U.S. officials about China’s access to AI technology, amid fears that it could be used by Beijing to upend elections in other countries, create bioweapons and launch cyberattacks.”

initial AI principles was the first broad (although nonbinding) international agreement on AI policy, first adopted in May 2019 by 42 countries[24] (Vincent, 2019). The United States, under the Trump administration, supported them (Thomas, 2019). The initial principles and recommendations, noted in the adjacent chart, were founded with the aim of benefitting people and the planet, respect for human dignity and the rule of law, transparency, accountability, and robust, secure, and safe systems (OECD, 2019). Those principles, summarized in the following chart, were updated in May 2024 and have 47 adherents, including the EU (OECD, 2024).

The OECD (n.d.) AI principles are important because “the European Union, the Council of Europe, the United States, and the United Nations and other jurisdictions use the OECD’s definition of an AI system and lifecycle … in their legislative and regulatory frameworks and guidance” (“Principles for trustworthy AI” section).

Finally, the UN has considered how to address AI since at least 2019, with a number of inter-agency working groups, task forces, and research units studying the issue (UN System Chief Executives Board for Coordination, n.d.). In March 2024, with unanimous support from all 193 member nations, the UN adopted its first resolution on AI, which was sponsored by the United States and the Biden administration (Lederer, 2024). The aim of the resolution was to show “global support to an international effort to ensure the powerful new technology benefits all nations, respects human rights and is ‘safe, secure and trustworthy’” (para. 1; Resolution A/78/L.49, 2024, p. 1).

China

Perhaps the greatest geopolitical challenge presented by AI comes from China, which is driven by its desire to be the dominant world leader in AI by 2030. According to the Center for Strategic & International Studies, Chinese President Xi Jinping “has characterized the AI-driven industrial revolution as an opportunity for China to ‘overtake on the curve’ (i.e., surpass incumbents)” (Arcesati & Creemers, 2024, para. 5). In addition to making significant investments in technological development, China was among the first countries to develop AI policies (The State Council of the People’s Republic of China, 2017b; The State Council of the People’s Republic of China, 2017a; The Economist, 2017). These efforts have increased this decade, with Sheehan (2022) reporting that since at least mid-2021, “the Chinese government has rolled out a series of policy documents and public pronouncements that are finally putting meat on the bone of the country’s governance regime for artificial intelligence” (para. 1). According to a translation of the seminal Chinese Communist Party (CCP) policy directive, China’s goal is

> to seize the major strategic opportunity for the development of AI, to build China’s first-mover advantage in the development of AI, to accelerate the construction of an innovative nation and global power in science and technology, in accordance with the requirements of the CCP Central Committee and the State Council. (The State Council of the People’s Republic of China, 2017, p. 2)

China is beginning to fulfill its goal to be the global AI leader by influencing areas in the developing and emerging world, including those on the Belt and Road Initiative and Digital Silk Road (p. 24; Arcesati & Creemers, 2024; Williams, 2024).

Reuters (2024b) reported that “worries are mounting among U.S. officials about China’s access to AI technology, amid fears that it could be used by Beijing to upend elections in other countries, create bioweapons and launch cyberattacks” (para. 4). For

[24] As of August 2024, 38 OECD member nations, eight non-member nations, and the EU adhere to the OECD (n.d.) AI principles.

example, the Microsoft Threat Analysis Center found that China is using generative AI visual materials for divisive political meddling in the U.S., particularly on social media (Watts, 2023). More broadly, FBI Director Christopher Wray has been very vocal about the Chinese threat, with Tucker (2023) reporting that "he was 'deeply concerned' about the Chinese government's artificial intelligence program, asserting that it was 'not constrained by the rule of law'" (para. 1; World Economic Forum, 2023). Tucker further reported that

> Wray said Beijing's AI ambitions were "built on top of massive troves of intellectual property and sensitive data that they've stolen over the years." He said that left unchecked, China could use artificial intelligence advancements to further its hacking operations, intellectual property theft and repression of dissidents inside the country and beyond. (paras. 2–3)

Furthermore, import and export controls on computer chips, hardware, software, and other AI technologies, tools, and investments have been a source of contention between the nations (Hawkins, 2024; Shivakumar et al., 2023). However, diplomatic channels are open, and U.S. and Chinese officials held their "first high-level talks over the risks of artificial intelligence" in May 2024 to find a sound bilateral and even a broader international approach to AI use (Dou, 2024, para. 1).

The European Union Artificial Intelligence Act

The remainder of this section will examine the European Union's Artificial Intelligence Act [hereinafter the "EU AI Act" or simply the "Act"], the most comprehensive AI framework to date (Regulation (EU) 2024/1689, 2024). In May 2018, the EU adopted the General Data Protection Regulation (GDPR), which is the world's most comprehensive data privacy law (Regulation (EU) 2016/679, 2016). Foreshadowing the AI debate to come, the GDPR "grants their citizens a 'right to explanation' if they are affected by algorithmic decision-making" (Xu et al., 2019, p. 566).

On the heels of GDPR, the concept of a European AI law emerged in 2021 and went through various draft versions as it moved through the legislative process. In March 2024, the European Parliament voted 523-46 to approve the final version, which went into effect on August 1, 2024 (Lomas, 2024; Long et al., 2024; Nahra et al., 2024).

The Act uses a four-tier, risk-based approach to AI systems, divided into those with minimal, limited, high, and unacceptable risk. The degree of regulation is influenced by the level of risk. Minimal risk AI is largely unregulated but developers and deployers "are encouraged to adhere to voluntary codes of conduct which would follow some of the high-risk AI systems requirements" (Garrod et al., 2024, "Minimal risk" section). Limited risk AI has light-touch transparency requirements. High-risk AI has strong safeguards. Unacceptable risk AI is prohibited. Under the Act, different obligations apply to different operators,[25] including providers[26] and deployers,[27] based on risk level and the impact on end users. For example, in high-risk AI systems, the obligations fall primarily on providers:

- Those that intend to place on the market or put into service high-risk AI systems in the EU, regardless of whether they are based in the EU or a third country.

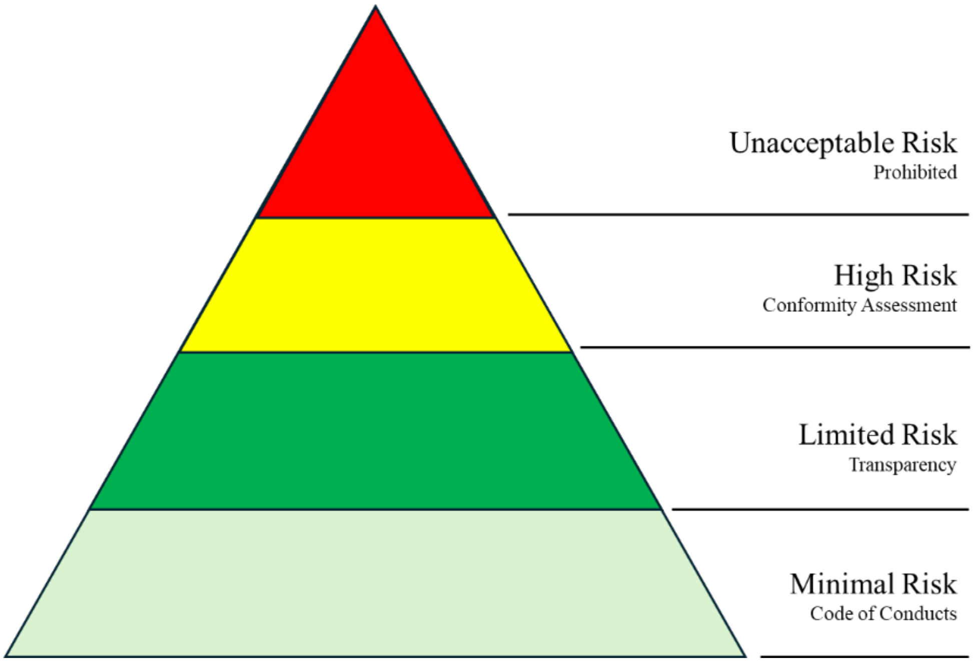
- And also third country providers where the high

[25] "Operator" is a broad term and includes "a provider, product manufacturer, deployer, authorised representative, importer or distributor" (Regulation (EU) 2024/1689, 2024, Art. 3(8)).

[26] While the EU AI Act uses the term "provider," it is substantially similar to a "developer," which is the common term the authors of this paper have employed. That said, a "provider" is

> a natural or legal person, public authority, agency or other body that develops an AI system or a general-purpose AI model or that has an AI system or a general-purpose AI model developed and places it on the market or puts the AI system into service under its own name or trademark, whether for payment or free of charge. (Regulation (EU) 2024/1689, 2024, Art. 3(3))

[27] A "deployer" is "a natural or legal person, public authority, agency or other body using an AI system under its authority except where the AI system is used in the course of a personal non-professional activity" (Regulation (EU) 2024/1689, 2024, Art. 3(4)).



Note. Chart reproduced from *The EU AI Act: What It Means for Your Business*, by K. Meier & R. Spichiger, Ernst & Young Limited, March 15, 2024 (https://www.ey.com/en_ch/forensic-integrity-services/the-eu-ai-act-what-it-means-for-your-business).

> risk AI system's output is used in the EU. (Future of Life Institute, 2024, "Four-point summary" section)

The Act also places certain privacy, safety, and transparency requirements on General Purpose Artificial Intelligence (GPAI).[28] GPAI includes many of the foundational, generative AI models in vogue, such as ChatGPT, DALL-E, Llama, Gemini, and the like. They are "characterised by their scale (a lot of memory, data and powerful hardware) as well as their reliance on transfer learning (applying knowledge from one task to another)" (Future of Life Institute, 2022, p. 3). Furthermore,

> a single general purpose AI system for language processing can be used as the foundation for several hundred applied models (e.g. chatbots, ad generation, decision assistants, spambots, translation, etc.), some of which can then be further fine-tuned into a number of applications. (p. 3)

The Act requires GPAI providers to "provide technical documentation, instructions for use, comply with the Copyright Directive, and publish a summary about the content used for training" (Future of Life Institute, 2024, "Four-point Summary" section). Providers of GPAI that includes a "systemic risk[29]—open or closed—must also conduct model evaluations, adversarial testing, track and report serious incidents and ensure cybersecurity protections" ("Four-point Summary" section).

[28] GPAI is

> an AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market. (Regulation (EU) 2024/1689, 2024, Art. 3(63))

[29] "Systemic risk" is

> a risk that is specific to the high-impact capabilities of general-purpose AI models, having a significant impact on the Union market due to their reach, or due to actual or reasonably foreseeable negative effects on public health, safety, public security, fundamental rights, or the society as a whole, that can be propagated at scale across the value chain. (Regulation (EU) 2024/1689, 2024, Art. 3(65))

Minimal risk AI, which is largely unregulated, is the catchall for systems that do not fall into one of the other three risk categories (McElligott, 2023, "Minimal or No Risk Systems" section). The European Commission (2024) noted that "the vast majority of AI systems currently used in the EU fall into this category," including "applications such as AI-enabled video games or spam filters" ("Minimal or no risk" section).

Limited risk AI has light-touch transparency requirements where, for example, "developers and deployers must ensure that end-users are aware that they are interacting with AI (chatbots and deepfakes)" (Future of Life Institute, 2024, "Four-point Summary" section). The European Commission (2024) further noted that

> providers also have to ensure that AI-generated content is identifiable. Besides, AI-generated text published with the purpose to inform the public on matters of public interest must be labelled as artificially generated. This also applies to audio and video content constituting deep fakes. ("Limited risk" section)

High-risk AI has strong safeguards, requiring a risk management system, sound data governance, record keeping, compliance documentation, quality management, instructions for downstream use, human oversight of the system, and system design that emphasizes accuracy and cybersecurity (Future of Life Institute, 2024, "High risk AI systems" section). Broadly, high-risk AI systems fall into the following categories:

- critical infrastructures (e.g. transport), that could put the life and health of citizens at risk

- educational or vocational training, that may determine the access to education and professional course of someone's life (e.g. scoring of exams)

- safety components of products (e.g. AI application in robot-assisted surgery)

- employment, management of workers and access to self-employment (e.g. CV-sorting software for recruitment procedures)

- essential private and public services (e.g. credit scoring denying citizens opportunity to obtain a loan)

- law enforcement that may interfere with people's fundamental rights (e.g. evaluation of the reliability of evidence)

- migration, asylum and border control management (e.g. automated examination of visa applications)

- administration of justice and democratic processes (e.g. AI solutions to search for court rulings). (European Commission, 2024, "High risk" section)

Other examples of high-risk applications include "AI deployed in medical devices, as a safety component in toys or in the management of critical infrastructure like the supply of electricity, in employment recruitment tools, credit scoring applications and grade prediction technology in education" (McElligott, 2023, "High Risk AI Systems" section).

Finally, "certain AI practices are considered to be a clear threat to fundamental rights and are prohibited" (Nahra et al., 2024, "What Is the EU Approach to AI Regulation?" section). These prohibited, unacceptable risk use cases include:

- deploying **subliminal, manipulative, or deceptive techniques** to distort behaviour and impair informed decision-making, causing significant harm.

- **exploiting vulnerabilities** related to age, disability, or socio-economic circumstances to distort behaviour, causing significant harm.

- **biometric categorisation systems** inferring sensitive attributes (race, political opinions, trade union membership, religious or philosophical beliefs, sex life, or sexual orientation), except labelling or filtering of lawfully acquired biometric datasets or when law enforcement categorises biometric data.

- **social scoring**, i.e., evaluating or classifying individuals or groups based on social behaviour or personal traits, causing detrimental or unfavourable treatment of those people.
- **assessing the risk of an individual committing criminal offenses** solely based on profiling or personality traits, except when used to augment human assessments based on objective, verifiable facts directly linked to criminal activity.
- **compiling facial recognition databases** by untargeted scraping of facial images from the internet or CCTV footage.
- **inferring emotions in workplaces or educational institutions**, except for medical or safety reasons.
- **'real-time' remote biometric identification (RBI) in publicly accessible spaces for law enforcement**, except when:
  - searching for missing persons, abduction victims, and people who have been human trafficked or sexually exploited;
  - preventing substantial and imminent threat to life, or foreseeable terrorist attack; or
  - identifying suspects in serious crimes (e.g., murder, rape, armed robbery, narcotic and illegal weapons trafficking, organised crime, and environmental crime, etc.). (Future of Life Institute, 2024, "Prohibited AI systems" section; Regulation (EU) 2024/1689, 2024, Art. 5)

In addition to the above categories, specific examples of unacceptable AI risk include "voice-activated toys that encourage dangerous behaviour in children," "social scoring by governments that might lead to discrimination," and "predictive policing" (European Parliament, 2024, "Unacceptable risk" section; McNally, 2023, "Unacceptable risk" section; Townson et al., 2024, "Risk Severity and Governance Requirements" section).

It is important to note that there are several exemptions from the EU AI Act, including AI developed for the military, defense, national security, and scientific research purposes, as well as "free and open-source AI, in which the code is in the public domain and available for anyone to use, modify, and distribute" (Butler, 2024, "Building an AI program" section).

Furthermore, according to Long et al. (2024), "the AI Act is mainly enforced at an EU Member State national level, with the exception of general-purpose AI models, which are enforced by the European AI Office" ("Enforcement" section). Compliance deadlines are also staggered. The AI Act went into force on August 1, 2024, and applies the following compliance deadlines: six months for prohibited AI systems, 12 months for GPAI, 24–36 months for high risk AI systems, and codes of practice established within nine months (Future of Life Institute, 2024, "Timelines" section). Finally, penalties are steep, "rang[ing] from €7.5 million to €35 million or 1% to 7%" of global annual revenue, whichever is greater, "depending on the severity of the infringement" (Meier & Spichiger, 2024, "What are the penalties in case of non-compliance?" section; Jain, 2024, "AI Act Basics" section).[30]

POLICY CONSIDERATIONS

The Foundation's guiding principle is that technology is a tool that should serve humanity, not the other way around. The moral, policy , and legal questions in the design and use of artificial intelligence

[30] For context, using the conversion rate as of October 26, 2024, enumerated fines range from $8,130,742.50 to $37,943,465 (or, as noted above, 1%–7% of global annual revenue, whichever is greater) (Mastercard, n.d.).

present a critical challenge not just to the state and nation but to humanity as a whole. It is incumbent upon lawmakers, policy experts, industry, and other key stakeholders to lay a sound policy foundation upon which to balance competing interests such as respecting privacy and human dignity with responsible technological advancements that unleash human flourishing. Indeed, light-touch regulatory certainty in the early stages of this AI boom will help foster an environment for investment, innovation, and job creation here in Texas.

To further this broader guiding principle and flesh out a policy framework, the following section first outlines why Texas should be the national leader in responsible AI policy. Second, it discusses policy considerations for responsible guardrails ungirded by values such as human dignity, privacy, transparency, and accountability. Third, it considers questions of applicability and enforcement. Finally, it assesses use cases within a risk-based framework, highlighting those that should receive additional scrutiny and those that should generally not be allowed.

The Why: Texas Should be the National Leader in Responsible AI Policy

According to the National Conference of State Legislatures (2024a; 2024b), at least 27 states and territories in 2023 and at least 43 states and territories in 2024 introduced AI-related bills. As discussed above in the state regulatory landscape section, some were in the form of study committees, others targeted to specific use cases like deepfakes or elections, and a few have attempted a more comprehensive framework. Because of its size, economy, job market, educational environment, and influence, Texas is positioned to be a leader in substantive AI policy among the states.

As with other technology policy issues, the Foundation advocates for a state-based solution, which preempts local regulations of AI. Industry will predictably argue for no local or state regulation or minimal federal regulation to avoid a patchwork of laws. While uniformity is preferable, industry's arguments are unpersuasive (as they were on data privacy and kids' online safety legislation). Texas need not wait for an ineffectual Congress to act—as it has failed to on comprehensive data privacy, kids' online safety, and AI reforms, for example. This issue is too important and timely to be neglected; Texas must act in the 89th Legislature. By enacting a strong, comprehensive, consumer protection law, Texas' law can become the model for sister states to follow, accomplishing the mutually desired goal of uniformity.

The Texas Legislature has numerous committees, councils, and stakeholder working groups studying AI in the lead up to the 2025 session, including the AI Advisory Council pursuant to HB 2060 (2023), the House Select Committee on Artificial Intelligence & Emerging Technologies, the Senate interim charges in three committees, and several AI stakeholder groups. The work of these groups will produce not only a comprehensive AI bill but also several targeted AI bills, underscoring the thoughtful process by which Texas lawmakers will act.

The How: Principles and Policy Considerations

Proviso: Definitions

It is important to note at the outset that legislation requires more than principles; it requires precise language. That is, an AI bill cannot contain abstractions but requires carefully crafted terms. Principles can be the foundation; definitions are the building blocks. Besides the worldview differences on how AI should be regulated, it is the definitions that will cause a major sticking point as AI legislation moves through the legislative process. Throughout this paper as well as in Appendix A,[31] the authors have endeavored to identify the key terms, and, where

[31] Appendix A (page 122) is meant as a resource guide to ensure lawmakers are apprised of essential terms in the realm of AI technology and regulation for legislative scope, precision, applicability, and enforceability. The authors first consulted the Texas code for relevant definitions; second, legislation and laws in other jurisdictions; and finally, terms as used by researchers, industry, and other key stakeholders.

appropriate, to include sample definitions from other legislation, industry, academic research, and other stakeholders. The following sections attempt to lay a principled foundation and identify definitional building blocks for a sound AI policy structure.

Human Dignity

Because human dignity is the overarching and animating principle of our AI policy considerations, it is worth sitting a moment with its importance. The Foundation has consistently maintained that technology is a tool which can be used for good or for ill, and it must always be used to serve humanity, not the other way around. AI is the latest technology that, like few technologies before it, raises serious challenges for humanity, and even existential questions about what it means to be human. It is for these very reasons that credible actors in the realms of innovation, technology, computer science, and other affiliated fields have made unflinching assertions that AI could eradicate humankind as we know it. The severity of this technology is palpable, and while the concerns should be taken into account while crafting public policy, emotions on either extreme of the AI spectrum—doom and gloom versus transhumanistic idealism—should not determine the guiding ethos of the public policy process.

This is not conjecture or hyperbole. Under Texas law, the definition of an artificial intelligence system is a system capable of "perceiving an environment through data acquisition and processing and interpreting the derived information to take an action or actions or to imitate intelligent behavior given a specific goal" and of "learning and adapting behavior by analyzing how the environment is affected by prior actions" (HB 2060, 2023, Sec. 2054.621). This legal definition recognizes core metaphysical and epistemological questions about AI technology. It further contemplates emerging and even hypothetical technologies beyond ubiquitous foundational, generative AI models like ChatGPT, including artificial general intelligence and superintelligence.

While a "hypothetical technology," artificial general intelligence (AGI) is "where artificial machine intelligence achieves human-level learning, perception and cognitive flexibility" (Mucci & Stryker, 2024, paras. 2–3). OpenAI is trying to move beyond the hypothetical, however, arguing that its "mission is to ensure that artificial general intelligence—AI systems that are generally smarter than humans—benefits all of humanity" (OpenAI, 2023, para. 1). For the authors, the noteworthy clause of that sentence is the desire to create "AI systems that are generally smarter than humans." Artificial superintelligence (ASI) would go even further, to the extent of having "cutting-edge cognitive functions and highly developed thinking skills more advanced than *any* [emphasis added] human" allowing it to surpass "human intelligence across all fields" and "capable of outperforming the best human minds in every domain" (Mucci & Stryker, 2023, para. 1; USC Libraries, n.d., para. 1).

Indeed, these forms of AI could present a serious risk to humanity. This underscores the importance of values-driven, responsible technological innovation that is pro-humanity at the core. This is necessary not merely in the future but *now*. This includes data integrity and the design, deployment, optimization, and use of AI tools. A human-centered approach recognizes that technology should never manipulate free will, surpass human intelligence, nor take away human agency—even if that means being accused of specism.[32]

Using AI to augment human efforts and humanity is good. Allowing it to surpass, replace, or defeat humanity is, in a word, bad. The authors firmly believe that, given all the advancements AI will precipitate, it

[32] According to Walter Isaacson's (2023) biography, *Elon Musk*, in 2013

> Musk argued that unless we built in safeguards, artificial intelligence systems might replace humans, making our species irrelevant or even extinct.
>
> [Then-Google CEO Larry] Page pushed back. Why would it matter, he asked, if machines someday surpassed humans in intelligence, even consciousness? It would simply be the next stage of evolution.
>
> Human consciousness, Musk retorted, was a precious flicker of light in the universe, and we should not let it be extinguished. Page considered that sentimental nonsense. … He accused Musk of being a "specist," someone who was biased in favor of their own species. "Well, yes, I am pro-human," Musk responded. (p. 241)

is unacceptable to realize any outcome other than one where AI only *furthers* human dignity and agency.

A human-centric, risk-based approach contemplates that some use cases should not be allowed or, perhaps, even developed or deployed at all. This approach requires respect for privacy, civil liberties, and constitutional rights. It requires protection against exploitation and discrimination. It protects against the misuse of granular human data such as facial and emotional recognition and biometrics and the data of vulnerable populations like minors. And it fights back against epistemic challenges, the erosion of truth, and threats to democratic institutions.

The authors have made this point at length in previous research publications, but it bears repeating (Whiting, 2023a; Dunmoyer & Whiting, 2022). With so many other recent advancements in emergent technology—whether social media, smart devices, or biometrics—unfettered innovation in the near term came at the expense of demonstrable consumer harm. Consequently, this has required legislators to redress grievances—like data privacy, kids' online safety, or cybersecurity requirements for critical information—after irreconcilable harm befell consumers. Furthermore, when businesses grow in an unfettered manner that allows for the chipping away of human dignity, privacy, transparency, and other key principles, it creates momentous headwinds that work against public policy solutions that will be dismissed as "too little, too late."

Privacy

The Foundation has long advocated for strong consumer and individual protections in technology policy. Accordingly, any AI bill should be the strongest consumer protection bill in the nation. Emphasis on data privacy has been a cornerstone of many technology bills in recent sessions. Because AI is so data dependent, data privacy must—once again—be the cornerstone of responsible AI policy. First, AI bills must emphasize data privacy and cybersecurity protections at all stages. As AI takes on various forms and unique applications arise through innovation; accordingly, data minimization, security, and cyber hygiene should be primary legislative considerations. Second, AI bills should include digital rights like those in the landmark TDPSA, including the rights of citizens to know, access, correct, delete, and opt-out of their data being used to train AI models, as well as an appeal process (HB 4, 2023, Sec. 541.051). Third, bills should incorporate the principles of HB 4 and HB 18 (2023) with additional protections relating to the collection and use of data from minors and limiting exposure to harmful content. Finally, AI developers and deployers should also be subject to a TDPSA data protection assessment-like requirement (HB 4, 2023, Sec, 541.105).

Transparency

Transparency requirements should be a key metric by which consumer protections are measured. Texans deserve to know when they are interacting with AI, full stop. To foster public trust and familiarity with emerging technologies, AI deployment should be made evident to the user, conspicuously, through some form of digital identification and notice at the very least. Even better is a notice followed by an affirmative consent of the user. And even better than that is developing explainable AI systems.

Importantly, this notice and consent process should be conspicuous, explainable, streamlined, and not onerous. First, it should be prominently displayed or otherwise easily navigable. Second, it should be written in clear, easily understandable language and not unnecessarily obscured by technobabble that the average user cannot understand (Kaling & Heckerling, 2005, 5:50). Finally, it should be a streamlined process that is not onerous for the user. In the data privacy context, the authors have written elsewhere that "it takes a total of 10 clicks and a proficient ability to navigate 'unclear, incomplete, and misleading' information to set up a more privacy-friendly option" (Dunmoyer & Whiting, 2022, p. 5; Lomas, 2022, para. 5). It should not take users that many clicks to know they are interacting with an AI system, consent to the interaction, or to exercise their digital rights.

Accountability and Compliance

Transparency is important in itself, but accountability helps concretize it. The goal of accountability measures is compliance, not punishment. Accordingly, Texas lawmakers should find ways to simplify compliance for good actors while carrying a big enough stick to dissuade bad actors. Here are some suggestions.

First, establish a rebuttable presumption for compliance if a developer or deployer adheres to certain standards, whether developed by industry or through entities like NIST or the OECD, for example. Second, establish a data protection, cyber hygiene, or impact assessment requirement similar to the specifications in the TDPSA (HB 4, 2023, Sec. 541.105). Third, to measure accountability, AI systems should be subject to an audit and report process. The gold standard is the use of third-party audits to enhance transparency. These audits should be accompanied by an independent reporting system that facilitates the submission of complaints and the oversight and prioritization of the audit process. Fourth, establish a broader human-driven appeal process for consumers for alleged infringements of digital rights (Sec. 541.053). Fifth, establish a cure process for developers and deployers (Section 541.154). Sixth, streamline the enforcement process in a single entity. The authors suggest utilizing the deceptive trade practices framework with enforcement through the consumer protection division of the Office of the Attorney General, as Texas has done in other recent technology bills. Ultimately, these accountability measures exist to incentivize good actors and responsible development and deployment while also insulating the process from immediate enforcement action, thereby freeing up resources to focus on the most egregious violations.

Such transparency and accountability requirements have been lamented as overly burdensome by industry in other contexts, notably data privacy and kids' online safety debates. Care must be taken by lawmakers to ensure the compliance burdens are not undue, but such measures are not unfamiliar to industry, which must comply with similar data privacy measures in other contexts (Whiting, 2023c).

Applicability, Liability, and Enforcement

As with other technology bills, questions of applicability, liability, and enforcement are often sticking points. Taking a few steps back, however, a frequent inquiry throughout this paper is whether existing laws already apply, negating the need for further lawmaking. As Rebecca Engrav, a privacy and data security attorney, argued, "what seems to get missed … in some of the public discourse these days about AI is that AI is already regulated by existing laws" (Krietzberg, 2023, para. 7). To that point, others argued that existing laws on fraud, defamation, deceptive trade practices, products liability, discrimination, intellectual property, and more already apply to AI technologies (McSherry, 2024; Smith, 2024; Mims, 2024). Furthermore, federal agencies such as the FTC, Department of Justice, Equal Employment Opportunity Commission, Consumer Financial Protection Bureau, Copyright Office, U.S. Patent and Trademark Office, Department of Labor, and others have noted existing laws already apply to AI, automated systems, and other emerging technologies (National Telecommunications and Information Administration, 2024, "AI Accountability Chain" section).

The authors of this paper question whether the AI legal landscape is at once so comprehensive and clean cut. Indeed, even if conceding that the above arguments are true, this leaves a tangled, even unworkable mess and patchwork of state versus federal laws, criminal versus civil laws, agency versus judicial enforcement, and the like. This raises more questions than it provides answers. What existing areas of the law govern AI? Are there any areas of the law to which AI does not apply? How is liability apportioned? Who enforces violations? What are the penalties, fines, or even jail time for violations? And, critically, should lawmakers address these questions or just leave them to the courts to decide?

To answer the last question first, Texas lawmakers should be much clearer. There are several ways to accomplish this. First, lawmakers could add

A Texas model should identify and place guardrails around general purpose AI and high-risk use cases, prohibit unacceptable use cases, and encourage all developers and deployers to adhere to industry standards or voluntary codes of conduct on privacy, transparency, and safety.

appropriate definitions of AI and similar terms to existing laws to expressly ensure that AI content is covered. For those who argued above that AI is already covered by existing law, an express statutory addition to a fraud, election, or CSAM statute should not be problematic. However, while there may be some utility to this, for the authors it is a less desirable approach, it risks missing areas of the law, and it still leaves a confusing enforcement patchwork. Instead, Texas should proceed with its comprehensive framework to lay out the principles that apply to the responsible development and use of AI and house enforcement under one entity, as the legislature did with other data privacy and kids' online safety bills in previous sessions.

Second, as noted above, it is important to determine to whom obligations apply, at what stage of development or deployment, and how to assess liability for violations. An important question for lawmakers to consider is to whom transparency and accountability should apply: to all businesses, some businesses, or big business but not startups or small business, for example. As mentioned above, compliance is the ultimate goal of any AI regulation, not punishment. Accordingly, lawmakers should consider a rebuttal presumption of reasonable care for complying with certain standards.

Finally, on the question of enforcement, lawmakers should avoid private rights of action which could create a court-clogging amount of litigation that flies in the face of decades of tort reform efforts in Texas. Rather, enforcing AI legislation with strong financial penalties through the deceptive trade practices framework makes the most sense. It is well-established, clear, and consistent with several key technology laws enacted in previous sessions. Critically, Texas already has the enforcement infrastructure in place. In June 2024, the attorney general's office created a dedicated data privacy and security team within the consumer protection division, which could naturally house AI enforcement (Office of the Attorney General, 2024).

A Risk-Based Approach

A risk-based approach provides a framework, adopted in other jurisdictions, on which to build substantive AI policy. As an initial matter, that the EU and Colorado utilize a risk-based approach is not an endorsement of their laws. However, they provide useful case studies and should be taken into consideration by Texas lawmakers.

Recall that, in a risk-based framework, the degree of regulation is influenced by the level of risk, which is "the combination of the probability of an occurrence of harm and the severity of that harm" (Regulation (EU) 2024/1689, 2024, Art. 3(2)). For example, the EU AI Act uses four-tiers—minimal, limited, high, and unacceptable risk. While the Colorado law focuses primarily on high-risk, the EU law also considers general purpose AI and unacceptable use cases. Below are some considerations and challenges lawmakers face in fleshing out a risk-based framework.

First, a Texas model should identify and place guardrails around general purpose AI and high-risk use cases, prohibit unacceptable use cases, and encourage all developers and deployers to adhere to industry standards or voluntary codes of conduct on privacy, transparency, and safety. While there may be philosophical opposition to a risk-based model broadly, the primary practical conflict among stakeholders will likely be how to define general purpose, high-risk, and unacceptable and what use cases constitute each category. For scholarly guidance on how to define these use cases, MIT researchers created "a comprehensive living database of over 700 AI risks categorized by their cause and risk domain" (AI Risk Repository, n.d., "What are the risks from Artificial Intelligence" section; Mulligan, 2024).

As noted above, general purpose AI includes many of the foundational, generative AI models in vogue, such as ChatGPT, DALL-E, Llama, Gemini, and the like. Because of their ubiquitous use, broad application, massive data requirements, cybersecurity risks, and hallucinations, some minimum safeguards should be put in place. The EU AI Act places certain privacy, safety, and transparency requirements on GPAI. Furthermore, NIST's (2023a) AI RMF 1.0, the GAI supplement (NIST, 2024a), and other rules and guidance may provide helpful state-based standards and metrics for AI systems (U.S. Department of Commerce, 2024; NIST, 2024b).

Additionally, Texas lawmakers must consider high-risk AI systems. The Colorado law, for example, defines a high-risk system as "any artificial intelligence system that, when deployed, makes, or is a substantial factor in making, a consequential decision" (SB 24-205, 2024, Sec. 6-1-1701(9)(a)). Under the Colorado law, a "consequential decision" is "a decision that has a material legal or similarly significant effect on the provision or denial to any consumer of, or the cost or terms of" educational, employment, financial, government, healthcare, housing, insurance, or legal services (SB 24-205, 2024, Sec. 6-1-1701(3)). The TDPSA includes a similar definition of a "decision that produces a legal or similarly significant effect concerning a consumer" and adds criminal justice and "access to basic necessities, such as food and water" (HB 4, 2023, Sec. 541.001(11)). The EU AI Act includes a lengthy list of high-risk use cases (discussed at length above in the "International" regulation section), which adds use cases such as critical infrastructure, product safety, immigration and border security, and the justice system (European Commission, 2024, "High risk" section).

Second, beyond the challenges of defining and identifying GPAI and high-risk use cases, another major sticking point will be determining what obligations apply, to whom they apply, at what stage of development or deployment they apply, and how to address violations. For example, the Software Alliance argued for delineating "a distinction between developers of AI and those entities putting AI systems into use ('deployers'), along with clear obligations for both" (BSA, 2023, para. 7). This is critical in a Texas bill. Under the EU AI Act, different obligations apply to different operators, including providers and users, based on risk level and the impact on end users . Under the EU framework, for example, the obligations fall primarily on providers in high-risk AI systems. Likewise, the Colorado law places different obligations on developers and deployers. According to Kohne et al. (2024), "the law will require developers of high-risk AI systems to use 'reasonable care' to protect consumers from any known or foreseeable risks of algorithmic discrimination resulting from the intended and contracted uses of those high-risk AI systems" ("Requirements for AI Developers" section). Furthermore, "the law will require deployers of high-risk AI systems to protect consumers from any known or reasonably foreseeable risks of algorithmic discrimination" ("Requirements for AI Deployers" section).

Texas lawmakers should carefully consider these obligations. Specifically, developers of high-risk AI systems should provide a statement of intended uses by deployers, have responsibilities to downstream deployers, provide a centralized system for reporting harms discovered in development or deployment, and may submit impact assessments to regulators on behalf of deployers. Furthermore, deployers of high-risk AI systems should be required to have a standardized risk management policy and program, provide consumers notice and appeal procedures, conduct impact assessments, and maintain a harm reporting process.

Third, various AI measures place some reasonable privacy, transparency, reporting, and accountability requirements as well as some that are quite burdensome. Some reasonable baseline requirements include conspicuous notice that a user is interacting

It is critical that Texas lawmakers do not adopt some of the overly burdensome requirements of the EU and California SB 1047 (2024), which would stifle innovation.

with an AI system, receiving affirmative consent from a user to interact, data minimization, additional safeguards on sensitive data, or even opting out of personal data being used to train the model. Other considerations include how to measure performance and accuracy.

In turn, it is critical that Texas lawmakers do not adopt some of the overly burdensome requirements of the EU and California SB 1047 (2024), which would stifle innovation. For example, these approaches have overly burdensome disclosure and reporting requirements on AI systems *before* initiating training or putting it on the market. In opposing an early iteration of California SB 1047, Anthropic "suggest[ed] the bill shift to 'outcome-based deterrence' from 'pre-harm enforcement,' letting AI companies develop and deploy safety protocols and be held liable for any catastrophes they cause" (Gold, 2024, para. 6). This is a careful balancing act that Texas lawmakers have done before in areas like data privacy and kids' online safety and must do again here with AI regulation.

Fourth, exemptions are less desirable but are realistically part of the AI stakeholder and legislative process. The most likely areas for exemptions are certain categories of business and data as well as specific use cases. Categorically, previous Texas data privacy measures have exempted state agencies; political subdivisions; educational institutions; small businesses as defined by the Small Business Administration; businesses based on number of employees, revenue, or percentage of data collection and processing activities; and entities with additional data requirements under federal health, financial, and educational laws, for example. Other measures, including the EU law, exempt AI systems used for military, defense, national security, and research purposes (Garrod et al., 2024). Lawmakers should also consider regulatory sandboxes. As to specific use cases, the EU, Colorado, and New York City laws provide a list of exemptions including procedural tasks, pattern detection, anti-fraud, anti-virus, video games, calculators, cybersecurity tools, data storage, firewalls, spam filters, spell-check, and even spreadsheets (Regulation (EU) 2024/1689, 2024, Art. 6(3); SB 24-205, 2024, Sec. 6-1-1701(9)(b); Local Law 2021/144, 2021, Sec. 20-870). Where these technologies do not have an undue material impact on individuals, these are reasonable use cases to exempt.

Finally, there are some use cases of AI that should never be allowed (certainly not at this time). Lawmakers in 2025 are right to restrain unfettered innovation in areas where fundamental human rights and dignity are at risk. Examples of unacceptable use cases include manipulation of human behavior to circumvent free will, exploiting vulnerabilities based on personal characteristics or protected classes, social credit scores, untargeted facial recognition, certain predictive analytics, emotional recognition in schools or the workplace, biometric categorization, and child sexual abuse material. It may be a leap at this time, but lawmakers may (or may not) consider exemptions for certain public safety and law enforcement exigencies and use cases such as active commission of a felony, missing persons, human trafficking, contraband smuggling, and other serious crimes. Unless the technology undergirding these use cases can be deployed in responsible, pro-human ways, then, and only then, should lawmakers consider allowing them to be used in Texas.

CONCLUSION

Generative AI systems like ChatGPT have firmly entered the zeitgeist. And while AI is ubiquitous—as it seems to be everywhere and in all things these days—it is not a new technology, tracing its history back to the era of World War II Nazi codebreaking. Today, AI is a technology with great promise and peril. From healthcare to education, employment to financial services, and national security to the justice system, AI has the potential to revolutionarily enhance human flourishing and potential. At the same time, AI presents critical challenges to privacy, security, the workforce, knowledge, trust, institutions, and humanity itself. Indeed, Neil Postman's ecology of technology theory applied to AI rings true: AI does not add to or subtract from the ecosystem, it fundamentally transforms it in its entirety.

Texas is fortunate to have all the ingredients to be a world leader in the development, deployment, and use of AI technologies. Because of its regulatory environment, infrastructure, capital investments, and more, Texas' dominance in the emerging technology space increases daily. For these reasons, along with the legislative thought leadership the state of Texas brings to the public policy arena globally, it is incumbent upon Texas lawmakers to carefully consider the necessary policy landscape within which this technology should be responsibly developed, deployed, and used. Specifically, this should occur within a risk-based framework emphasizing values such as privacy, transparency, accountability and human dignity. Indeed, Texas has the legislative leadership and thoughtful interim and stakeholder processes to be a national leader in crafting sound policy solutions that will serve as a model for sister states—and perhaps even Congress and the rest of the world—to follow. ■

REFERENCES

AI Elections Accord. (2024a, February 16). *Technology industry to combat deceptive use of AI in 2024 elections* [Press release]. https://www.aielectionsaccord.com/uploads/2024/02/Press-Release-AI-Elections-Accord-16-Feb-2024.pdf

AI Elections Accord. (2024b, February 16). *A tech accord to combat deceptive use of AI in 2024 elections*. https://www.aielectionsaccord.com/uploads/2024/02/A-Tech-Accord-to-Combat-Deceptive-Use-of-AI-in-2024-Elections.FINAL_.pdf

AI Risk Repository. (n.d). *What are the risks from artificial intelligence?* MIT. Retrieved August 22, 2024, from https://airisk.mit.edu/

Aich, S., & Burch, G. F. (2023, January 1). *Looking inside the magical black box: A systems theory guide to managing AI*. ISACA. https://www.isaca.org/resources/isaca-journal/issues/2023/volume-1/looking-inside-the-magical-black-box

AIWS. (2021, April 16). *This week in the history of AI at AIWS.net – MIT receives a $2.2 million grant in June 1963 from DARPA*. https://aiws.net/the-history-of-ai/this-week-in-the-history-of-ai-at-aiws-net-mit-receives-a-2-2-million-grant-in-june-1963-from-darpa/

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). *Molecular biology of the cell (4th edition)*. Garland Science. https://www.ncbi.nlm.nih.gov/books/NBK21054/

Allen, G. C. (2022). *Choking off China's access to the future of AI*. Center for Strategic & International Studies. https://csis-website-prod.s3.amazonaws.com/s3fs-public/2023-04/221011_Allen_China_AccesstoAI.pdf

Alper, A. (2024, May 8). *Exclusive: US eyes curbs on China's access to AI software behind apps like ChatGPT*. Reuters. https://www.reuters.com/technology/us-eyes-curbs-chinas-access-ai-software-behind-apps-like-chatgpt-2024-05-08/

American Civil Liberties Union. (2021, May 18). *ACLU statement on extended Amazon face recognition moratorium* [Press release]. https://www.aclu.org/press-releases/aclu-statement-extended-amazon-face-recognition-moratorium

American College of Radiology. (n.d.). *Jobs*. Retrieved August 19, 2024, from https://jobs.acr.org/jobs/

Amod, F. (2024, July 11). *What is voice cloning?* Paubox. https://www.paubox.com/blog/what-is-voice-cloning

Anyoha, R. (2017, August 28). *The history of artificial intelligence*. Science in the News. https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/

Arcesati, R., & Creemers, R. (2024, April 2). *Chinese assessments of AI: Risks and mitigation strategies*. Center for Strategic & International Studies. https://interpret.csis.org/chinese-assessments-of-ai-risks-and-mitigation-strategies/

Arm. (n.d.). *What is voice recognition?* Retrieved August 30, 2024, from https://www.arm.com/glossary/voice-recognition

Armitage, H. (2018, November 20). *Artificial intelligence rivals radiologists in screening x-rays for certain diseases*. Stanford Medicine. https://med.stanford.edu/news/all-news/2018/11/ai-outperformed-radiologists-in-screening-x-rays-for-certain-diseases.html

Artificial Intelligence Policy Institute. (n.d.). *Vast majority of US voters of all political affiliations support President Biden's executive order on AI*. Retrieved October 29, 2024, from https://theaipi.org/poll-biden-ai-executive-order-10-30/

Artist Rights Alliance. (2024, April 1). *200+ artists urge tech platforms: Stop devaluing music*. https://artistrightsnow.medium.com/200-artists-urge-tech-platforms-stop-devaluing-music-559fb109bbac

Arundel, K. (2023, December 12). *How are high schoolers using AI?* K-12 Dive. https://www.k12dive.com/news/artificial-intelligence-high-school-students/702228/

Association of Southeast Asian Nations. (n.d.). *ASEAN member states*. Retrieved August 21, 2024, from https://asean.org/member-states/

Aswad, J. (2024, April 2). Billie Eilish, Nicki Minaj, Stevie Wonder, dozens more call on AI developers to respect artists' rights. *Variety*. https://variety.com/2024/music/news/billie-eilish-nicki-minaj-ai-respect-artists-rights-1235957451/

Autonomous Weapons. (n.d.). *Slaughterbots are here*. Retrieved August 19, 2024, from https://autonomousweapons.org/

Awati, R., & Yasar, K. (2024). *Black box AI*. TechTarget. https://www.techtarget.com/whatis/definition/black-box-AI

AWS. (n.d.-a). *What is structured data?* Retrieved August 19, 2024, from https://aws.amazon.com/what-is/structured-data/

AWS. (n.d.-b). *What are foundational models?* Retrieved August 19, 2024, from https://aws.amazon.com/what-is/foundation-models/

AWS. (n.d.-c). *What are large language models (LLM)?* Retrieved August 30, 2024, from https://aws.amazon.com/what-is/large-language-model/

AWS. (n.d.-d). *What is a neural network?* Retrieved August 30, 2024, from https://www.ibm.com/topics/neural-networks

Ayoub, E., & Goitein, E. (2024, February 13). *Closing the data broker loophole*. Brennan Center for Justice. https://www.brennancenter.org/our-work/research-reports/closing-data-broker-loophole

Bagchi, S. (2023, May 22). *What is a black box? A computer scientist explains what it means when the inner workings of AIs are hidden*. The Conversation. https://theconversation.com/what-is-a-black-box-a-computer-scientist-explains-what-it-means-when-the-inner-workings-of-ais-are-hidden-203888

Bank of America. (2023, July 13). *BofA's Erica surpasses 1.5 billion client interactions, totaling more than 10 million hours of conversations* [Press release]. https://newsroom.bankofamerica.com/content/newsroom/press-releases/2023/07/bofa-s-erica-surpasses-1-5-billion-client-interactions--totaling.html

Bannon, L. (2023, June 15). When AI overrules the nurses caring for you. *The Wall Street Journal*. https://www.wsj.com/articles/ai-medical-diagnosis-nurses-f881b0fe

Barker, E. (2020). *Guideline for using cryptographic standards in the federal government: Cryptographic mechanisms*. National Institute of Standards and Technology. https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-175Br1.pdf

Barrett, D., Protess, B., & Haberman, M. (2024, October 29). Trump family members and Biden aides among China hack targets. *The New York Times*. https://www.nytimes.com/2024/10/29/us/politics/trump-biden-hacking-china.html

Bates, A-J. (2020). *Investigating the ASL interpreter shortage for legal settings*. Southern Methodist University. https://www.smu.edu/-/media/site/provost/saes/academic-enrichment/engagedlearning/ugr/research-week/research-days-2020/bates---research-days-2020.pdf

Bethea, C. (2024, March 7). The terrifying A.I. scam that uses your loved one's voice. *The New Yorker*. https://www.newyorker.com/science/annals-of-artificial-intelligence/the-terrifying-ai-scam-that-uses-your-loved-ones-voice

Bhuiyan, J. (2023, September 7). Lost in AI translation: Growing reliance on language apps jeopardizes some asylum applications. *The Guardian*. https://www.theguardian.com/us-news/2023/sep/07/asylum-seekers-ai-translation-apps

Bieser, J. (2022). *Creative through AI: How artificial intelligence can support the development of new ideas*. Gottlieb Duttweiler Institute. http://doi.org/10.59986/CCHA2271

Blinder, A., & Perlroth, N. (2018, March 29). Hard choice for cities under cyberattack: Whether to pay ransom. *The New York Times*. https://www.nytimes.com/2018/03/29/us/atlanta-cyberattack-ransom.html

Blouin, L. (2023, March 6). *AI's mysterious 'black box' problem, explained*. University of Michigan-Dearborn. https://umdearborn.edu/news/ais-mysterious-black-box-problem-explained

Boak, J., & O'Brien, M. (2023, October 30). *Biden wants to move fast on AI safeguards and signs an executive order to address his concerns*. The Associated Press. https://apnews.com/article/biden-ai-artificial-intelligence-executive-order-cb86162000d894f238f28ac029005059

Bond, S. (2023, May 22). *Fake viral images of an explosion at the Pentagon were probably created by AI*. NPR. https://www.npr.org/2023/05/22/1177590231/fake-viral-images-of-an-explosion-at-the-pentagon-were-probably-created-by-ai

Booth, H. (2024, July 11). Republicans' vow to repeal Biden's AI executive order has some experts worried. *Time*. https://time.com/6996927/republicans-repeal-biden-ai-executive-order/

Bordelon, B. (2023, October 30). The politics of Biden's vast new AI order. *Politico*. https://www.politico.com/news/2023/10/30/bidens-executive-order-artificial-intelligence-00124395

Bovbjerg, M. (2023, May 24). *How AI can help organizations adapt and recover from cyberattacks*. Dark Reading. https://www.darkreading.com/cyber-risk/how-ai-can-help-organizations-adapt-and-recover-from-cyberattacks

Brennan Center for Justice. (2024, October 1). *Artificial intelligence legislation tracker*. https://www.brennancenter.org/our-work/research-reports/artificial-intelligence-legislation-tracker

Brewster, T. (2023, December 5). This AI watches millions of cars daily and tells cops if you're driving like a criminal. *Forbes*. https://www.forbes.com/sites/thomasbrewster/2023/07/17/license-plate-reader-ai-criminal/

Brewster, T. (2024, June 20). FedEx's secretive police force is helping cops build an AI car surveillance network. *Forbes*. https://www.forbes.com/sites/thomasbrewster/2024/06/19/fedex-police-help-cops-build-an-ai-car-surveillance-network/

BSA. (2023, September 27). *BSA analysis: State AI legislation surges by 440% in 2023*. https://www.bsa.org/news-events/news/bsa-analysis-state-ai-legislation-surges-by-440-in-2023

Buehler, K., Corsi, A., Weintraub, B., Jurisic, M., Siani, A., & Lerner, L. (2024, March 22). *Scaling gen AI in banking: Choosing the best operating model*. McKinsey & Company. https://www.mckinsey.com/industries/financial-services/our-insights/scaling-gen-ai-in-banking-choosing-the-best-operating-model

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research 81*, 1–15. https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

Burgess, S. (2022, March 17). *Ukraine war: Deepfake video of Zelenskyy telling Ukrainians to 'lay down arms' debunked*. Sky News. https://news.sky.com/story/ukraine-war-deepfake-video-of-zelenskyy-telling-ukrainians-to-lay-down-arms-debunked-12567789

Business Wire. (2021, March 2). *Iktos announces collaboration with Pfizer in AI for drug design* [Press release]. https://www.businesswire.com/news/home/20210302005501/en/Iktos-Announces-Collaboration-With-Pfizer-in-AI-for-Drug-Design

Butler, R. (2023, August 28). *Generative AI and the courts: Balancing efficiency and legal obligations*. Thomson Reuters. https://www.thomsonreuters.com/en-us/posts/government/generative-ai-courts/

Butler, R. (2024, June 24). *Forum: Global impact of the EU AI Act*. Thomson Reuters. https://www.thomsonreuters.com/en-us/posts/corporates/forum-eu-ai-act-impact/

Camello, M. L., Houston-Kolnik, J. D., & Planty, M. (2021). *Chatbots in the criminal justice system*. Criminal Justice Testing and Evaluation Consortium. https://cjtec.org/files/chatbots-criminal-justice

Cannestra, B. (2024, April 12). *Sunnyvale turns to AI technology to translate public meetings, saving money in the process*. Local News Matters. https://localnewsmatters.org/2024/04/12/sunnyvale-turns-to-ai-technology-to-translate-public-meetings-saving-money-in-the-process/

Capps, M. (2023, August 24). *Making AI trustworthy: Can we overcome black-box hallucinations?* TechCrunch. https://techcrunch.com/2023/08/24/making-ai-trustworthy-can-we-overcome-black-box-hallucinations/

Carbon Robotics. (n.d.). [Home page]. Retrieved August 20, 2024, from https://carbonrobotics.com/

Carmichael, M. (2023, November 14). *There's strong bipartisan support for Biden's executive order on AI*. Ipsos. https://www.ipsos.com/en-us/theres-strong-bipartisan-support-bidens-executive-order-ai

Casalino, L. P., Gans, D., Weber, R., Cea, M., Tuchovsky, A., Bishop, T. F., Miranda, Y., Frankel, B. A., Ziehler, K., Wong, M. M., & Evenson, T. B. US physician practices spend more than $15.4 billion annually to report quality measures. *Health Affairs*, *35*(3). https://doi.org/10.1377/hlthaff.2015.1258

Castro, D. (2022, October 5). *White House AI Bill of Rights is all wrong, says Center for Data Innovation*. Center for Data Innovation. https://datainnovation.org/2022/10/white-house-ai-bill-of-rights-is-all-wrong-says-center-for-data-innovation/

CAT Lab. (n.d.). *Studying how employers comply with NYC's new hiring algorithm law*. Retrieved August 21, 2024, from https://citizensandtech.org/research/2024-algorithm-transparency-law/

Chandler, S. (2020, March 9). Why deepfakes are a net positive for humanity. *Forbes*. https://www.forbes.com/sites/simonchandler/2020/03/09/why-deepfakes-are-a-net-positive-for-humanity/

Chandra, A. (2024, February 23). College admissions trends: AI, college essays and going international. *Forbes*. https://www.forbes.com/councils/forbesbusinesscouncil/2024/02/23/college-admissions-trends-ai-college-essays-and-going-international/

Chatterjee, M., & Bordelon, B. (2024, January 26). The campaign to take down the Biden AI executive order. *Politico*. https://www.politico.com/news/2024/01/25/conservatives-prepare-attack-on-bidens-ai-order-00137935

Chavez, K. (2023, October 30). *Biden releases AI red tape wishlist in new executive order*. NetChoice. https://netchoice.org/biden-releases-ai-red-tape-wishlist-in-new-executive-order/

Chee, F. Y., & Hummel, T. (2024, May 22). *Europe sets benchmark for rest of the world with landmark AI laws*. Reuters. https://www.reuters.com/world/europe/eu-countries-back-landmark-artificial-intelligence-rules-2024-05-21/

Chen, H., & Magramo, K. (2024, February 4). *Finance worker pays out $25 million after video call with deepfake 'chief financial officer.'* CNN. https://www.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html

Chilson, N., & Stout, K. (2024, June 20). *Coalition letter opposing California SB 1047*. International Center for Law & Economics. https://laweconcenter.org/resources/coalition-letter-opposing-california-sb-1047/

Choi, D., Mirbod, O., Ilodibe, U., & Kinsey, S. (2023). Understanding artificial intelligence: What it is and how it is used in agriculture: AE589, 10/2023. *EDIS*, *2023*(6). https://doi.org/10.32473/edis-ae589-2023

ChrysaLabs. (n.d.). [Technology page]. Retrieved August 20, 2024, from https://www.chrysalabs.com/technology/

Citi. (2024, June 17). *AI in finance*. https://www.citigroup.com/global/insights/citigps/ai-in-finance

City of New York. (n.d.). *Automated employment decision tools (AEDT)*. Retrieved August 21, 2024, from https://www.nyc.gov/site/dca/about/automated-employment-decision-tools.page

City of New York. (2023). *Automated employment decision tools: Frequently asked questions*. https://www.nyc.gov/assets/dca/downloads/pdf/about/DCWP-AEDT-FAQ.pdf

Clark, J. (2023, November 2). *DOD releases AI adoption strategy*. U.S. Department of Defense. https://www.defense.gov/News/News-Stories/Article/Article/3578219/dod-releases-ai-adoption-strategy/

Cole, D. (2024, October 23). *The Chinese room argument*. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/chinese-room/

Columbus, L. (2020, October 31). The state of AI adoption in financial services. *Forbes*. https://www.forbes.com/sites/louiscolumbus/2020/10/31/the-state-of-ai-adoption-in-financial-services/

CompTIA. (n.d.). *Artificial intelligence (AI) glossary: Terms & definitions for beginners*. Retrieved August 30, 2024, from https://comptiacdn.azureedge.net/webcontent/docs/default-source/research-reports/ai-glossary.pdf

CookieYes. (2024, June 13). *Opt-in vs opt-out: Differences and examples*. https://www.cookieyes.com/blog/opt-in-opt-out/

Cottrill, J. (2024, July 30). *Nearly two in five Americans turn to AI for financial management advice*. Ipsos. https://www.ipsos.com/en-us/nearly-two-five-americans-turn-ai-financial-management-advice

Coursera. (2024, March 19). *Artificial intelligence (AI) terms: A to Z glossary*. https://www.coursera.org/articles/ai-terms

Craig, L. (2023). *AI watermarking*. TechTarget. https://www.techtarget.com/searchenterpriseai/definition/AI-watermarking

Cybersecurity & Infrastructure Security Agency. (n.d.). *Critical infrastructure sectors*. U.S. Department of Homeland Security. Retrieved August 19, 2024, from https://www.cisa.gov/topics/critical-infrastructure-security-and-resilience/critical-infrastructure-sectors

Cybersecurity & Infrastructure Security Agency. (2024, January 18). *Risk in focus: Generative A.I. and the 2024 election cycle*. U.S. Department of Homeland Security. https://www.cisa.gov/sites/default/files/2024-05/Consolidated_Risk_in_Focus_Gen_AI_ElectionsV2_508c.pdf

Cybersecurity & Infrastructure Security Agency & United States Coast Guard Cyber Command. (2023). *CISA analysis: Fiscal year 2022 risk and vulnerability assessments*. https://www.cisa.gov/sites/default/files/2023-07/FY22-RVA-Analysis%20-%20Final_508c.pdf

Darrach, B. (1970). Meet Shaky, the first electronic person. *Life Magazine*, *69*(21), 58b–68. https://books.google.com/books?id=2FMEAAAAMBAJ

Davidson, N. (2023, October 30). *Map: How are state and local governments navigating AI regulation?* Government Technology. https://www.govtech.com/biz/data/how-are-state-and-local-governments-navigating-ai-regulation

Davidson, N. (2024, April 16). *ALPR audit takeaways: What we learned about policy gaps*. Government Technology. https://www.govtech.com/biz/data/alpr-audit-takeaways-what-we-learned-about-policy-gaps

Dejeux, S. (2022, September 19). *Team interpreting for magistrate courts in Texas*. American Translators Association. https://www.atanet.org/interpreting/team-interpreting-for-magistrate-courts-in-texas/

Deloitte. (2023, November 22). *Open vs. closed-source generative AI*. https://www.deloitte.com/uk/en/Industries/technology/blogs/open-vs-closed-source-generative-ai.html

Demarest, C. (2024, July 11). *One-third of U.S. military could be robotic, Milley predicts*. Axios. https://www.axios.com/2024/07/11/military-robots-technology

Densen, P. (2011). Challenges and opportunities facing medical education. *Transactions of the American Clinical and Climatological Association, 122*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3116346/pdf/tacca122000048.pdf

Dictionary.com. (n.d.). Per ardua ad astra. In *Dictinary.com dictionary*. Retrieved August 19, 2024, from https://www.dictionary.com/browse/per-ardua-ad-astra

Digi.City. (n.d.). *Definitions*. Retrieved August 20, 2024, from https://www.digi.city/smart-city-definitions

Digital Recognition Network. (n.d.). [Home page]. Retrieved October 18, 2024, from https://drndata.com/

Diliberti, M. K., Schwartz, H. L., Doan, S., Shapiro, A., Rainey, L. R., & Lake, R. J. (2024). *Using artificial intelligence tools in K-12 classrooms*. RAND. https://www.rand.org/pubs/research_reports/RRA956-21.html

DISCO. (2024, October 7). *Best AI curriculum generator for modern educators - create courses in minutes like magic!* https://www.disco.co/blog/the-best-ai-curriculum-generator-for-modern-educators-2024

Dou, E. (2024, May 13). U.S.-China talks on AI risks set to begin in Geneva. *The Washington Post*. https://www.washingtonpost.com/technology/2024/05/13/us-china-ai-talks/

Douglas, T. (2018, February 16). *Los Angeles chatbot deputized to help with police recruitment*. Government Technology. https://www.govtech.com/products/los-angeles-chatbot-deputized-to-help-with-police-recruitment.html

Dow Jones. (n.d.). *Understanding the steps of a "know your customer" process*. Retrieved August 21, 2024, from https://www.dowjones.com/professional/risk/glossary/know-your-customer/

Downey, R. (2024a, March 29). Texas lawmakers and agency leaders experiment, ponder policies for an AI future. *San Antonio Express-News*. https://www.expressnews.com/news/article/texas-ai-uses-regulations-19375570.php

Downey, R. (2024b, April 30). *False ad depicting Dade Phelan with Nancy Pelosi could inspire new anti-deepfake legislation*. The Texas Tribune. https://www.texastribune.org/2024/04/30/dade-phelan-nancy-pelosi-deep-fake/

Dunmoyer, D. (2023, September 13). *Finding our freedoms: Americans must regain agency over their lives*. The Cannon. https://thecannononline.com/finding-our-freedoms-americans-must-regain-agency-over-their-lives/

Dunmoyer, D. (2024). *Responsibly ushering in autonomous vehicles in Texas*. Texas Public Policy Foundation. https://www.texaspolicy.com/responsibly-ushering-in-autonomous-vehicles-in-texas/

Dunmoyer, D., & Whiting, Z. (2022). *Why Texas needs a digital bill of rights*. Texas Public Policy Foundation. https://www.texaspolicy.com/why-texas-needs-a-digital-bill-of-rights/

Eastwood, B. (2024, May 1). *Financial services' deliberate approach to AI*. MIT. https://mitsloan.mit.edu/ideas-made-to-matter/financial-services-deliberate-approach-to-ai

Edinger, J. (2023, November 8). *AI-enabled automation streamlines local government finance*. Government Technology. https://www.govtech.com/budget-finance/ai-enabled-automation-streamlines-local-government-finance

Edwards, B. (2015, June 26). Who needs GPS? The forgotten story of Etak's amazing 1985 car navigation system. *Fast Company*. https://www.fastcompany.com/3047828/who-needs-gps-the-forgotten-story-of-etaks-amazing-1985-car-navigation-system

Electronic Frontier Foundation. (2023, October 1). *Automated license plate readers*. https://sls.eff.org/technologies/automated-license-plate-readers-alprs

Ellery, S. (2023, March 28). *Fake photos of Pope Francis in a puffer jacket go viral, highlighting the power and peril of AI*. CBS News. https://www.cbsnews.com/news/pope-francis-puffer-jacket-fake-photos-deepfake-power-peril-of-ai/

EMARKETER. (2023, January 2). *Artificial intelligence in financial services: Applications and benefits of AI in finance*. https://www.emarketer.com/insights/ai-in-finance/

European Commission. (2024, October 14). *AI Act*. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

European Parliament. (2024, June 18). *EU AI Act: First regulation on artificial intelligence*. https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

Evans, M. (2023, September 10). *Where is all your health data going? The Google and Fitbit scandal explained*. TechRadar. https://www.techradar.com/health-fitness/fitness-trackers/google-and-fitbit-scandal-explained

Exec. Order No. 13859, 84 Fed. Reg. 3,967 (Feb. 14, 2019). https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence

Exec. Order No. 13960, 85 Fed. Reg. 78,939 (Dec. 8, 2020). https://www.federalregister.gov/documents/2020/12/08/2020-27065/promoting-the-use-of-trustworthy-artificial-intelligence-in-the-federal-government

Exec. Order No. 14083, 87 Fed. Reg. 57,369 (Sep. 20, 2022). https://www.federalregister.gov/documents/2022/09/20/2022-20450/ensuring-robust-consideration-of-evolving-national-security-risks-by-the-committee-on-foreign

Exec. Order No. 14105, 88 Fed. Reg. 54,867 (Aug. 11, 2023). https://www.federalregister.gov/documents/2023/08/11/2023-17449/addressing-united-states-investments-in-certain-national-security-technologies-and-products-in

Exec. Order No. 14110, 88 Fed. Reg. 75,191 (Nov. 1, 2023). https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence

Exscientia. (2021, May 13). *Exscientia announces second molecule created using AI from Sumitomo Dainippon Pharma collaboration to enter phase 1 clinical trial* [Press release]. https://investors.exscientia.ai/press-releases/press-release-details/2021/Exscientia-announces-second-molecule-created-using-AI-from-Sumitomo-Dainippon-Pharma-collaboration-to-enter-Phase-1-clinical-trial/default.aspx

Faguy, A., & Ray, S. (2023, September 27). Hollywood writers strike ends: Deal finalized after 148 days of work stoppage. *Forbes*. https://www.forbes.com/sites/anafaguy/2023/09/27/hollywood-writers-strike-ends-deal-finalized-after-148-day-work-stoppage/

Farrell, F. (2023, November 7). *How Russia's homegrown Lancet drone became so feared in Ukraine*. Yahoo! News. https://www.yahoo.com/news/russia-homegrown-lancet-drone-became-224320518.html

Federal Bureau of Investigation. (2023). *2023 Internet crime report*. Department of Justice. https://www.ic3.gov/Media/PDF/AnnualReport/2023_IC3Report.pdf

Federal Communications Commission. (2024). *In the matter of implications of artificial intelligence technologies on protecting consumers from unwanted robocalls and robotexts*. https://docs.fcc.gov/public/attachments/FCC-24-17A1.pdf

Federal Trade Commission. (2023, April 25). *FTC Chair Khan and officials from DOJ, CFPB and EEOC release joint statement on AI*. https://www.ftc.gov/news-events/news/press-releases/2023/04/ftc-chair-khan-officials-doj-cfpb-eeoc-release-joint-statement-ai

Field, H. (2024, February 26). *Google to relaunch Gemini AI picture generator in a 'few weeks' following mounting criticism of inaccurate images*. CNBC. https://www.cnbc.com/2024/02/26/googles-gemini-ai-picture-generator-to-relaunch-in-a-few-weeks.html

Finio, M., & Downie, A. (2023, December 8). *What is artificial intelligence (AI) in finance?* IBM. https://www.ibm.com/topics/artificial-intelligence-finance

Finklea, K. (2023). *Law enforcement use of artificial intelligence and directives in the 2023 executive order*. Congressional Research Service. https://crsreports.congress.gov/product/pdf/IN/IN12289

Firth-Butterfield, K., & Silverman, K. (2022). *Artificial intelligence and the courts: Materials for judges*. American Association for the Advancement of Science. https://doi.org/10.1126/aaas.adf0782

Fitzpatrick, D. J., Gorr, W. L., & Neill, D. B. (2019). Keeping score: Predictive analytics in policing. *Annual Review of Criminology*, *2*, 473–491. https://doi.org/10.1146/annurev-criminol-011518-024534

Forristal, L. (2023, November 7). *Waze gets a new safety feature that warns you if a road has a history of crashes*. TechCrunch. https://techcrunch.com/2023/11/07/waze-crash-history-alerts/

Fortino, A. (2023, November 2). *Embracing creativity: How AI can enhance the creative process*. NYU School of Professional Studies. https://www.sps.nyu.edu/homepage/emerging-technologies-collaborative/blog/2023/embracing-creativity-how-ai-can-enhance-the-creative-process.html

Friedland, A. (2022, October 6). *Text-to-video generators from Meta and Google, the White House's "Blueprint for an AI Bill of Rights," and IARPA's project to identify anonymous authors with AI*. Center for Security and Emerging Technology. https://cset.georgetown.edu/newsletter/october-6-2022/

Fung, B. (2023, September 13). *Bill Gates, Elon Musk and Mark Zuckerberg meeting in Washington to discuss future AI regulations*. CNN. https://www.cnn.com/2023/09/13/tech/schumer-tech-companies-ai-regulations/index.html

Futr. (n.d.). *Police*. Retrieved August 20, 2024, from https://futr.ai/police/

Future of Life Institute. (2022). *General purpose AI and the AI Act*. https://artificialintelligenceact.eu/wp-content/uploads/2022/05/General-Purpose-AI-and-the-AI-Act.pdf

Future of Life Institute. (2024, May 30). *High-level summary of the AI Act*. https://artificialintelligenceact.eu/high-level-summary/

Future of Privacy Forum. (2023). *Best practices for AI and workplace assessment technologies*. https://fpf.org/wp-content/uploads/2023/09/FPF-Best-Practices-for-AI-and-HR-Final.pdf

Garner, R. (2015). *Early popular computers, 1950-1970*. ETHW. Retrieved August 19, 2024, from https://ethw.org/Early_Popular_Computers,_1950_-_1970

Garrod, D., Arlington, J., Jamooji, J., Odubanjo, O., Kohne, N. G., Rickhoff, H. C., Babin, R., & Dowell, R. (2024, May 23). *Final approval of ground-breaking EU AI Act*. Akin Gump Strauss Hauer & Feld. https://www.akingump.com/en/insights/alerts/final-approval-of-ground-breaking-eu-ai-act

Gaudet, A. (2024, January 17). AI-powered traffic control on I-30? A 'digital corridor' proposal could make it possible. *The Dallas Morning News*. https://www.dallasnews.com/news/transportation/2024/01/17/ai-powered-traffic-control-on-i-30-a-digital-corridor-proposal-could-make-it-possible/

Gent, E. (2023, March 31). *When AI's large language models shrink*. IEEE Spectrum. https://spectrum.ieee.org/large-language-models-size

Gent, E. (2024, March 25). The tech industry can't agree on what open-source AI means. That's a problem. *MIT Technology Review*. https://www.technologyreview.com/2024/03/25/1090111/tech-industry-open-source-ai-definition-problem/

Georgetown University. (2024, June 4). *OpenAI v. Scarlett Johansson? Georgetown L aw professor answers legal questions on AI-generated content*. https://www.georgetown.edu/news/ask-a-professor-openai-v-scarlett-johansson/

German, K. (2023, July 3). *How the world's largest ports are using AI to keep the global supply chain humming*. TechFinitive. https://www.techfinitive.com/features/how-the-worlds-largest-ports-are-using-ai-to-keep-the-global-supply-chain-humming/

Ghose, R., Bantanidis, S., Master, K., Shah, R. S., Yu, Y., Sharma, P., Dilaj, K., Van Kooij, B., Zajac, A., Ren, B., Birch, D. G. W., Savova, D., Kornbacher, D., Chang, E., Adylov, E., Sena, K., Rutter, K., Van Buskirk, M., Alur, S., Kundu, S., Austin, S., Reich, S., Balasubramaniyan, V., & Gorelov, Z. (2024). *AI in finance: Bot, bank & beyond*. Citi GPS. https://ir.citi.com/gps/9j79xHla-vfPi785TYiSciffO0j4I0D52fI9LrahsLZEo6MpT4aM7SpwSFagAL9Clukqn2fwiJ_GNvDsLy4b6XEjftdK1abu

Gibbs, A. (2024, March 21). TN Gov. Lee signs ELVIS Act into law in honky-tonk, protects musicians from AI abuses. *The Tennessean*. https://www.tennessean.com/story/entertainment/music/2024/03/21/elvis-act-tennessee-gov-lee-signs-act-musicians-ai/73019388007/

Gokhale, N., Gajjaria, A., Kaye, R., & Kuder, D. (2019). *AI leaders in financial services: Common traits of frontrunners in the artificial intelligence race*. Deloitte Insights. https://www2.deloitte.com/content/dam/insights/us/articles/4687_traits-of-ai-frontrunners/DI_AI-leaders-in-financial-services.pdf

Gold, A. (2024, July 25). *Exclusive: Anthropic weighs in on California AI bill*. Axios. https://www.axios.com/2024/07/25/exclusive-anthropic-weighs-in-on-california-ai-bill

Gold, M., & Nehamas, N. (2024, March 12). Donald Trump and Joe Biden clinch their party nominations. *The New York Times*. https://www.nytimes.com/2024/03/12/us/politics/trump-republican-nomination.html

Google AI. (n.d.). *Our principles*. Retrieved August 21, 2024, from https://ai.google/responsibility/principles/

Google Cloud. (n.d.-a). *What is supervised learning?* Retrieved August 19, 2024, from https://cloud.google.com/discover/what-is-supervised-learning

Google Cloud. (n.d.-b). *What is artificial intelligence (AI) in finance?* Retrieved August 20, 2024, from https://cloud.google.com/discover/finance-ai

Google Cloud. (n.d.-c). *What is artificial general intelligence (AGI)?* Retrieved August 30, 2024, from https://cloud.google.com/discover/what-is-artificial-general-intelligence

Google Cloud. (n.d.-d). *What is unsupervised learning?* Retrieved August 30, 2024, from https://cloud.google.com/discover/what-is-unsupervised-learning

Google for Education. (n.d.). *Ivy Tech develops machine learning algorithm to identify at-risk students and provide early intervention*. Retrieved August 20, 2024, from https://edu.google.com/why-google/customer-stories/ivytech-gcp/

Gordon, R. (2023, July 12). *Generative AI imagines new protein structures.* MIT News. https://news.mit.edu/2023/generative-ai-imagines-new-protein-structures-0712

Govindiah, M. (2023, August 10). *8 ways AI and ML are transforming child welfare*. Unisys. https://www.unisys.com/blog-post/cis/8-ways-ai-and-ml-are-transforming-child-welfare/

Gunning, D. (2016, August 11) *Explainable Artificial Intelligence (XAI)* [PowerPoint slides]. DARPA. https://www.darpa.mil/attachments/XAIIndustryDay_Final.pptx

Hammer, A. (2024, July 31). How FedEx is helping cops build a mass surveillance network. *The Daily Mail*. https://www.dailymail.co.uk/news/article-13552111/FedEx-trucks-spying-cameras-police.html

Hassan, J. (2023, August 9). AI is being used to give dead, missing kids a voice they didn't ask for. *The Washington Post*. https://www.washingtonpost.com/technology/2023/08/09/ai-dead-children-tiktok-videos/

Harris, L., & Jaikaran, C. (2024). *Highlights of the 2023 executive order on artificial intelligence for Congress*. Congressional Research Service. https://crsreports.congress.gov/product/pdf/R/R47843

Haven, J. (2022, October 4). *How to read the White House's Blueprint for an AI Bill of Rights*. Medium. https://medium.com/datasociety-points/how-to-read-the-white-houses-blueprint-for-an-ai-bill-of-rights-bf6cf312544a

Hawkins, M. (2024, June 18). *US lawmakers look to bar Chips Act winners from using Chinese tools*. Bloomberg. https://www.bloomberg.com/news/articles/2024-06-18/lawmakers-look-to-bar-chips-act-winners-from-using-chinese-tools

HB 4. Enrolled. 88th Texas Legislature. Regular. (2023). https://capitol.texas.gov/tlodocs/88R/billtext/pdf/HB00004F.pdf

HB 18. Enrolled. 88th Texas Legislature. Regular. (2023). https://capitol.texas.gov/tlodocs/88R/billtext/pdf/HB00018F.pdf

HB 1181. Enrolled. 88th Texas Legislature. Regular. (2023). https://capitol.texas.gov/tlodocs/88R/billtext/pdf/HB01181F.pdf

HB 2060. Enrolled. 88th Texas Legislature. Regular. (2023). https://capitol.texas.gov/tlodocs/88R/billtext/pdf/HB02060F.pdf

HB 2060 Bill Analysis. 2023. Senate Research Center. 88th Texas Legislature. Regular. (2023, May). https://capitol.texas.gov/tlodocs/88R/analysis/html/HB02060E.htm

HB 2700. Enrolled. 88th Texas Legislature. Regular. (2023). https://capitol.texas.gov/tlodocs/88R/billtext/pdf/HB02700F.pdf

Heath, R. (2023, August 9). *Exclusive poll: Americans distrust AI giants*. Axios. https://www.axios.com/2023/08/09/ai-voters-trust-government-regulation

Heath, R. (2024, May 1). *AI hits trust hurdles with U.S. military*. Axios. https://www.axios.com/2024/05/01/pentagon-military-ai-trust-issues

Heilweil, R., & Alder, M. (2024, January 25). *Justice Department discloses FBI project with Amazon Rekognition tool*. FedScoop. https://fedscoop.com/doj-fbi-amazon-rekognition-technology-ai-use-case/

Hicks, K. (2023, August 28). *Deputy Secretary of Defense Kathleen Hicks keynote address: 'The urgency to innovate' (as delivered)* [Speech]. U.S. Department of Defense. https://www.defense.gov/News/Speeches/Speech/Article/3507156/deputy-secretary-of-defense-kathleen-hicks-keynote-address-the-urgency-to-innov/

Holdsworth, J. (2023, December 22). *What is AI bias?* IBM. https://www.ibm.com/topics/ai-bias

Holdsworth, J., & Scapicchio, M. (2024, June 17). *What is deep learning?* IBM. Retrieved August 19, 2024, from https://www.ibm.com/topics/deep-learning

House Select Committee on Artificial Intelligence & Emerging Technologies. (2024). *Interim report to the eighty-ninth Texas Legislature*. Texas House of Representatives. https://www.house.texas.gov/pdfs/committees/reports/interim/88interim/House-Select-Committee-on-Artificial-Intelligence-Emerging-Technologies.pdf

Hsiung, C., & Chen, F. (2023, September 20). Exploring AI for law enforcement. *Police Chief*. https://www.policechiefmagazine.org/exploring-ai-law-enforcement-interview/

Huberman, A. (2023, December 25). *Rick Rubin: Protocols to access creative energy and process* [Video]. YouTube. https://www.youtube.com/watch?v=GpgqXCkRO-w

Huff, J., Ward, L., Lo, T. A., Dong, N., & Townsend, D. (2022, September 23). *Key takeaways from President Biden's Executive Order 14083 on ensuring robust consideration of evolving national security risks by the Committee on Foreign Investment in the United States ("CFIUS")*. Dorsey & Whitney. https://www.dorsey.com/newsresources/publications/client-alerts/2022/09/takeaways-from-bidens-eo-14083

Huff, N. (n.d.). *AI in radiology: Shaping the future of medical imaging*. Segmed. Retrieved October 15, 2024, from https://www.segmed.ai/resources/blog/ai-radiology-and-the-future

Hutson, M. (2024). How AI is being used to accelerate clinical trials. *Nature*, *627*, S2–S5. https://doi.org/10.1038/d41586-024-00753-x

IAPP. (2024). *Key terms for AI governance*. https://iapp.org/media/pdf/resource_center/key_terms_for_ai_governance.pdf

IBM. (2021, June 29). *Structured vs unstructured data*. https://www.ibm.com/think/topics/structured-vs-unstructured-data

IBM. (n.d.-a). *What is machine learning (ML)?* Retrieved August 19, 2024, from https://www.ibm.com/topics/machine-learning

IBM. (n.d.-b). *What is unsupervised learning?* Retrieved August 19, 2024, from https://www.ibm.com/topics/unsupervised-learning

IBM. (n.d.-c). *What is a neural network?* Retrieved August 19, 2024, from https://www.ibm.com/topics/neural-networks

IBM. (n.d.-d). *What is explainable AI?* Retrieved August 19, 2024, from https://www.ibm.com/topics/explainable-ai

IBM. (n.d.-e). *What is a chatbot?* Retrieved August 30, 2024, from https://www.ibm.com/topics/chatbots

IBM. (n.d.-f). *What is deep learning?* Retrieved August 30, 2024, from https://www.ibm.com/topics/deep-learning

IBM. (n.d.-g). *What is explainable AI?* Retrieved August 30, 2024, from https://www.ibm.com/topics/explainable-ai

IBM. (n.d.-h). *What are large language models (LLMs)?* Retrieved August 30, 2024, from https://www.ibm.com/topics/large-language-models

IBM. (n.d.-i). *What is a neural network?* Retrieved August 19, 2024, from https://www.ibm.com/topics/neural-networks

IBM. (n.d.-j). *What is predictive analytics?* Retrieved August 30, 2024, from https://www.ibm.com/topics/predictive-analytics

IBM. (n.d.-k). *What is supervised learning?* Retrieved August 30, 2024, from https://www.ibm.com/topics/supervised-learning

IBM. (n.d.-l). *What is speech recognition?* Retrieved August 30, 2024, from https://www.ibm.com/topics/speech-recognition

Iktos. (n.d.). *Iktos' mission*. Retrieved August 19, 2024, from https://iktos.ai/about

InCrowd. (2019, August 6). *InCrowd study shows a 79% burnout level among PCPs, and 68% across all specialties, spotlighting a national problem* [Press release]. https://incrowdnow.com/news/incrowd-study-shows-a-79-burnout-level-among-pcps-and-68-across-all-specialties-spotlighting-a-national-problem

Indeed. (n.d.). *How we're championing responsible AI in HR Tech*. Retrieved October 21, 2024, from https://www.indeed.com/esg/responsible-ai

Intel. (2023, September 18). *Moore's law*. https://www.intel.com/content/www/us/en/newsroom/resources/moores-law.html

Intelligent. (2023, September 27). *8 in 10 colleges will use AI in admissions by 2024*. https://www.intelligent.com/8-in-10-colleges-will-use-ai-in-admissions-by-2024/

International Association of Chiefs of Police. (n.d.). *Automated license plate recognition*. Retrieved August 20, 2024, from https://www.theiacp.org/projects/automated-license-plate-recognition

Internet Watch Foundation. (2022). *The annual report 2022*. https://annualreport2022.iwf.org.uk/wp-content/uploads/2023/04/IWF-Annual-Report-2022_FINAL.pdf

Isaacson, W. (2023). *Elon Musk*. Simon & Schuster.

Jain, R. (2024, March 15). *The European Union's AI Act: What you need to know*. Holland & Knight. https://www.hklaw.com/en/insights/publications/2024/03/the-european-unions-ai-act-what-you-need-to-know

Jargon, J. (2024, June 18). 'I felt shameful and fearful': Teen who saw AI fake nudes of herself speaks out. *The Wall Street Journal*. https://www.wsj.com/politics/policy/teen-deepfake-ai-nudes-bill-ted-cruz-amy-klobuchar-3106eda0

Johnson, K. B., Wei, W-Q., Weeraratne, D., Frisse, M. E., Misulis, K., Rhee, K., Zhao, J., & Snowdon, J. L. (2021). Precision medicine, AI, and the future of personalized health care. *Clinical and Translational Science*, *14*(1), 86–93. https://doi.org/10.1111/cts.12884

Jones, J. (2024, February 6). *New Hampshire AG places blame for deepfake Biden robocalls*. MSNBC. https://www.msnbc.com/the-reidout/reidout-blog/biden-robocalls-new-hampshire-ai-rcna137517

Jones, O. T., Matin, R. N., van der Schaar, M., Bhayankaram, K. P., Ranmuthu, C. K. I., Islam, M. S., Behiyat, D., Boscott, R., Calanzani, N., Emery, J., Williams, H. C., & Walter, F. M. (2022). Artificial intelligence and machine learning algorithms for early detection of skin cancer in community and primary care settings: A systematic review. *The Lancet Digital Health*, *4*(6), e466–e476. https://doi.org/10.1016/S2589-7500(22)00023-1

Joudaki, H., Rashidian, A., Minaei-Bidgoli, B., Mahmoodi, M., Geraili, B., Nasiri, M., & Arab, M. (20165). Improving fraud and abuse detection in general physician claims: A data mining study. *International Journal of Health Policy and Management*, *5*(3), 165–172. https://doi.org/10.15171/ijhpm.2015.196

Joynes, G., Anderson, B., McKemey, T., Russo, G., & Lay, C. (2024, June 4). *Senate AI working group releases landmark policy roadmap*. Brownstein Hyatt Farber Schreck. https://www.bhfs.com/insights/alerts-articles/2024/senate-ai-working-group-releases-landmark-policy-roadmap

Justice Innovation. (n.d.). *Legal tasks for AI in the access to justice domain*. Retrieved August 20, 2024, from https://justiceinnovation.law.stanford.edu/projects/ai-access-to-justice/tasks/

Kaling, M. (Writer), & Heckerling, A. (Director). (2005, April 26). Hot girl (Season 1, Episode 6) [TV series episode]. In Daniels, G., Gervais, R., Klein, H., Merchant, S., & Silverman, B. (Executive Producers), *The Office*. Reveille Productions; Deedle-Dee Productions; both in association with NBC Universal Television Studios. https://vimeo.com/498212706

Kallenborn, Z. (2022, February 1). Giving an AI control of nuclear weapons: What could possibly go wrong? *Bulletin of the Atomic Scientists*. https://thebulletin.org/2022/02/giving-an-ai-control-of-nuclear-weapons-what-could-possibly-go-wrong/

Kang, C. (2024, September 29). California governor vetoes sweeping A.I. legislation. *The New York Times*. https://www.nytimes.com/2024/09/29/technology/california-ai-bill.html

Kapko, M. (2024, March 11). *Ransomware attacks are hitting critical infrastructure more often, FBI says* [Brief]. Cybersecurity Dive. https://www.cybersecuritydive.com/news/ransomware-hitting-critical-infrastructure-fbi/709814/

Karimi, F. (2023, April 29). *'Mom, these bad men have me': She believes scammers cloned her daughter's voice in a fake kidnapping*. CNN. https://www.cnn.com/2023/04/29/us/ai-scam-calls-kidnapping-cec/index.html

Karjian, R. (2023, August 16). *The history of artificial intelligence: Complete AI timeline*. TechTarget. https://www.techtarget.com/searchenterpriseai/tip/The-history-of-artificial-intelligence-Complete-AI-timeline

Kaur, R. (n.d.). *Is it good for lawyers to use AI for legal briefs?* CaseFox. Retrieved August 20, 2024, from https://www.casefox.com/blog/ai-legal-brief/

Kawooy, E. N. (2016, April 14). *To X-ray or not to X-ray?* World Health Organization. https://www.who.int/news-room/feature-stories/detail/to-x-ray-or-not-to-x-ray-

Kelly, J. (2024, April 24). *What is black box AI?* Invoca Blog. https://www.invoca.com/blog/what-is-black-box-ai

Kelly, S. M. (2024, January 25). *Explicit, AI-generated Taylor Swift images spread quickly on social media*. CNN. https://www.cnn.com/2024/01/25/tech/taylor-swift-ai-generated-images/index.html

Klepper, D. (2024, June 13). *Presidential election a prime target for foreign disinformation, intelligence officials say*. The Associated Press. https://apnews.com/article/disinformation-election-interference-russia-china-2925ca62c1ba8c1893454ccb2f76785a

Kohne, N. G., Dowell, R., & Hold, J. (2024, May 28). *Colorado passes new watershed AI consumer protection bill*. Akin Gump Strauss Hauer & Feld. https://www.akingump.com/en/insights/alerts/colorado-passes-new-watershed-ai-consumer-protection-bill

Krall, A. (2024, February 26). *St. Louis County mother warns of new AI kidnapping phone scam*. KSDK. https://www.ksdk.com/article/tech/ai-phone-call-scam-kidnapping-ransom-st-louis-county-parents/63-3a293efd-eac1-4dd1-9ea8-67a58398ad2a

Kremer, A., Govindarajan, A., Singh, H., & Kristensen, I., & Li, E. (2024, July 1). *Embracing generative AI in credit risk*. McKinsey & Company. https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/embracing-generative-ai-in-credit-risk

Krietzberg, I. (2023, September 18). *The laws to regulate AI are already in place, expert argues*. TheStreet. https://www.thestreet.com/technology/the-laws-to-regulate-ai-are-already-in-place-expert-argues

Krishan, N. (2023, November 9). *Tech groups push back on Biden AI executive order, raising concerns that it could crush innovation*. FedScoop. https://fedscoop.com/tech-groups-push-back-on-biden-ai-executive-order-raising-concerns-that-it-could-crush-innovation/

Kuipers, B. (2019, September 15). *Progress in AI*. University of Michigan. https://web.eecs.umich.edu/~kuipers/opinions/AI-progress.html

Lamparth, M., & Schneider, J. (2024, April 29). Why the military can't trust AI. *Foreign Affairs*. https://www.foreignaffairs.com/united-states/why-military-cant-trust-ai

LaPedis, R. (2023, May 19). *Argus BWC: Empowering law enforcement with advanced AI, redaction and officer safety features*. Police1. https://www.police1.com/police-products/body-cameras/articles/argus-bwc-empowering-law-enforcement-with-advanced-ai-redaction-and-officer-safety-features-xfdRA7dhUWbgc9Ks/

Lat, D. (2024a, February 28). *AI use in law practice needs common sense, not more court rules*. Bloomberg Law. https://news.bloomberglaw.com/business-and-practice/ai-use-in-law-practice-needs-common-sense-not-more-court-rules

Lat, D. (2024b, February 29). *New court rules for AI: Stop the insanity*. Original Jurisdiction. https://davidlat.substack.com/p/new-court-rules-for-ai-stop-the-insanity

Lederer, E. M. (2024, March 21). *The UN adopts a resolution backing efforts to ensure artificial intelligence is safe*. The Associated Press. https://apnews.com/article/united-nations-artificial-intelligence-safety-resolution-vote-8079fe83111cced0f0717fdecefffb4d

Lee, J. (2024, March 6). *AI-driven credit risk decisioning: What you need to know*. Experian. https://www.experian.com/blogs/insights/ai-driven-credit-risk-decisioning/

Lee, N. T., & Chijioke, O. (2023, December 15). *Why states and localities are acting on AI*. The Brookings Institution. https://www.brookings.edu/articles/why-states-and-localities-are-acting-on-ai/

Let's Talk Science. (2023, November 9). *AI and personal vehicles*. https://letstalkscience.ca/educational-resources/backgrounders/ai-and-personal-vehicles

Levitt, K. (2024, January 11). *AI takes center stage: Survey reveals financial industry's top trends for 2024*. NVIDIAvidia. https://blogs.nvidia.com/blog/ai-in-financial-services-survey-2024/

Lippmann, W. (1943). *U.S. foreign policy: Shield of the republic*. Little, Brown and Company. https://archive.org/details/in.ernet.dli.2015.74564/page/n67/mode/2up

Liptak, K., & Atwood, K. (2023, March 17). *Biden administration skeptical of Xi's intentions ahead of his summit with Putin*. CNN. https://www.cnn.com/2023/03/17/politics/biden-putin-xi/index.html

Local Law 2021/144. Enacted. City of New York. (2021). https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=4344524&GUID=B051915D-A9AC-451E-81F8-6596032FA3F9&Options=ID%7cText%7c&Search=

Lockhart, J. (2024, April 11). Pioneering the future of smart cities with AI and generative AI. *Forbes*. https://www.forbes.com/sites/delltechnologies/2024/04/11/pioneering-the-future-of-smart-cities-with-ai-and-generative-ai/

Lodge, M. (2024, January 15). Google to remove sick AI-generated images of James Bulger from its platforms after campaign by murdered toddler's mother. *The Daily Mail*. https://www.dailymail.co.uk/news/article-12963745/Google-removes-James-Bulger-AI-generated-images-videos-TikTok.html

Lohr, S. (2018, February 9). Facial recognition is accurate, if you're a white guy. *The New York Times*. https://www.nytimes.com/2018/02/09/technology/facial-recognition-race-artificial-intelligence.html

Lomas, N. (2022, June 29). *Google's 'deceptive' account sign-up process targeted with GDPR complaints*. TechCrunch. https://techcrunch.com/2022/06/29/google-account-gdpr-complaint/

Lomas, N. (2024, August 1). *The EU's AI Act is now in force*. TechCrunch. https://techcrunch.com/2024/08/01/the-eus-ai-act-is-now-in-force/

Long, W. R. M., Cuyvers, L., Bruynseraede, M., & Kumar, Subhalakshmi, K. (2024, March 21). *EU formally adopts world's first AI law*. Sidley Austin. https://datamatters.sidley.com/2024/03/21/eu-formally-adopts-worlds-first-ai-law/

Luna, A. (2024, May 2). *The open or closed AI dilemma*. Bipartisan Policy Center. https://bipartisanpolicy.org/blog/the-open-or-closed-ai-dilemma/

Lyngaas, S. (2024, February 7). *Chinese hackers have lurked in some US infrastructure systems for 'at least five years.'*. CNN. https://www.cnn.com/2024/02/07/politics/china-hacking-us-agencies-report/index.html

Maass, D. (2021, April 22). *Data driven 2: California dragnet—New data set shows scale of vehicle surveillance in the Golden State*. Electronic Frontier Foundation. https://www.eff.org/deeplinks/2021/04/data-driven-2-california-dragnet-new-dataset-shows-scale-vehicle-surveillance

Malewar, A. (2023, July 13). *FrameDiff: A generative AI to craft new protein structures*. Tech Explorist. https://www.techexplorist.com/framediff-generative-ai-craft-new-protein-structures/64141/

Mancebo, A., Jr. (2023, September 1). *With arms wide open (or not): Navigating open vs. closed development of powerful AI models*. Data Science Alliance. https://www.datasciencealliance.org/work/opensource-vs-closedsource-article

Manning, C. (2020). *Artificial intelligence definitions*. Standford University. https://hai.stanford.edu/sites/default/files/2020-09/AI-Definitions-HAI.pdf

Manson, K. (2023, October 11). *US Space Force pauses generative AI use based on security concerns*. Bloomberg. https://www.bloomberg.com/news/articles/2023-10-11/space-force-pauses-generative-ai-based-on-security-concerns

Martin Heinrich. (n.d.). *Artificial intelligence caucus*. Retrieved August 21, 2024, from https://www.heinrich.senate.gov/artificial-intelligence-caucus

Martineau, K. (2023, April 20). *What is generative AI?* IBM. https://research.ibm.com/blog/what-is-generative-AI

Martoccio, A. (2023, December 14). Taylor Swift proposal tears Pennsylvania Legislature apart. *Rolling Stone*. https://www.rollingstone.com/music/music-news/taylor-swift-pennsylvania-legislature-cringe-1234928841/

Mastercard. (n.d.). *Mastercard® currency converter calculator*. Retrieved October 26, 2024, from https://www.mastercard.us/en-us/personal/get-support/convert-currency.html

Mayer Brown. (2024, February 8). *Biden artificial intelligence executive order action tracker*. https://www.mayerbrown.com/en/insights/publications/2024/02/biden-artificial-intelligence-executive-order-action-tracker

MCA. (n.d.). *ALPR cameras for public safety*. Retrieved August 20, 2024, from https://callmc.com/mss/vehicle-upfitting/vehicle-video-camera-systems/alpr-and-lpr-cameras/

McCullom, R. (2024, August 7). *In some cities, second thoughts about gunshot detection sensors*. Undark. https://undark.org/2024/08/07/second-thoughts-gunshot-detection-technology/

McElligott, B. (2023, March 29). *EU AI Act: Risk categories*. Mason Hayes & Curran. https://www.mhc.ie/hubs/the-eu-artificial-intelligence-act/eu-ai-act-risk-categories

McKinsey & Company. (2024). *What is AI (artificial intelligence)?* https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-ai/

McLellan, J. (2024, June 25). *AI: Pope expresses concern over 'technocratic' future*. United States Conference of Catholic Bishops. https://www.usccb.org/news/2024/ai-pope-expresses-concern-over-technocratic-future

McNally, L. (2023, November 30). *Understanding the EU AI Act's risk levels*. OneTrust. https://www.onetrust.com/blog/understanding-the-eu-ai-acts-risk-levels/

McNerney, J., Moeller, E. V., Daniels, B. L., Saperstein, C. J., Cote, B. J., Zimmer, E., & Trozzo, A. (2024, March 13). *Congress continues to address AI in bipartisan fashion, launching House AI task force as latest step*. Pillsbury Winthrop Shaw Pittman. https://www.pillsburylaw.com/en/news-and-insights/congress-ai-task-force.html

McPherson-Smith, O. (2024, January 11). *Manufacturing a crisis: The Biden administration's abuse of the Defense Production Act*. America First Policy Institute. https://americafirstpolicy.com/issues/manufacturing-a-crisis-the-biden-administrations-abuse-of-the-defense-production-act

McSherry, C. (2024, January 19). *The No AI Fraud Act creates far more problems than it solves*. Electronic Frontier Foundation. https://www.eff.org/deeplinks/2024/01/no-ai-fraud-act-creates-way-more-problems-it-solves

McSilver Institute. (2021, November 12). *Experts discuss pros and cons of predictive risk tools in child welfare practice*. https://mcsilver.nyu.edu/predictive-risk-tools-in-child-welfare-practice/

Meier, K., & Spichiger, R. (2024, March 15). *The EU AI Act: What it means for your business*. Ernst & Young Limited. https://www.ey.com/en_ch/forensic-integrity-services/the-eu-ai-act-what-it-means-for-your-business

Meng, X. (2023). Data science and engineering with human in the loop, behind the loop, and above the loop. *Harvard Data Science Review*, *5*(2). https://doi.org/10.1162/99608f92.68a012eb

Merriam-Webster. (n.d.). Ad astra per aspera. In *Merriam-Webster.com dictionary*. Retrieved August 19, 2024, from https://www.merriam-webster.com/dictionary/ad%20astra%20per%20aspera

Metz, C., & Issac, M. (2023, May 18). In battle over A.I., Meta decides to give away its crown jewels. *The New York Times*. https://www.nytimes.com/2023/05/18/technology/ai-meta-open-source.html

Metz, R. (2019, October 7). *The number of deepfake videos online is spiking. Most are porn*. CNN. https://www.cnn.com/2019/10/07/tech/deepfake-videos-increase/index.html

Miller, L. (2023, July 5). *Arnold Schwarzenegger believes AI has made Terminator's dystopian future 'a reality.'* CBR. https://www.cbr.com/arnold-schwarzenegger-terminator-ai-reality/

Mims, C. (2024, March 29). The AI industry is steaming toward a legal iceberg. *The Wall Street Journal*. https://www.wsj.com/tech/ai/the-ai-industry-is-steaming-toward-a-legal-iceberg-5d9a6ac1

Minsky, M. L. (1967). *Computation*: *Finite and infinite machines*. Prentice-Hall.

Mitchell, B. (Host). (2024, July 29). Biden's AI executive order milestones; USAID's new digital policy [Audio podcast episode]. In *The Daily Scoop Podcast*. FedScoop. https://fedscoop.com/radio/apples-ai-safety-commitment-and-usaids-10-year-tech-vision/

Moor, J. (2006). The Dartmouth College artificial intelligence conference: The next fifty years. *AI Magazine*, *27*(4), 87–91. https://doi.org/10.1609/aimag.v27i4.1911

Morning Joe. (2024, January 26). *'Is this going to be the deepfake election?': Analyzing AI's potential influence on 2024* [Video]. MSNBC. https://www.msnbc.com/morning-joe/watch/is-2024-going-to-be-the-deepfake-election-chris-krebs-is-gravely-concerned-202994757875

Morris, M. R., Sohl-Dickstein, J., Fiedel, N., Warkentin, T., Dafoe, A., Faust, A., Farabet, C., & Legg, S. (2024). *Levels of AGI for operationalizing progress on the path to AGI*. Cornell University. https://doi.org/10.48550/arXiv.2311.02462

Mount, I. (2022, December 19). *Farmers use $60 billion of pesticides each year. 2 MIT scientists have developed a new technology that could cut that number in half*. MIT MechE. https://meche.mit.edu/news-media/farmers-use-60-billion-pesticides-each-year-2-mit-scientists-have-developed-new

Mucci, T., & Stryker, C. (2023, December 18). *What is artificial superintelligence?* IBM. https://www.ibm.com/topics/artificial-superintelligence

Mucci, T., & Stryker, C. (2024, April 18). *Getting ready for artificial general intelligence with examples*. IBM. https://www.ibm.com/blog/artificial-general-intelligence-examples/

Mulligan, S. J. (2024, August 14). *A new public database lists all the ways AI could go wrong*. MIT Technology Review. https://www.technologyreview.com/2024/08/14/1096455/new-database-lists-ways-ai-go-wrong/

Murakami, K. (2024, January 26). *AI is helping police solve more crimes, but some are still worried*. Route Fifty. https://www.route-fifty.com/emerging-tech/2024/01/ai-helping-police-solve-more-crimes-some-are-still-worried/393670/

Nahra, K. J., Evers, A., Jessani, A. A., Braun, M., Vallery, A., & Benizri, I. (2024, March 14). *The European Parliament adopts the AI Act*. Wilmer Cutler Pickering Hale and Dorr. https://www.wilmerhale.com/en/insights/blogs/wilmerhale-privacy-and-cybersecurity-law/20240314-the-european-parliament-adopts-the-ai-act

National Center for Missing & Exploited Children. (2024, March 11). *Generative AI CSAM is CSAM*. https://www.missingkids.org/blog/2024/generative-ai-csam-is-csam

National Conference of State Legislatures. (2024a, January 12). *Artificial intelligence 2023 legislation*. https://www.ncsl.org/technology-and-communication/artificial-intelligence-2023-legislation

National Conference of State Legislatures. (2024b, September 9). *Artificial intelligence 2024 legislation*. https://www.ncsl.org/technology-and-communication/artificial-intelligence-2024-legislation

National Conference of State Legislatures. (2024c, October 10). *Deceptive audio or visual media ('deepfakes') 2024 legislation*. https://www.ncsl.org/technology-and-communication/deceptive-audio-or-visual-media-deepfakes-2024-legislation

National Institute of Standards and Technology. (2022, January 11). *About NIST*. https://www.nist.gov/about-nist

National Institute of Standards and Technology. (2023a). *Artificial intelligence risk management framework (AI RMF 1.0)*. https://doi.org/10.6028/NIST.AI.100-1

National Institute of Standards and Technology. (2023b, October 20). *NIST risk management framework aims to improve trustworthiness of artificial intelligence*. https://www.nist.gov/news-events/news/2023/01/nist-risk-management-framework-aims-improve-trustworthiness-artificial

National Institute of Standards and Technology. (2024a). *Artificial intelligence risk management framework: Generative artificial intelligence profile*. https://doi.org/10.6028/NIST.AI.600-1

National Institute of Standards and Technology. (2024b). *Artificial intelligence: The vitals*. https://www.nist.gov/system/files/documents/2023/11/02/AI%20Fact%20Sheet%200615%20FINAL.pdf

National Telecommunications and Information Administration. (2024). *NTIA artificial intelligence accountability policy report*. https://www.ntia.gov/sites/default/files/publications/ntia-ai-report-final.pdf

NBC News. (2024, January 22). *Listen: Fake Biden robocall tells voters to skip New Hampshire primary* [Video]. https://www.nbcnews.com/video/listen-fake-biden-robocall-tells-new-hampshire-not-to-vote-in-primary-202609733664

NetChoice. (n.d.). *Responsible AI principles to maintain America's tech dominance*. Retrieved August 21, 2024, from https://netchoice.org/responsible-ai-principles-to-maintain-americas-tech-dominance/

New York Foundation for the Arts. (2023, September 28). *How a new law regulating AI hiring software impacts employers*. LinkedIn. https://www.linkedin.com/pulse/how-new-law-regulating-ai-hiring-software/

Noelle, C. (2019, July 31). *5 ways artificial intelligence (AI) is revolutionizing the banking industry*. ProcessMaker. https://www.processmaker.com/blog/5-ways-artificial-intelligence-ai-is-revolutionizing-the-banking-industry/

Norden, L. & Harris, D. E. (2024, February 29). *New tech accord to fight AI threats to 2024 lacks accountability for companies*. Brennan Center for Justice. https://www.brennancenter.org/our-work/analysis-opinion/new-tech-accord-fight-ai-threats-2024-lacks-accountability-companies

Norris, D. (2019, June 12). Artificial intelligence and community-police relations. *Police Chief*. https://www.policechiefmagazine.org/ai-community-police-relations/

Office of the Attorney General. (2024, June 4). *Attorney General Ken Paxton launches data privacy and security initiative to protect Texans' sensitive data from illegal exploitation by tech, AI, and other companies* [Press release]. https://www.texasattorneygeneral.gov/news/releases/attorney-general-ken-paxton-launches-data-privacy-and-security-initiative-protect-texans-sensitive

Office of the Texas Governor. (2023, June 13). *Governor Abbott establishes new artificial intelligence advisory council* [Press release]. https://gov.texas.gov/news/post/governor-abbott-establishes-new-artificial-intelligence-advisory-council

Open Source Initiative. (2024, February 16). *The open source definition*. https://opensource.org/osd

Open Source Initiative. (n.d.). *The open source AI definition—1.0-RC2*. Retrieved November 4, 2024, from https://opensource.org/ai/drafts/the-open-source-ai-definition-1-0-rc2

OpenAI. (n.d.-a). *What is ChatGPT?* Retrieved August 19, 2024, from https://help.openai.com/en/articles/6783457-what-is-chatgpt

OpenAI. (n.d.-b). *Develop safe, beneficial AI systems*. Retrieved August 21, 2024, from https://openai.com/careers/

OpenAI. (2023, February 24). *Planning for AGI and beyond*. https://openai.com/index/planning-for-agi-and-beyond/

Organisation for Economic Co-operation and Development. (2019). *C/MIN(2019)3/FINAL*. https://one.oecd.org/document/C/MIN(2019)3/FINAL/en/pdf

Organisation for Economic Co-operation and Development. (2020, March 2). *What are the OECD principles on AI?* https://www.oecd-ilibrary.org/sites/6ff2a1c4-en/index.html?itemId=/content/paper/6ff2a1c4-en

Organisation for Economic Co-operation and Development. (2024, May 3). *OECD updates AI Principles to stay abreast of rapid technological developments* [Press release]. https://www.oecd.org/en/about/news/press-releases/2024/05/oecd-updates-ai-principles-to-stay-abreast-of-rapid-technological-developments.html

Organisation for Economic Co-operation and Development. (n.d.). *OECD AI principles overview*. Retrieved August 30, 2024, from https://oecd.ai/en/ai-principles

Patrick, D. (2024). *2024 interim legislative charges*. Office of the Texas Lieutenant Governor. https://www.ltgov.texas.gov/wp-content/uploads/2024/04/2024-Interim-Legislative-Charges.pdf

Pérez-Peña, R., & Rosenberg, M. (2018, January 29). Strava fitness app can reveal military sites, analysts say. *The New York Times*. https://www.nytimes.com/2018/01/29/world/middleeast/strava-heat-map.html

Perrigo, B. (2024, May 21). No one truly knows how AI systems work. A new discovery could change that. *Time*. https://time.com/6980210/anthropic-interpretability-ai-safety-research/

Piddubna, A. (2024, August 12). *AI in agriculture — the future of farming*. Intellias. https://intellias.com/artificial-intelligence-in-agriculture/

Plotinsky, D., & Cinelli, G. M. (2024, April 9). *Existing and proposed federal AI regulation in the United States*. Morgan, Lewis & Bockius. https://www.morganlewis.com/pubs/2024/04/existing-and-proposed-federal-ai-regulation-in-the-united-states

Polcyn, B. (2023, August 2). *Wisconsin AI-powered Flock cameras are tracking where you drive*. Fox6 Milwaukee. https://www.fox6now.com/news/wisconsin-ai-powered-flock-cameras

Polis, J. (2024, May 17). [Signing statement from Governor Jared Polis to the Honorable Colorado General Assembly]. https://progresschamber.org/wp-content/uploads/2024/05/SB24-205-Signing-Statement.pdf

Postman, N. (1998). *Five things we need to know about technological change*. University of California, Davis. https://web.cs.ucdavis.edu/~rogaway/classes/188/materials/postman.pdf

Potkin, F., & Mukherjee, S. (2023, October 11). *Exclusive: Southeast Asia eyes hands-off AI rules, defying EU ambitions*. Reuters. https://www.reuters.com/technology/southeast-asia-eyes-hands-off-ai-rules-defying-eu-ambitions-2023-10-11/

Ramer, H., & Swenson, A. (2024, May 23). *Political consultant behind fake Biden robocalls faces $6 million fine and criminal charges*. The Associated Press. https://apnews.com/article/biden-robocalls-ai-new-hampshire-charges-fines-9e9cc63a71eb9c78b9bb0d1ec2aa6e9c

Rao, S., & Ramstad, A. (2023, December 21). *Legal fictions and ChatGPT hallucinations: 'Mata v. Avianca' and generative AI in the courts*. Law.com. https://www.law.com/newyorklawjournal/2023/12/21/legal-fictions-and-chatgpt-hallucinations-mata-v-avianca-and-generative-ai-in-the-courts/

Rathbun, J. O. (2023, September 6). *DON guidance on the use of generative artificial intelligence and large language models*. Department of the Navy Chief Information Officer. https://www.doncio.navy.mil/ContentView.aspx?id=16442

Raymond, N. (2024, April 19). *US judicial panel wrestles with how to police AI-generated evidence*. Reuters. https://www.reuters.com/legal/transactional/us-judicial-panel-wrestles-with-how-police-ai-generated-evidence-2024-04-19/

Readocracy. (n.d.). *FAQ/guide*. Retrieved August 20, 2024, from https://readocracy.com/faqs

Redden, J., Aagaard, B., & Taniguchi, T. (2020a). *Artificial intelligence applications in law enforcement*. Criminal Justice Testing and Evaluation Center. https://cjtec.org/files/5f5f94aa4c69b

Redden, J., Banks, D., & Criminal Justice Testing and Evaluation Consortium. (2020b). *Artificial intelligence applications for criminal courts*. Criminal Justice Technology Testing and Evaluation Center. https://cjtec.org/files/65532c7675a44

Regulation (EU) 2016/679. Adopted. European Parliament. (2016). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

Regulation (EU) 2024/1689. Adopted. European Parliament. (2024). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689

Reisner, A. (2023, September 25). These 183,000 books are fueling the biggest fight in publishing and tech. *The Atlantic*. https://www.theatlantic.com/technology/archive/2023/09/books3-database-generative-ai-training-copyright-infringement/675363/

Research@Texas A&M. (2024, March 8). *Improving livestock management efficiency using artificial intelligence and 'smart' technology*. https://research.tamu.edu/2024/03/08/improving-livestock-management-efficiency-using-artificial-intelligence-and-smart-technology/

Resolution A/78/L.49. Adopted. United Nations General Assembly. (2024). https://documents.un.org/doc/undoc/ltd/n24/065/92/pdf/n2406592.pdf

Reuters. (2024a, May 9). *US lawmakers unveil bill to make it easier to restrict exports of AI models*. https://www.reuters.com/technology/us-lawmakers-unveil-bill-make-it-easier-restrict-exports-ai-models-2024-05-10/

Reuters. (2024b, May 15). *US raises concerns to Chinese officials about AI misuse*. Reuters. https://www.reuters.com/technology/us-china-hold-ai-risk-safety-talks-white-house-says-2024-05-15/

Rigano, C. (2019). Using artificial intelligence to address criminal justice needs. *NIJ Journal*, *280*. https://www.ojp.gov/pdffiles1/nij/252038.pdf

Rippy, S. (2021, May 10). *Opt-in vs. opt-out approaches to personal information processing*. IAPP. https://iapp.org/news/a/opt-in-vs-opt-out-approaches-to-personal-information-processing

Robert, J. (2024, February 26). *2024 EDUCAUSE AI landscape study*. EDUCAUSE. https://www.educause.edu/ecar/research-publications/2024/2024-educause-ai-landscape-study/the-future-of-ai-in-higher-education

Roberts, J., Jr. (2023). *2023 year-end report on the federal judiciary*. Supreme Court of the United States. https://www.supremecourt.gov/publicinfo/year-end/2023year-endreport.pdf

Rothschild, D. M. (2021, June 3). *Policy is also downstream of culture*. Discourse. https://www.discoursemagazine.com/p/policy-is-also-downstream-of-culture

Rouhiainen, L. (2019, October 14). How AI and data could personalize higher education. *Harvard Business Review*. https://hbr.org/2019/10/how-ai-and-data-could-personalize-higher-education

Ryan, W. A., Garrett, A., & Sears, B. (2023, August 8). *Practical lessons from the attorney AI missteps in Mata v. Avianca*. Association of Corporate Counsel. https://www.acc.com/resource-library/practical-lessons-attorney-ai-missteps-mata-v-avianca

S. 2691. AI Labeling Act of 2023. 118th Congress. (2023). https://www.govinfo.gov/content/pkg/BILLS-118s2691is/pdf/BILLS-118s2691is.pdf

Saeidi, M. (2024, May 17). *Voice cloning scams are a growing threat. Here's how you can protect yourself*. CBS News New York. https://www.cbsnews.com/newyork/news/ai-voice-clone-scam/

Salinas, M. P., Sepúlveda, J., Hidalgo, L., Periano, D., Morel, M., Uribe, P., Rotemberg, V., Briones, J., Mery, D., & Navarrete-Dechent, C. (2024). A systematic review and meta-analysis of artificial intelligence versus clinicians for skin cancer diagnosis. *npj Digital Medicine*, *7*(125). https://doi.org/10.1038/s41746-024-01103-x

Sanofi. (2022, January 7). *Exscientia and Sanofi establish strategic research collaboration to develop AI-driven pipeline of precision-engineered medicines* [Press release]. https://www.sanofi.com/en/media-room/press-releases/2022/2022-01-07-06-00-00-2362917

Savage, N. (2021, May 27). Tapping into the drug discovery potential of AI. *Nature Portfolio*. https://doi.org/10.1038/d43747-021-00045-7

SB 24-205. Signed. 74th Colorado General Assembly. Second Regular. (2024). https://leg.colorado.gov/sites/default/files/2024a_205_signed.pdf

SB 64. Enrolled. 86th Texas Legislature. Regular. (2019). https://capitol.texas.gov/tlodocs/86R/billtext/pdf/SB00064F.pdf

SB 79. Enrolled. 99th South Dakota Legislature. Regular. (2024). https://sdlegislature.gov/Session/Bill/24991/264651

SB 1047. Introduced. 2023–2024 California Legislature. Regular. (2024). https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202320240SB1047

SB 2105. Enrolled. 88th Texas Legislature. Regular. (2023). https://capitol.texas.gov/tlodocs/88R/billtext/pdf/SB02105F.pdf

Schumer, C., Rounds, M., Heinrich, M., & Young, T. (2024). *Driving U.S. innovation in artificial intelligence: A roadmap for artificial intelligence policy in the United States Senate*. The Bipartisan Senate AI Working Group. https://www.schumer.senate.gov/imo/media/doc/Roadmap_Electronic1.32pm.pdf

Schwartz, D. E., Jeffers, A. V., & Safko, E. D. (2023). *AI and the workplace: Employment considerations*. Skadden, Arps, Slate, Meagher & Flom. https://www.skadden.com/insights/publications/2023/06/quarterly-insights/ai-and-the-workplace

Security. (2023, September 19). *Energy sector faces 39% of critical infrastructure attacks*. https://www.securitymagazine.com/articles/99915-energy-sector-faces-39-of-critical-infrastructure-attacks

Security Hero. (2023). *2023 state of deepfakes*. https://www.securityhero.io/state-of-deepfakes/

Seitz-Wald, A. (2024, February 6). *N.H. attorney general says he found source of fake Biden robocalls*. NBC News. https://www.nbcnews.com/politics/2024-election/nh-attorney-general-says-found-source-fake-biden-robocalls-rcna137499

Selissen, A. (2024). *AI at TxDOT*. Texas Department of Transportation. https://txdir.widen.net/s/bdkgjhvwcd/3.-20240321_txdot_ai_advisory_council

Shanfeld, E. (2023, November 1). Scarlett Johansson takes legal action against AI app that ripped off her likeness in advertisement. *Variety*. https://variety.com/2023/digital/news/scarlett-johansson-legal-action-ai-app-ad-likeness-1235773489/

Shaw, A. (2023, September 16). *'I f----ing like humanity, dude': Elon Musk's friendship with then-Google CEO ended over AI: book*. Fox Business. https://www.foxbusiness.com/fox-news-tech/humanity-dude-elon-musks-friendship-then-google-ceo-ended-ai-book

Shaw, C., Yuan, L., Brennan, D., Martin, S., Janson, N., Fox, K., & Bryant, G. (2023). *GenAI in higher education: Fall 2023 update of time for class study*. Tyton Partners. https://tytonpartners.com/time-for-class-2023/GenAI-Update

Sheehan, M. (2022, January 4). *China's new AI governance initiatives shouldn't be ignored*. Carnegie Endowment for International Peace. https://carnegieendowment.org/posts/2022/01/chinas-new-ai-governance-initiatives-shouldnt-be-ignored

Shepardson, D. (2024, January 26). *Eying China, US proposes 'know your customer' cloud computing requirements*. Reuters. https://www.reuters.com/technology/us-propose-know-your-customer-requirements-cloud-computing-companies-2024-01-26/

Shivakumar, S., Wessner, C., & Howell, T. (2023, November 7). *"Guardrails" on CHIPS Act funding to restrict investments in China may restrict participation in CHIPS Act incentives*. Center for Strategic & International Studies. https://www.csis.org/blogs/perspectives-innovation/guardrails-chips-act-funding-restrict-investments-china-may-restrict

Sidley Austin. (2022, September 20). *Executive order directs CFIUS to conduct broad national security analysis*. https://www.sidley.com/en/insights/publications/2022/09/executive-order-directs-cfius-to-conduct-broad-national-security-analysis

Sidley Austin. (2023, November 6). *President Biden signs sweeping artificial intelligence executive order*. https://www.sidley.com/en/insights/newsupdates/2023/11/president-biden-signs-sweeping-artificial-intelligence-executive-order

Silberling, A. (2023, September 26). *The writers' strike is over; here's how AI negotiations shook out*. TechCrunch. https://techcrunch.com/2023/09/26/writers-strike-over-ai/

Simmons, S. (2024, May 13). *TEA to use AI to grade written portion of STAAR test*. KHOU 11. https://www.khou.com/article/news/education/texas-education-agency-artificial-intelligence-staar-tests/285-df68555c-3da2-4fc0-ba15-465bbf1aebfd

Simonite, T. (2020, January 27). AI license plate readers are cheaper—so drive carefully. *Wired*. https://www.wired.com/story/ai-license-plate-readers-cheaper-drive-carefully/

Singer, N. (2023, August 24). Despite cheating fears, schools repeal ChatGPT bans. *The New York Times*. https://www.nytimes.com/2023/08/24/business/schools-chatgpt-chatbot-bans.html

Singer, N. (2024, April 8). Teen girls confront an epidemic of deepfake nudes in schools. *The New York Times*. https://www.nytimes.com/2024/04/08/technology/deepfake-ai-nudes-westfield-high-school.html

Sinsky, C., Colligan, L., Ling, L., Prgomet, M., Reynolds, S., Goeders, L., Westbrooke, J., Tutty, M., & Blike, G. (2016). Allocation of physician time in ambulatory practice: A time and motion study in 4 specialties. *Annals of Internal Medicine*, *165*(11), 753–760. https://doi.org/10.7326/M16-0961

Sinwar, D., Dhaka, V. S., Tesfaye, B. A., Raghuwanshi, G., Kumar, A., Maakar, S. K., & Agrawal, S. (2022). Artificial intelligence and deep learning assisted rapid diagnosis of COVID-19 from chest radiographical images: A survey. *Contrast Media & Molecular Imaging*. https://doi.org/10.1155/2022/1306664

Siripurapu, A. (2021, December 22). *What is the Defense Production Act?* Council on Foreign Relations. https://www.cfr.org/in-brief/what-defense-production-act

Sisson, P. (2024, April 16). AI was supposed to make policy bodycams better. What happened? *MIT Technology Review*. https://www.technologyreview.com/2024/04/16/1090846/ai-police-body-cams-cops-transparency/

Siwicki, B. (2019, February 13). *IBM Watson Health's chief health officer talks healthcare challenges and AI*. Healthcare IT News. https://www.healthcareitnews.com/news/ibm-watson-health%E2%80%99s-chief-health-officer-talks-healthcare-challenges-and-ai

Sloly, P. (n.d.). *Emerging tech that can make smart cities safer*. Deloitte. Retrieved August 20, 2024, from https://www2.deloitte.com/ca/en/pages/public-sector/articles/emerging-tech-smart-cities-safer.html

Smith, A. (2024, April 16). *Coordinate compliance among laws regulating AI*. Society for Human Resource Management. https://www.shrm.org/topics-tools/employment-law-compliance/compliance-laws-regulating-ai

Software & Information Industry Association. (2023, October 30). *SIIA statement on White House AI executive order*. https://www.siia.net/siia-statement-on-white-house-ai-executive-order/

Sokler, B. D., Hecht, A., Fjeld, C. T., & Gambhir, R. (2023, December 7). *A timeline of Biden's AI executive order — AI: The Washington report*. Mintz, Levin, Cohn, Ferris, Glovsky and Popeo. https://www.mintz.com/insights-center/viewpoints/2191/2023-12-06-timeline-bidens-ai-executive-order-ai-washington-report

Sokler, B. D., Hecht, A., Fjeld, C. T., & Gambhir, R. (2024, February 2). *Biden's AI executive order achieves first major milestones (AI EO January update) — AI: The Washington report*. Mintz, Levin, Cohn, Ferris, Glovsky and Popeo. https://www.mintz.com/insights-center/viewpoints/54731/2024-02-01-bidens-ai-executive-order-achieves-first-major

Solaiman, I. (2023, February 5). *The gradient of generative AI release: Methods and considerations*. arXiv. https://arxiv.org/pdf/2302.04844

Somers, M. (2020, July 21). *Deepfakes, explained*. Ideas Made to Matter, MIT Sloan School of Management. https://mitsloan.mit.edu/ideas-made-to-matter/deepfakes-explained

Sorensen, E. (2019, September 9). *What is a neobank? How does it differ from traditional banks?* MobileTransaction. https://www.mobiletransaction.org/what-is-a-neo-bank/

SoundThinking. (n.d.). *ShotSpotter*. Retrieved August 20, 2024, from https://www.soundthinking.com/law-enforcement/leading-gunshot-detection-system/

Spinks, R. (2023, October 20). *What you need to know about automatic license plate readers*. American Police Beat. https://apbweb.com/2023/10/what-you-need-to-know-about-automatic-license-plate-readers/

Stanley, J. (2022). *Fast-growing company Flock is building a new AI-driven mass-surveillance system*. American Civil Liberties Union. https://assets.aclu.org/live/uploads/publications/flock_1.pdf

Statt, N. (2020, June 10). *Amazon bans police from using its facial recognition technology for the next year*. The Verge. https://www.theverge.com/2020/6/10/21287101/amazon-rekognition-facial-recognition-police-ban-one-year-ai-racial-bias

Steele, C. (2023, October 18). *The internet is full of deepfakes, and most of them are porn*. PCMag. https://www.pcmag.com/news/the-internet-is-full-of-deepfakes-and-most-of-them-are-porn

Stelloh, T. (2024, April 2). *Washington state judge blocks use of AI-enhanced video as evidence in possible first-of-its-kind ruling*. NBC News. https://www.nbcnews.com/news/us-news/washington-state-judge-blocks-use-ai-enhanced-video-evidence-rcna141932

Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., Hirschberg, J., Kalyanakrishnan, S., Kamar, E., Kraus, S., Leyton-Brown, K., Parkes, D., Press, W., Saxenian, A., Shah, J., Tambe, M., & Teller, A. (2016). *Artificial intelligence and life in 2030*. Stanford University. https://ai100.stanford.edu/sites/g/files/sbiybj18871/files/media/file/ai100report10032016fnl_singles.pdf

Stoyanov, L. (2024, February 4). AI vs. human translators: Who will win the translation battle? (J. Fell, Ed.). *Entrepreneur Europe*. https://www.entrepreneur.com/en-gb/science-technology/ai-vs-human-translators-who-will-win-the-translation/467210

Strickland, E. (2022, October 14). 6 reactions to the White House's AI Bill of Rights . *IEEE Spectrum*. https://spectrum.ieee.org/white-house-ai

Stryker, C., & Kavlakoglu, E. (2024, August 16). *What is artificial intelligence (AI)?* IBM. Retrieved August 19, 2024, from https://www.ibm.com/topics/artificial-intelligence

Surfshark. (n.d.). *Children and online risks: Global statistics*. Retrieved August 19, 2024, from https://surfshark.com/research/cybersecurity-for-kids/statistics

Surveillance Technology Oversight Project. (2022, October 4). *S.T.O.P. expresses concern that White House AI Bill of Rights normalizes AI abuses* [Press release]. https://www.stopspying.org/latest-news/2022/10/4/stop-expresses-concern-that-white-house-ai-bill-of-rights-normalizes-ai-abuses

Susarla, A. (2023, November 4). *Analysis: How Biden's new executive order tackles AI risks, and where it falls short*. PBS News. https://www.pbs.org/newshour/politics/analysis-how-bidens-new-executive-order-tackles-ai-risks-and-where-it-falls-short

Swarns, C. (2023, September 19). *When artificial intelligence gets it wrong*. Innocence Project. https://innocenceproject.org/when-artificial-intelligence-gets-it-wrong/

Takahama, E. (2024, February 25). Why health care has become a top target for cybercriminals. *The Seattle Times*. https://www.seattletimes.com/seattle-news/health/why-health-care-has-become-a-top-target-for-cybercriminals/

Taylor, M. (2024, August 2). Hospitals grapple with radiologist shortage. *Becker's Hospital Review*. https://www.beckershospitalreview.com/radiology/hospitals-grapple-with-radiologist-shortage.html

TechNet. (2024, January 1). *Artificial intelligence*. https://www.technet.org/policy/artificial-intelligence-federal/

TechTarget. (2021). *Kill switch*. https://www.techtarget.com/whatis/definition/kill-switch

Teixeira, C., & Boyas, M. (2017). *Predictive analytics in child welfare: An assessment of current efforts, challenges and opportunities*. U.S. Department of Health and Human Services. https://aspe.hhs.gov/sites/default/files/migrated_legacy_files/178496/PACWAnAssessmentCurrentEffortsChallengesOpportunities.pdf

Tenbarge, K. (2023, March 27). *Found through Google, bought with Visa and Mastercard: Inside the deepfake porn economy*. NBC News. https://www.nbcnews.com/tech/internet/deepfake-porn-ai-mr-deep-fake-economy-google-visa-mastercard-download-rcna75071

Tenbarge, K. (2024, January 16). *Teen deepfake victim pushes for federal law targeting AI-generated explicit content*. NBC News. https://www.nbcnews.com/tech/tech-news/deepfake-law-ai-new-jersey-high-school-teen-image-porn-rcna133706

Terech, K. (2024, February 27). *Google's Gemini will be right back after these hallucinations: Image generator to make a return after historical blunders*. TechRadar. https://www.techradar.com/computing/artificial-intelligence/googles-gemini-will-be-right-back-after-these-hallucinations-image-generator-to-make-a-return-after-historical-blunders

Teshome, E. (2024, July 26). Apple to support Biden AI safety guidelines. *The Hill*. https://thehill.com/homenews/4794915-apple-ai-safety-guidelines-biden-executive-order/

Tex. Business and Commerce Code § 1.201 (1967 & rev. 1973, 1983, 1989, 1995, 1999, 2003, 2005, 2015, 2021). https://statutes.capitol.texas.gov/Docs/BC/htm/BC.1.htm

Tex. Business and Commerce Code § 509.001 (2023). https://statutes.capitol.texas.gov/Docs/BC/htm/BC.509.htm

Tex. Business and Commerce Code § 541.001 (2023). https://statutes.capitol.texas.gov/Docs/BC/htm/BC.541.htm

Tex. Election Code § 255.004 (1987 & rev. 2019). https://statutes.capitol.texas.gov/docs/el/htm/el.255.htm

Tex. Government Code § 2054.621 (2023). https://statutes.capitol.texas.gov/Docs/GV/htm/GV.2054.htm

Tex. Penal Code § 21.165 (2023). https://statutes.capitol.texas.gov/Docs/PE/htm/PE.21.htm

Texas Department of Information Resources [@TexasDIR]. (2024, June 6). *The AI Advisory Council heard from representatives of from @TexasDFPS, @TexasHHSC, @TXAG, @TxDPS, @TDCJ, and @TexasTDI regarding their current and* [Image attached] [Post]. X. https://x.com/TexasDIR/status/1798827087542768073

Texas Education Agency. (2024). *Hybrid scoring key questions*. Texas Education Agency. https://tea.texas.gov/student-assessment/testing/hybrid-scoring-key-questions.pdf

Texas Public Policy Foundation. (2023, October 31). *TPPF statement on Biden administration's A.I. executive order* [Press release]. https://www.texaspolicy.com/press/tppf-statement-on-biden-administrations-a-i-executive-order

The Authors Guild. (2023, December 21). *New Authors Guild AI survey reveals that authors overwhelmingly want consent and compensation for use of their works*. https://authorsguild.org/news/ag-ai-survey-reveals-authors-overwhelmingly-want-consent-and-compensation-for-use-of-their-works/

The Economist. (2017, July 27). Why China's AI push is worrying. https://www.economist.com/leaders/2017/07/27/why-chinas-ai-push-is-worrying

The Economist. (2019, September 5). Artificial intelligence and war. https://www.economist.com/leaders/2019/09/05/artificial-intelligence-and-war

The National Museum of Computing. (n.d.). *The Turing-Welchman Bombe*. Retrieved August 19, 2024, from https://www.tnmoc.org/bombe

The Princeton Review. (n.d.). *Intelligent tutoring systems: Enhancing learning through AI.* Retrieved August 20, 2024, from https://www.princetonreview.com/ai-education/intelligent-tutoring-systems

The State Council of the People's Republic of China. (2017a, July 20). *A next generation artificial intelligence development plan* (Creemers, R., Webster, G., Triolo, P., & Kania, E., Trans.). New America. https://d1y8sb8igg2f8e.cloudfront.net/documents/translation-fulltext-8.1.17.pdf (Original work published 2017)

The State Council of the People's Republic of China. (2017b, July 20). *China issues guideline on artificial intelligence development*. https://english.www.gov.cn/policies/latest_releases/2017/07/20/content_281475742458322.htm

The Texas Senate. (2024, June 6). *Senate Committee on Criminal Justice* [Video]. https://tlcsenate.granicus.com/MediaPlayer.php?clip_id=18513

The White House. (2023a). *Ensuring safe, secure, and trustworthy AI.* https://www.whitehouse.gov/wp-content/uploads/2023/07/Ensuring-Safe-Secure-and-Trustworthy-AI.pdf

The White House. (2023b, July 21). *Fact Sheet: Biden-Harris administration secures voluntary commitments from leading artificial intelligence companies to manage the risks posed by AI.* https://www.whitehouse.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/

The White House. (2023c, September 12). *Fact Sheet: Biden-Harris administration secures voluntary commitments from eight additional artificial intelligence companies to manage the risks posed by AI.* https://www.whitehouse.gov/briefing-room/statements-releases/2023/09/12/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-eight-additional-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/

The White House. (2023d, October 30). *Fact Sheet: President Biden issues executive order on safe, secure, and trustworthy artificial intelligence*. 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/

The White House. (2024, July 26). *Fact Sheet: Biden-Harris administration announces new AI actions and receives additional major voluntary commitment on AI*. https://www.whitehouse.gov/briefing-room/statements-releases/2024/07/26/fact-sheet-biden-harris-administration-announces-new-ai-actions-and-receives-additional-major-voluntary-commitment-on-ai/

The White House Office of Science and Technology Policy. (2022). *Blueprint for an AI bill of rights: Making automated systems work for the American people*. https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf

Thomas, L. (2019, May 24). *OECD members, including U.S., back guiding principles to make AI safer*. Reuters. https://www.reuters.com/article/technology/oecd-members-including-us-back-guiding-principles-to-make-ai-safer-idUSKCN1SS1V5/

Thompson, M. (2021, November 1). *AI-enabled ground combat vehicles demonstrate agility and synergy at PC21*. U.S. Army. https://www.army.mil/article/251632/ai_enabled_ground_combat_vehicles_demonstrate_agility_and_synergy_at_pc21

Thomson Reuters. (2024, February 21). *State of the courts report 2024: What do courts think of gen AI?* https://www.thomsonreuters.com/en-us/posts/government/state-of-the-courts-report-2024/

Tierno, P. (2024). *Artificial intelligence and machine learning in financial services*. Congressional Research Service. https://crsreports.congress.gov/product/pdf/R/R47997

Today. (2024, May 21). *Scarlett Johansson says OpenAI used her voice without permission* [Video]. https://www.today.com/video/scarlett-johansson-says-openai-used-her-voice-without-permission-211299909920

Torode, G. (2024, May 2). *US official urges China, Russia to declare only humans, not AI, control nuclear weapons*. Reuters. https://www.reuters.com/world/us-official-urges-china-russia-declare-only-humans-not-ai-control-nuclear-2024-05-02/

Townson, S., Holton, L., & Durman, P. (2024). *How the EU AI Act will impact your business*. Oliver Wyman. https://www.oliverwyman.com/our-expertise/insights/2024/may/how-eu-ai-act-affect-business.html

Trapassi, G., Pincetti, M., Fontanini, G., Falcone, G., Bartolucci, M., Rigoni, C., Boccia, N., Gabbiani, G., Da Silva, I. L., Mascaro, D., Mangini, S., & Papaveri, L. (2021). *Digital banking maturity 2020*. Deloitte. https://www2.deloitte.com/content/dam/Deloitte/it/Documents/financial-services/DBM2020Deloitte.pdf

Tucker, E. (2023, January 19). *FBI chief says he's 'deeply concerned' by China's AI program*. The Associated Press. https://apnews.com/article/technology-science-christopher-wray-beijing-china-3c30625e842b08a834e715230d584847

Tufts, S. (2023, July 6). *Critical infrastructure attacks are ramping up*. Security. https://www.securitymagazine.com/articles/99597-critical-infrastructure-attacks-are-ramping-up

Tung, L. (2018, December 7). *Microsoft: Here's why we need AI facial-recognition laws right now*. ZDNET. https://www.zdnet.com/article/microsoft-heres-why-we-need-ai-facial-recognition-laws-right-now/

Tunguz, B. [@tunguz]. (2024, July 12). *America makes software. Asia makes hardware. Europe makes it difficult* [Post]. X. https://x.com/tunguz/status/1811741275982561599

Turing, A. (1950). Computing machinery and intelligence. *Mind*, *49*(236), 433–460. https://doi.org/10.1093/mind/LIX.236.433

Twomey, J., Ching, D., Aylett, M. P., Quayle, M., Linehan, C., & Murphy, G. (2023). Do deepfake videos undermine our epistemic trust? A thematic analysis of tweets that discuss deepfakes in the Russian invasion of Ukraine. *PLoS One*, *18*(10): e0291668. https://doi.org/10.1371/journal.pone.0291668

Tyson, A., Pasquini, G., Spencer, A., & Funk, C. (2023). *60% of Americans would be uncomfortable with provider relying on AI in their own health care.* Pew Research Center. https://www.pewresearch.org/wp-content/uploads/sites/20/2023/02/PS_2023.02.22_AI-health_REPORT.pdf

U.S. Chamber of Commerce. (2023, October 30). *AI executive order addresses important priorities but needs more work*. https://www.uschamber.com/technology/artificial-intelligence/ai-executive-order-addresses-important-priorities-but-needs-more-work

U.S. Department of Commerce. (2023, October 30). *Department of Commerce to undertake key responsibilities in historic artificial intelligence executive order* [Press release]. https://www.commerce.gov/news/press-releases/2023/10/department-commerce-undertake-key-responsibilities-historic-artificial

U.S. Department of Commerce. (2024, April 29). *Department of Commerce announces new actions to implement President Biden's executive order on AI* [Press release]. https://www.commerce.gov/news/press-releases/2024/04/department-commerce-announces-new-actions-implement-president-bidens

U.S. Department of Defense. (n.d.). *Summary of the 2018 Department of Defense Artificial Intelligence Strategy* . Retrieved August 20, 2024, from https://media.defense.gov/2019/feb/12/2002088963/-1/-1/1/summary-of-dod-ai-strategy.pdf

U.S. Department of Homeland Security. (2020). *Review of CBP's major cybersecurity incident during a 2019 biometric pilot*. https://www.oig.dhs.gov/sites/default/files/assets/2020-09/OIG-20-71-Sep20.pdf

U.S. Department of Homeland Security. (2023a). *Homeland threat assessment 2024*. https://www.dhs.gov/sites/default/files/2023-09/23_0913_ia_23-333-ia_u_homeland-threat-assessment-2024_508C_V6_13Sep23.pdf

U.S. Department of Homeland Security. (2023b, December 1). *Secure cyberspace and critical infrastructure*. Retrieved August 19, 2024, from https://www.dhs.gov/secure-cyberspace-and-critical-infrastructure

U.S. Department of Homeland Security. (2024, April 29). *DHS publishes guidelines and report to secure critical infrastructure and weapons of mass destruction from AI-related threats*. https://www.dhs.gov/news/2024/04/29/dhs-publishes-guidelines-and-report-secure-critical-infrastructure-and-weapons-mass

U.S. Department of State. (2023, November 9). *Political declaration on responsible military use of artificial intelligence and autonomy*. https://www.state.gov/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy-2/

U.S. Department of State. (2024, October 17). *Political declaration on responsible military use of artificial intelligence and autonomy*. Retrieved October 25, 2024, from https://www.state.gov/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy/

U.S. Equal Employment Opportunity Commission. (n.d.). *3. Who is protected from employment discrimination?* Retrieved August 20, 2024, from https://www.eeoc.gov/employers/small-business/3-who-protected-employment-discrimination

UN System Chief Executives Board for Coordination. (n.d.). *Artificial intelligence*. Retrieved August 21, 2024, from https://unsceb.org/topics/artificial-intelligence

United States Courts. (n.d.). *Federal judicial caseload statistics 2023*. Retrieved August 21, 2024, from https://www.uscourts.gov/statistics-reports/federal-judicial-caseload-statistics-2023

USC Libraries. (n.d.). *Artificial superintelligence*. Retrieved October 29, 2024, from https://libraries.usc.edu/events/artificial-superintelligence

University of San Diego. (n.d.). *Artificial intelligence in finance [15 examples]*. Retrieved August 20, 2024, from https://onlinedegrees.sandiego.edu/artificial-intelligence-finance/

University of Virginia. (n.d.). *What the heck is a deepfake?* Retrieved August 19, 2024, from https://security.virginia.edu/deepfakes

van der Schaar, M. (2023, June 26). AI-powered personalised medicine could revolutionise healthcare (and no, we're not putting ChatGPT in charge). *The Guardian*. https://www.theguardian.com/commentisfree/2023/jun/26/ai-personalise-medicine-patient-lab-health-diagnosis-cambridge

Velasquez, S. (2023, July 18). *How AI is bringing film stars back from the dead*. BBC. https://www.bbc.com/future/article/20230718-how-ai-is-bringing-film-stars-back-from-the-dead

Veltman, C. (2024, April 30). *AI is contentious among authors. So why are some feeding it their own writing?* WAMU 88.5. https://wamu.org/story/24/04/30/ai-is-contentious-among-authors-so-why-are-some-feeding-it-their-own-writing/

Vermna, P. (2023, November 5). AI fake nudes are booming. It's ruining real teens' lives. *The Washington Post*. https://www.washingtonpost.com/technology/2023/11/05/ai-deepfake-porn-teens-women-impact/

Vergun, D. (2023, November 22). *U.S. endorses responsible AI measures for global militaries*. U.S. Department of Defense. https://www.defense.gov/News/News-Stories/Article/Article/3597093/us-endorses-responsible-ai-measures-for-global-militaries/

Vincent, B. (2019, May 22). *42 countries agree to international principles for artificial intelligence*. Nextgov/FCW. https://www.nextgov.com/artificial-intelligence/2019/05/42-countries-agree-international-principles-artificial-intelligence/157189/

Volpicelli, G. (2024, February 12). Meet the Vatican's AI mentor. *Politico*. https://www.politico.eu/article/meet-the-vatican-ai-mentor-diplomacy-friar-paolo-benanti-pope-francis/

Waddell, K. (2016, April 22). How license-plate readers have helped police and lenders target the poor. *The Atlantic*. https://www.theatlantic.com/technology/archive/2016/04/how-license-plate-readers-have-helped-police-and-lenders-target-the-poor/479436/

Wadhwani, K. (n.d.). *AI use cases and applications in key industries*. SoluLab. Retrieved August 19, 2024, from https://www.solulab.com/ai-use-cases-and-applications/

Wagman, J., & Nicula-Golovei, T. (2022). The evolution of safeguards technology. *IAEA Bulletin, 63*(3). https://www.iaea.org/bulletin/the-evolution-of-safeguards-technology

Wall, S., & Schellmann, H. (2021, July 7). We tested AI interview tools. Here's what we found. *MIT Technology Review*. https://www.technologyreview.com/2021/07/07/1027916/we-tested-ai-interview-tools/

Watts, C. (2023, September 7). *China, North Korea pursue new targets while honing cyber capabilities*. Microsoft. https://blogs.microsoft.com/on-the-issues/2023/09/07/digital-threats-cyberattacks-east-asia-china-north-korea/

Weber, L. (2023, July 5). New York City starts to regulate AI used in hiring tools. *The Wall Street Journal*. https://www.wsj.com/articles/new-york-city-starts-to-regulate-ai-used-in-hiring-tools-79a2260f

Weber, L. (2024, January 22). New York City passed an AI hiring law. So far, few companies are following it. *The Wall Street Journal*. https://www.wsj.com/business/new-york-city-passed-an-ai-hiring-law-so-far-few-companies-are-following-it-7e31a5b7

Wehner, G. (2024, April 23). *Bipartisan lawmakers seek answers from Mayorkas after Russian cyberattacks on water systems in US*. Fox News. https://www.foxnews.com/politics/bipartisan-lawmakers-seek-answers-mayorkas-russian-cyberattacks-water-systems-us

Weiser, B. (2023, May 27). Here's what happens when your lawyer uses ChatGPT. *The New York Times*. https://www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html

White & Case. (2024, May 13). *AI watch: Global regulatory tracker - United States*. https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-united-states

Whiting, Z. (2023a). *Social media is a harmful product: Texas should prohibit companies from granting access to minors*. Texas Public Policy Foundation. https://www.texaspolicy.com/social-media-is-a-harmful-product-texas-should-prohibit-companies-from-granting-access-to-minors/

Whiting, Z. (2023b, March 20). *Testimony before the Texas House Youth Health & Safety Committee*. Texas Public Policy Foundation. https://www.texaspolicy.com/wp-content/uploads/2023/03/2023-03-BTT-Testimony-HB18.pdf

Whiting, Z. (2023c, April 5). *Margaritaville t-shirts vs. dead kids: How big business tried to kill a kids' online safety bill – in their own words*. Texas Public Policy Foundation. https://www.texaspolicy.com/margaritaville-t-shirts-vs-dead-kids-how-big-business-tried-to-kill-a-kids-online-safety-bill-in-their-own-words/

Whyman, B. (2023, October 10). *AI regulation is coming- What is the likely outcome?* Center for Strategic & International Studies. https://www.csis.org/blogs/strategic-technologies-blog/ai-regulation-coming-what-likely-outcome

Wilding, M. (2023, August 31). *IBM promised to back off facial recognition — then it signed a $69.8 million contract to provide it*. The Verge. https://www.theverge.com/2023/8/31/23852955/ibm-uk-government-contract-biometric-facial-recognition

Williams, E. (2024, May 9). *China's Digital Silk Road taking its shot at the global stage*. East Asia Forum. https://eastasiaforum.org/2024/05/09/chinas-digital-silk-road-taking-its-shot-at-the-global-stage/

Williams, K. (2023, November 7). *Summary: What does Biden's executive order on artificial intelligence actually say?* Electronic Privacy Information Center. https://epic.org/summary-what-does-bidens-executive-order-on-artificial-intelligence-actually-say/

Witherspoon, T. (2024, March 8). *Texas courts facing shortage of court reporters*. KWTX. https://www.kwtx.com/2024/03/08/texas-courts-facing-shortage-court-reporters/

Wong, S., Thorpe, F.V, Nobles, R., & Brown-Kaiser, L. (2023, September 13). *Elon Musk warns of 'civilizational risk' posed by AI in meeting with tech CEOs and senators*. NBC News. https://www.nbcnews.com/politics/congress/big-tech-ceos-ai-meeting-senators-musk-zuckerberg-rcna104738

World Economic Forum. (2023, January 19). *In the name of national security* [Video]. https://www.weforum.org/events/world-economic-forum-annual-meeting-2023/sessions/in-the-name-of-national-security/

Wright, L., Muenster, R. M., Vecchione, B., Qu, T., Cai, P., Smith, A., COMM/INFO 2450 Student Investigators, Metcalf, J., & Matias, J. N. (2024). Null compliance: NYC Local Law 144 and the challenges of algorithm accountability. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)* (pp. 1701–1713). Association for Computing Machinery. https://doi.org/10.1145/3630106.3658998

Wright, R. (2023, December 6). *Artificial intelligence in the states: Emerging legislation*. The Council of State Governments. https://www.csg.org/2023/12/06/artificial-intelligence-in-the-states-emerging-legislation/

Wu, J. (2019, August 1). *AI goes to court: The growing landscape of AI for access to justice*. Justice Innovation. https://justiceinnovation.law.stanford.edu/ai-goes-to-court-the-growing-landscape-of-ai-for-access-to-justice/

Xu, F., Uszkoreit, H., Du, Y., Fan, W., Zhao, D., & Zhu, J. (2019). Explainable AI: A brief survey on history, research areas, approaches and challenges. In J. Tang, M.-Y. Kan, D. Zhao, S. Li, & H. Zan (Eds.), *Natural Language Processing and Chinese Computing*, (pp. 563–574). Springer. https://doi.org/10.1007/978-3-030-32236-6_51

Yang, A. (2024, June 25). *U.S. record labels are suing AI music generators, alleging copyright infringement*. NBC News. https://www.nbcnews.com/tech/tech-news/us-record-labels-are-suing-ai-music-generators-alleging-copyright-infr-rcna158660

Yang, A., & Hamedy, S. (2024, April 26). *Drake pulls 'Taylor Made Freestyle' after Tupac estate threatens action for apparent use of AI voice*. NBC News. https://www.nbcnews.com/pop-culture/pop-culture-news/drake-pulls-taylor-made-freestyle-tupac-estate-threatens-action-appare-rcna149592

Yasar, K. (2023). *Image recognition*. TechTarget. https://www.techtarget.com/searchenterpriseai/definition/image-recognition

Yehoshua, R. (2023, August 14). *Why detection and response technology won't solve all ransomware attacks*. SC Magazine. https://www.scmagazine.com/perspective/why-detection-and-response-technology-wont-solve-all-ransomware-attacks

Zakrzewski, C. (2024, May 15). Senators studied AI for a year. Critics call the result 'pathetic.' *The Washington Post*. https://www.washingtonpost.com/technology/2024/05/15/congress-ai-road-map-regulation-schumer/

ZeroEyes. (n.d.). [Home page]. Retrieved August 20, 2024, from https://zeroeyes.com/

Zetoony, D. (2023, August 18). *Understanding AI terms: What is output data?* Greenberg Traurig. https://www.gtlaw-dataprivacydish.com/2023/08/understanding-ai-terms-what-is-output-data/

Zhou, L. (2024, February 4). Russia and China compare notes on 'military use of artificial intelligence.' *South China Morning Post*. https://www.scmp.com/news/china/diplomacy/article/3250893/russia-and-china-compare-notes-military-use-artificial-intelligence

APPENDIX A: KEY TERMS

Proviso. These are not necessarily definitions the Foundation endorses. Rather, this is meant as a resource guide to ensure lawmakers are apprised on essential terms in the realm of AI technology and regulation for legislative scope, precision, applicability, and enforceability. The authors first consulted the Texas code for relevant definitions; second, legislation and laws in other jurisdictions; and finally, terms as used by researchers, industry, and other key stakeholders.

Term	Sample definition or framework
Algorithm	"'Algorithm' means a computerized procedure consisting of a set of steps used to accomplish a determined task" (Texas Government Code, Sec. 2054.621(1)). "A clearly specified mathematical process for computation; a set of rules that, if followed, will give a prescribed result" (Barker, 2020, p. 3).
Algorithmic discrimination	"Algorithmic discrimination occurs when automated systems contribute to unjustified different treatment or impacts disfavoring people based on their race, color, ethnicity, sex (including pregnancy, childbirth, and related medical conditions, gender identity, intersex status, and sexual orientation), religion, age, national origin, disability, veteran status, genetic information, or any other classification protected by law" (The White House Office of Science and Technology Policy, 2022, p. 5).
Artificial intelligence bias	"AI bias, also called machine learning bias or algorithm bias, refers to the occurrence of biased results due to human biases that skew the original training data or AI algorithm—leading to distorted outputs and potentially harmful outcomes" (Holdsworth, 2023, para. 1).
Artificial general intelligence	"Artificial general intelligence (AGI) refers to the hypothetical intelligence of a machine that possesses the ability to understand or learn any intellectual task that a human being can. It is a type of artificial intelligence (AI) that aims to mimic the cognitive abilities of the human brain" (Google Cloud, n.d.-c, para. 1).
Artificial intelligence	"A term coined by emeritus Stanford Professor John McCarthy in 1955, was defined by him as 'the science and engineering of making intelligent "machines" (Manning, 2020, para. 2). "AI is a machine's ability to perform the cognitive functions we associate with human minds, such as perceiving, reasoning, learning, interacting with the environment, problem-solving, and even exercising creativity" (McKinsey & Company, 2024, para. 4).

Artificial intelligence system	“‘Artificial intelligence systems’ means systems capable of: (A) perceiving an environment through data acquisition and processing and interpreting the derived information to take an action or actions or to imitate intelligent behavior given a specific goal; and (B) learning and adapting behavior by analyzing how the environment is affected by prior actions” (Texas Government Code, Sec. 2054.621(2)). “An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment” (OECD, n.d., “AI Terms & Concepts” section). “The AI RMF refers to an AI system as an engineered or machine-based system that can, for a given set of objectives, generate outputs such as predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy” (NIST, 2023a, p. 1). “‘AI system’ means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments” (Regulation (EU) 2024/1689, Article 3(1), 2024). “‘Artificial intelligence system’ means any machine-based system that, for any explicit or implicit objective, infers from the inputs the system received how to generate outputs, including content, decisions, predictions, or recommendations, that can influence physical or virtual environments” (SB 24-205 § 6-1-1701(2), 2024).
Artificial superintelligence	“Artificial superintelligence (ASI) is a hypothetical software-based artificial intelligence (AI) system with an intellectual scope beyond human intelligence. At the most fundamental level, this superintelligent AI has cutting-edge cognitive functions and highly developed thinking skills more advanced than any human” (Mucci & Stryker, 2023, para. 1).
Automated decision system	“‘Automated decision system’ means an algorithm, including an algorithm incorporating machine learning or other artificial intelligence techniques, that uses data-based analytics to make or support governmental decisions, judgments, or conclusions” (Texas Government Code, Sec. 2054.621(3)).
Automated final decision system	“‘Automated final decision system’ means an automated decision system that makes final decisions, judgments, or conclusions without human intervention” (Texas Government Code, Sec. 2054.621(4)).

Biometric data	"'Biometric data' means data generated by automatic measurements of an individual's biological characteristics. The term includes a fingerprint, voiceprint, eye retina or iris, or other unique biological pattern or characteristic that is used to identify a specific individual. The term does not include a physical or digital photograph or data generated from a physical or digital photograph, a video or audio recording or data generated from a video or audio recording, or information collected, used, or stored for health care treatment, payment, or operations under the Health Insurance Portability and Accountability Act of 1996 (42 U.S.C. Section 1320d et seq.)" (Texas Business and Commerce Code, Sec. 541.001(3)). "'Biometric data' means data generated by automatic measurements of an individual's biological patterns or characteristics, including fingerprint, voiceprint, retina or iris scan, information pertaining to an individual's DNA, or another unique biological pattern or characteristic that is used to identify a specific individual" (Texas Business and Commerce Code, Sec. 509.001(1)). "'Biometric data' means personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, such as facial images or dactyloscopic data" (Regulation (EU) 2024/1689, Article 3(34), 2024). "'Biometric identification' means the automated recognition of physical, physiological, behavioural, or psychological human features for the purpose of establishing the identity of a natural person by comparing biometric data of that individual to biometric data of individuals stored in a database" (Regulation (EU) 2024/1689, Article 3(35), 2024). "'Biometric verification' means the automated, one-to-one verification, including authentication, of the identity of natural persons by comparing their biometric data to previously provided biometric data" (Regulation (EU) 2024/1689, Article 3(36), 2024).
Black box problem	The black box "refers to an AI system whose internal workings or decision-making processes are opaque or not easily understandable to humans. In other words, when you put data into the black box AI system and get outputs in return, you don't really know how the AI arrived at the conclusions or decisions that it presents to you" (Kelly, 2024, para. 3).
Chatbot	"A chatbot is a computer program that simulates human conversation with an end user. Not all chatbots are equipped with artificial intelligence (AI), but modern chatbots increasingly use conversational AI techniques such as natural language processing (NLP) to understand user questions and automate responses to them" (IBM, n.d.-e, para. 1). "The term 'artificial intelligence chatbot' means generative artificial intelligence system with which users can interact by or through an interface that approximates or simulates conversation" (S. 2691, 2023, Sec. 2(c)(2)).
Closed source	"Closed-source refers to a software development approach where an application's source code is proprietary and not publicly available" (Deloitte, 2023, para. 3). "When all aspects and components of a system are inaccessible outside the developer organization, or even closed outside a specific subsection of an organization, the system is fully closed" (Solaiman, 2023, p. 5).
Computer vision	"An interdisciplinary scientific field that deals with how computers can be made to gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to automate tasks that the human visual system can do" (CompTIA, n.d., p. 2).

Confabulation (Hallucination)	“‘Confabulation’ refers to a phenomenon in which GAI systems generate and confidently present erroneous or false content in response to prompts. Confabulations also include generated outputs that diverge from the prompts or other input or that contradict previously generated statements in the same context. These phenomena are colloquially also referred to as ‘hallucinations’ or ‘fabrications’” (NIST, 2024a, p. 6).
Consent (Informed consent)	“‘Consent,’ when referring to a consumer, means a clear affirmative act signifying a consumer ’s freely given, specific, informed, and unambiguous agreement to process personal data relating to the consumer. The term includes a written statement, including a statement written by electronic means, or any other unambiguous affirmative action. The term does not include: (A) acceptance of a general or broad terms of use or similar document that contains descriptions of personal data processing along with other, unrelated information; (B) hovering over, muting, pausing, or closing a given piece of content; or (C) agreement obtained through the use of dark patterns” (Texas Business and Commerce Code, Sec. 541.001(6)). “‘[I]nformed consent’ means a subject’s freely given, specific, unambiguous and voluntary expression of his or her willingness to participate in a particular testing in real-world conditions, after having been informed of all aspects of the testing that are relevant to the subject’s decision to participate” (Regulation (EU) 2024/1689, Article 3(59), 2024).
Consequential decision (Decision that produces a legal or similarly significant effect concerning a consumer)	“‘Consequential decision’ means a decision that has a material legal or similarly significant effect on the provision or denial to any consumer of, or the cost or terms of: education enrollment or an education opportunity; employment or an employment opportunity; a financial or lending service; an essential government service; health-care services; housing; insurance; or a legal service” (SB 24-205, 2024, § 6-1-1701(3)). “‘Decision that produces a legal or similarly significant effect concerning a consumer’ means a decision made by the controller that results in the provision or denial by the controller of: (A) financial and lending services; (B) housing, insurance, or health care services; (C) education enrollment; (D) employment opportunities; (E) criminal justice; or (F) access to basic necessities, such as food and water” (Texas Business and Commerce Code, Sec. 541.001(11)).
Consumer	“‘Consumer’ means an individual who enters into a transaction primarily for personal, family, or household purposes” (Texas Business and Commerce Code, Sec. 1.201(11)).
Controller	“‘Controller’ means an individual or other person that, alone or jointly with others, determines the purpose and means of processing personal data” (Texas Business and Commerce Code, Sec. 541.001(8)).
Data provenance	“A process that tracks and logs the history and origin of records in a dataset, encompassing the entire life cycle from its creation and collection to its transformation to its current state. It includes information about sources, processes, actors and methods used to ensure data integrity and quality. Data provenance is essential for data transparency and governance, and it promotes better understanding of the data and eventually the entire AI system” (IAPP, 2024, p. 4).

Deep learning	“Deep learning is an artificial intelligence function that imitates the workings of the human brain in processing data and creating patterns for use in decision making. Deep learning is a subset of machine learning in AI that has networks capable of learning unsupervised from data that is unstructured or unlabeled. Also known as deep neural learning or deep neural network” (CompTIA, n.d., p. 2). “Deep learning is a subset of machine learning that uses multilayered neural networks, called deep neural networks, to simulate the complex decision-making power of the human brain. Some form of deep learning powers most of the artificial intelligence (AI) applications in our lives today” (IBM, n.d.-f, para. 1).
Deepfake	“‘[D]eep fake’ means AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places, entities or events and would falsely appear to a person to be authentic or truthful” (Regulation (EU) 2024/1689, Article 3(60), 2024). The current definition in Texas law only speaks to “deep fake video”—“a video, created with the intent to deceive, that appears to depict a real person performing an action that did not occur in reality” (Texas Penal Code, Sec. 21.165(1); Texas Election Code, Sec. 255.004(e)). This definition needs to be expanded considerably to 1) capture video, image, voice, etc. and 2) to include more than a depiction of a real person, including depictions that are “indistinguishable” from a real person (SB 79, 2024, §§ 1(5)(c), 1(8)).
Deploy(er) (Use)	“‘Deployer’ means a natural or legal person, public authority, agency or other body using an AI system under its authority except where the AI system is used in the course of a personal non-professional activity” (Regulation (EU) 2024/1689, Article 3(4), 2024). “‘Deployer’ means a person doing business in this state that deploys a high-risk artificial intelligence system” (SB 24-205, 2024, § 6-1-1701(6)). “‘Deploy’ means to use a high-risk artificial intelligence system” (SB 24-205, 2024, § 6-1-1701(5)).
Developer (Provider)	“‘Developer’ means a person doing business in this state that develops or intentionally and substantially modifies an artificial intelligence system” (SB 24-205, 2024, § 6-1-1701(6)). “‘Provider’ means a natural or legal person, public authority, agency or other body that develops an AI system or a general-purpose AI model or that has an AI system or a general-purpose AI model developed and places it on the market or puts the AI system into service under its own name or trademark, whether for payment or free of charge” (Regulation (EU) 2024/1689, Article 3(3), 2024).
Explainability	“Explainability refers to a representation of the mechanisms underlying AI systems’ operation” (NIST, 2023a, p. 16). “Explainable artificial intelligence (XAI) is a set of processes and methods that allows human users to comprehend and trust the results and output created by machine learning algorithms” (IBM, n.d.-g, para. 1).
General-purpose AI model	“‘General-purpose AI model’ means an AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market” (Regulation (EU) 2024/1689, Article 3(63), 2024).

General-purpose AI system	"'General-purpose AI system' means an AI system which is based on a general-purpose AI model and which has the capability to serve a variety of purposes, both for direct use as well as for integration in other AI systems" (Regulation (EU) 2024/1689, Article 3(66), 2024).
Generative artificial intelligence	"Generative AI refers to deep-learning models that can generate high-quality text, images, and other content based on the data they were trained on" (Martineau, 2023, para. 1). "EO 14110 defines Generative AI as 'the class of AI models that emulate the structure and characteristics of input data in order to generate derived synthetic content. This can include images, videos, audio, text, and other digital content.' While not all GAI is derived from foundation models, for purposes of this document, GAI generally refers to generative foundation models. The foundation model subcategory of 'dual-use foundation models' is defined by EO 14110 as 'an AI model that is trained on broad data; generally uses self-supervision; contains at least tens of billions of parameters; is applicable across a wide range of contexts'" (NIST, 2024a, p. 1, fn. 1; Exec. Order No. 14110, 2023, pp. 75194-75195).
Hallucination (Confabulation)	"'Confabulation' refers to a phenomenon in which GAI systems generate and confidently present erroneous or false content in response to prompts. Confabulations also include generated outputs that diverge from the prompts or other input or that contradict previously generated statements in the same context. These phenomena are colloquially also referred to as 'hallucinations' or 'fabrications'" (NIST, 2024a, p. 6).
High-risk artificial intelligence system	"'High-risk artificial intelligence system' means any artificial intelligence system that, when deployed, makes, or is a substantial factor in making, a consequential decision" (SB 24-205, 2024, § 6-1-1701(9)(a)).
Human-in-the-loop	"The term human-in-the-loop (HITL) generally refers to the need for human interaction, intervention, and judgment to control or change the outcome of a process, and it is a practice that is being increasingly emphasized in machine learning, generative AI, and the like" (Meng, 2023, para. 1).
Image recognition	"Image recognition, in the context of machine vision, is the ability of software to identify objects, places, people, writing and actions in digital images. Computers can use machine vision technologies in combination with a camera and artificial intelligence (AI) software to achieve image recognition" (Yasar, 2023, para. 1).
Informed consent (Consent)	"'Informed consent' means a subject's freely given, specific, unambiguous and voluntary expression of his or her willingness to participate in a particular testing in real-world conditions, after having been informed of all aspects of the testing that are relevant to the subject's decision to participate" (Regulation (EU) 2024/1689, Article 3(59), 2024). "'Consent,' when referring to a consumer, means a clear affirmative act signifying a consumer 's freely given, specific, informed, and unambiguous agreement to process personal data relating to the consumer. The term includes a written statement, including a statement written by electronic means, or any other unambiguous affirmative action. The term does not include: (A) acceptance of a general or broad terms of use or similar document that contains descriptions of personal data processing along with other, unrelated information; (B) hovering over, muting, pausing, or closing a given piece of content; or (C) agreement obtained through the use of dark patterns" (Texas Business and Commerce Code, Sec. 541.001(6)).

Intended purpose	"'Intended purpose' means the use for which an AI system is intended by the provider, including the specific context and conditions of use, as specified in the information supplied by the provider in the instructions for use, promotional or sales materials and statements, as well as in the technical documentation" (Regulation (EU) 2024/1689, Article 3(12), 2024).
Interpretability	"Interpretability refers to the meaning of AI systems' output in the context of their designed functional purposes" (NIST, 2023a, p. 1).
Input data	"'Input data' means data provided to or directly acquired by an AI system on the basis of which the system produces an output" (Regulation (EU) 2024/1689, Article 3(33), 2024).
Kill switch	"A kill switch in an IT context is a mechanism used to shut down or disable a device or program. The purpose of a kill switch is usually to prevent theft of a machine or data or shut down machinery in an emergency. The degree to which a kill switch limits, alters or stops an action or activity depends on the production, process or program it is intended to protect" (TechTarget, 2021, paras. 1-2).
Large language model	"Large language models (LLMs) are a category of foundation models trained on immense amounts of data making them capable of understanding and generating natural language and other types of content to perform a wide range of tasks" (IBM, n.d.-h, para. 1). "Large language models (LLM) are very large deep learning models that are pre-trained on vast amounts of data. The underlying transformer is a set of neural networks that consist of an encoder and a decoder with self-attention capabilities. The encoder and decoder extract meanings from a sequence of text and understand the relationships between words and phrases in it" (AWS, n.d.-c, para. 1).
Machine learning	"Machine learning is a branch of AI that allows systems to automatically process data and analyze for insights without being programmed explicitly. Machine learning is concerned with learning functions and patterns to do things like classification and prediction" (CompTIA, n.d., p. 2).
Neural network	"A neural network is a machine learning program, or model, that makes decisions in a manner similar to the human brain, by using processes that mimic the way biological neurons work together to identify phenomena, weigh options and arrive at conclusions" (IBM, n.d.-i, para. 1). "A neural network is a method in artificial intelligence that teaches computers to process data in a way that is inspired by the human brain. It is a type of machine learning (ML) process, called deep learning, that uses interconnected nodes or neurons in a layered structure that resembles the human brain. It creates an adaptive system that computers use to learn from their mistakes and improve continuously" (AWS, n.d.-d, para. 1).

Open source	“Open-source refers to an approach to software development where all or part of an application’s source code is released openly to the general public” (Deloitte, 2023, “What is Open-Source?” section). “An Open Source AI is an AI system made available under terms and in a way that grant the freedoms to: • Use the system for any purpose and without having to ask for permission. • Study how the system works and inspect its components. • Modify the system for any purpose, including to change its output. • Share the system for others to use with or without modifications, for any purpose. These freedoms apply both to a fully functional system and to discrete elements of a system. A precondition to exercising these freedoms is to have access to the preferred form to make modifications to the system” (Open Source Initiative, n.d., “What Is Open Source AI” section).
Operator	“‘Operator’ means a provider, product manufacturer, deployer, authorised representative, importer or distributor” (Regulation (EU) 2024/1689, Article 3(8), 2024).
Opt-in	“Opt-in is to give permission or accept something. In other words, it is an affirmative action of giving or asking for user consent” (CookieYes, 2024, “Opt-in” section). “An opt-in regime sets the default rule such that an entity must obtain the consent of a consumer before performing processing activities” (Rippy, 2021, “Strict Opt-in” section).
Opt-out	“Opt-out is to refuse permission or cancel something. In other words, it is an act of refusing or withdrawing consent in response to a particular event or process” (CookieYes, 2024, “Opt-out” section).
Output data	“Output data is new data an artificial intelligence (AI) creates or synthesizes based on input data and the AI’s algorithm” (Zetoony, 2023, para. 2).
Pattern recognition	“Pattern recognition is the method of using computer algorithms to analyze, detect, and label regularities in data. This informs how the data gets classified into different categories” (Coursera, 2024, “Pattern Recognition” section). “Pattern recognition is the label given to the activity of machines detecting patterns from data. It is often used synonymously with machine learning” (CompTIA, n.d., p. 2).
Predictive analytics	“Predictive analytics is the process of using data to forecast future outcomes. The process uses data analysis, machine learning, artificial intelligence, and statistical models to find patterns that might predict future behavior” (IBM, n.d.-j, para. 1).
Process(ing)	“‘Process’ or ‘processing’ means an operation or set of operations performed, whether by manual or automated means, on personal data or on sets of personal data, such as the collection, use, storage, disclosure, analysis, deletion, or modification of personal data” (Texas Business and Commerce Code, Sec. 541.001(22)). “‘Processing’ means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organization [sic], structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction” (Regulation (EU) 2016/679, Article 4(2)).

Processor	"'Processor' means a person that processes personal data on behalf of a controller" (Texas Business and Commerce Code, Sec. 541.001(23)).
Profiling	"'Profiling' means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements" (Regulation (EU) 2016/679, Article 4(4)).
Provider (Developer)	"'Provider' means a natural or legal person, public authority, agency or other body that develops an AI system or a general-purpose AI model or that has an AI system or a general-purpose AI model developed and places it on the market or puts the AI system into service under its own name or trademark, whether for payment or free of charge" (Regulation (EU) 2024/1689, Article 3(3), 2024). "'Developer' means a person doing business in this state that develops or intentionally and substantially modifies an artificial intelligence system"(SB 24-205, 2024, § 6-1-1701(6)).
Reasonably foreseeable misuse	"'Reasonably foreseeable misuse' means the use of an AI system in a way that is not in accordance with its intended purpose, but which may result from reasonably foreseeable human behaviour or interaction with other systems, including other AI systems" (Regulation (EU) 2024/1689, Article 3(13), 2024).
Risk	"'Risk' means the combination of the probability of an occurrence of harm and the severity of that harm" (Regulation (EU) 2024/1689, Article 3(2), 2024). "Risk refers to the composite measure of an event's probability of occurring and the magnitude or degree of the consequences of the corresponding event. The impacts, or consequences, of AI systems can be positive, negative, or both and can result in opportunities or threats. … When considering the negative impact of a potential event, risk is a function of 1) the negative impact, or magnitude of harm, that would arise if the circumstance or event occurs and 2) the likelihood of occurrence" (NIST, 2023a, p. 4).
Sandbox (AI regulatory sandbox)	"'AI regulatory sandbox' means a controlled framework set up by a competent authority which offers providers or prospective providers of AI systems the possibility to develop, train, validate and test, where appropriate in real-world conditions, an innovative AI system, pursuant to a sandbox plan for a limited time under regulatory supervision" (Regulation (EU) 2024/1689, Article 3(55), 2024).
Serious incident	"'Serious incident' means an incident or malfunctioning of an AI system that directly or indirectly leads to any of the following: (a) the death of a person, or serious harm to a person's health; (b) a serious and irreversible disruption of the management or operation of critical infrastructure; (c) the infringement of obligations under [European] Union law intended to protect fundamental rights; (d) serious harm to property or the environment" (Regulation (EU) 2024/1689, Article 3(49), 2024).
Structured data	Structured data is well-organized, easily interpretable (usually quantitative) data that is organized into rows and columns within databases, or labeled such that machine learning algorithms can clearly identify, consistent input data. For instance, structured data can include numerical values, categorically labeled photos or video, and timestamps, all of which can be used to train models to recognize patterns, make predictions, and improve decision-making processes (IBM, 2021; AWS, n.d.-a).

Substantial factor	"'Substantial factor' means a factor that: (I) assists in making a consequential decision; (II) is capable of altering the outcome of a consequential decision; and (III) is generated by an artificial intelligence system. 'Substantial factor' includes any use of an artificial intelligence system to generate any content, decision, prediction, or recommendation concerning a customer that is used as a basis to make a consequential decision concerning the customer" (SB 24-205 § 6-1-1701(11), 2024).
Supervised learning	"Supervised learning, also known as supervised machine learning, is a subcategory of machine learning and artificial intelligence. It is defined by its use of labeled data sets to train algorithms that to classify data or predict outcomes accurately. As input data is fed into the model, it adjusts its weights until the model has been fitted appropriately, which occurs as part of the cross validation process" (IBM, n.d.-k, paras. 1-2).
Systemic risk	"'Systemic risk' means a risk that is specific to the high-impact capabilities of general-purpose AI models, having a significant impact on the Union market due to their reach, or due to actual or reasonably foreseeable negative effects on public health, safety, public security, fundamental rights, or the society as a whole, that can be propagated at scale across the value chain" (Regulation (EU) 2024/1689, Article 3(65), 2024).
Testing data	"'Testing data' means data used for providing an independent evaluation of the AI system in order to confirm the expected performance of that system before its placing on the market or putting into service" (Regulation (EU) 2024/1689, Article 3(32), 2024).
Training data	"'Training data' means data used for training an AI system through fitting its learnable parameters" (Regulation (EU) 2024/1689, Article 3(29), 2024).
Turing test	"A test of a machine's ability to exhibit intelligent behavior equivalent to or indistinguishable from that of a human. Alan Turing (1912-1954) originally thought of the test to be an AI's ability to converse through a written text, such that a human reader would not be able to tell a computer-generated response from that of a human" (IAPP, 2024, p. 11).
Unstructured data	Unstructured data, typically characterized as qualitative data, is defined as information that lacks a predefined format or organizational structure, making it more challenging to analyze and process compared to structured data. Examples of unstructured data include text, images, videos, and emails. Experts estimate that 80-90% of data is unstructured (IBM, 2021; AWS, n.d.-a).
Unsupervised learning	"A subset of machine learning in which the model is trained by looking for patterns in an unclassified dataset with minimal human supervision. The AI is provided with preexisting unlabeled datasets and then analyzes those datasets for patterns" (IAPP, 2024, p. 11). "Unsupervised learning in artificial intelligence is a type of machine learning that learns from data without human supervision. Unlike supervised learning, unsupervised machine learning models are given unlabeled data and allowed to discover patterns and insights without any explicit guidance or instruction" (Google Cloud, n.d.-d, para. 1).

Voice cloning	“Voice cloning is the process of creating a synthetic replica of a person’s voice using artificial intelligence and machine learning techniques. … Voice cloning builds a digital copy of a person’s unique voice, including speech patterns, accents, voice inflection, and even breathing, by training an algorithm with a sample of a person’s speech” (Amod, 2024, paras. 1, 15).
Voice recognition (Speech recognition)	“Voice recognition is a deep learning technique used to identify, distinguish, and authenticate a particular person’s voice. It evaluates an individual’s unique voice biometrics, including frequency and flow of pitch, and natural accent. Although the terms “voice recognition” and “speech recognition” are often used interchangeably, they are distinct: Speech recognition recognizes spoken words; voice recognition identifies the speaker” (Arm, n.d., para. 1). “Speech recognition, also known as automatic speech recognition (ASR), computer speech recognition or speech-to-text, is a capability that enables a program to process human speech into a written format. While speech recognition is commonly confused with voice recognition, speech recognition focuses on the translation of speech from a verbal format to a text one whereas voice recognition just seeks to identify an individual user’s voice” (IBM, n.d.-l, paras. 1-2).
Watermark	“AI watermarking is the process of embedding a recognizable, unique signal into the output of an artificial intelligence model, such as text or an image, to identify that content as AI generated. That signal, known as a watermark, can then be detected by algorithms designed to scan for it. Ideally, an AI watermark should be invisible to the naked eye, but extractable using specialized software or algorithms” (Craig, 2023, paras. 1-2).

ABOUT THE AUTHORS

The Honorable Zach Whiting is Policy Director and Senior Fellow for Better Tech for Tomorrow at the Texas Public Policy Foundation.

Prior to joining the Foundation, he served as a state senator in his native state of Iowa. He served as Assistant Majority Leader, chair of the Labor and Business Relations Committee, and vice chair of the Administrative Rules Review Committee. Prior to the senate, Zach worked as a Legislative Assistant and Policy Advisor to a member of Congress. He graduated summa cum laude with a B.A. in political science from Stetson University and earned a J.D. from the Regent University School of Law.

David Dunmoyer is the campaign director for Better Tech for Tomorrow and water policy at the Texas Public Policy Foundation. In this role, he publishes research and commentary, provides expert testimony, and advocates for responsible technology and water policy in the Texas legislature. His portfolio includes artificial intelligence, data privacy, cybersecurity, critical infrastructure protection, national security and tech, and other emerging technology issues.

Dunmoyer serves as a committee member of the American Society of Civil Engineer's Water Infrastructure Security Enhancements (WISE) Stand, where he works alongside thought leaders and policymakers to address security for water infrastructure nationwide.

Dunmoyer has been with the Foundation for more than four years, previously serving as Chief of Staff to the executive team and CEOs Kevin Roberts and Greg Sindelar. He joined the Foundation after working for Republican leadership in Congress and as a public affairs professional. David received undergraduate degrees at Texas Christian University and graduated summa cum laude with a Master of Public Affairs from the University of Texas at Austin's LBJ School of Public Affairs.

Texas Public
POLICY FOUNDATION

901 Congress Avenue | Austin, Texas 78701 | (512) 472-2700 | www.TexasPolicy.com