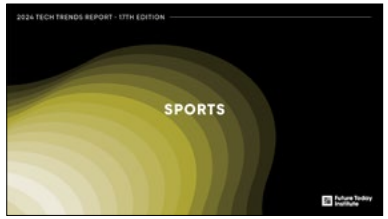
ARTIFICIAL INTELLIGENCE

FUTURE TODAY INSTITUTE'S 2024 TECH TREND REPORT

Our 2024 edition includes nearly 700 trends, which are published individually in 16 volumes and as one comprehensive report with all trends included.

Download all sections of Future Today Institute's 2024 Tech Trends report at **http://www.futuretodayinstitute.com/trends.**



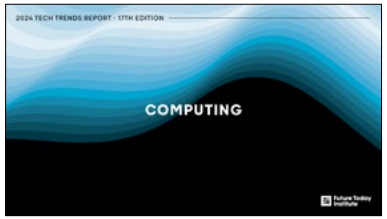


TABLE OF CONTENTS

TABLE OF CONTENTS

YOUR GUIDE TO THE FUTURE OF AI

For two decades, our commitment at Future Today Institute to understanding and leveraging artificial intelligence has been unwavering—even as general excitement about AI has wavered considerably. We've watched interest ebb and flow across industries, among executive leadership and boards of directors, and with investors, legislators, and academia.

Today, we're at a crucial inflection point in AI's developmental journey. This moment isn't marked by a single technological breakthrough but rather by a development that at first may seem less intuitive. What changed in the past year is our perception of what AI is and how it will change everyday life. AI models are now accessible to consumers and businesses alike, so their value can be immediately understood. What's followed: investment, new partnerships, and the grand expansion of value networks.

This rapid escalation in activity has left leaders feeling caught off guard, prompting an urgent need for strategic decision-making. In our conversations with clients and partners, a common theme has emerged: Leaders, understandably concerned about missing out on the next wave of innovation, need clarity about a complex area of technology that will continue to evolve for many years to come.

Recognizing this, we've reimagined our approach for the AI section of our 2024 Tech Trends report. By actively listening to leaders and experts in our network, we've curated the most common questions we're hearing from our clients and grouped our AI trends accordingly. Our goal is to guide leaders through a thoughtful exploration of these questions, enabling a deeper understanding of the implications.

While we don't claim to hold all the answers—every organization's journey with AI will be somewhat unique—what we do offer is a foundation of extensive research and insightful, strategic analysis. You will encounter questions that mirror your own, as well as those you may not have considered but will likely confront in the coming year. Our aim is to equip you with the insights necessary to navigate the upcoming impacts on your organization.

We are confident that this report will serve as an invaluable tool for leaders looking to identify strategic opportunities, achieve competitive advantages, and enhance organizational resilience in the foreseeable future.

Welcome to your guide on the future of AI.

Amy Webb
Chief Executive Officer

Sam Jordan
Manager and Advanced Computing Practice Lead

TOP HEADLINES

This year, the AI landscape could undergo a significant consolidation. Strategic investments, groundbreaking innovations, and regulatory maneuvers will further empower a select group of power brokers, intensifying the competitive dynamics and shaping the trajectory of global AI dominance.

01 OpenAI Seeks $7 Trillion Investment

OpenAI's Sam Altman wants to revamp the semiconductor sector with trillions in investment, targeting global chip capacity to boost AI growth, engaging with investors like the United Arab Emirates.

02 Mainstream Multimodality

For the first time, the public can interact with advanced multimodal AI models capable of understanding and generating various media types, including images and videos. This innovation is akin to human learning processes, enabling AI to learn from visual and auditory information, not just text—just like us.

03 Nvidia GPUs in High Demand for AI Training

The rush to acquire Nvidia's powerful GPUs for AI model training has intensified, as everyone from tech giants to startups seeks the computational horsepower these units offer for advanced machine learning tasks.

04 Open-Source AI Uprising

Meta sparked an open-source large language model movement by releasing the weights for LLaMA, enabling researchers to freely build off the model, fine-tune it, and create customized versions.

05 US Tightens Grip on AI

The US intensifies its stance on AI by restricting access to crucial enabling technologies like semiconductors. In a strategic move, the US also pressures allies to implement similar restrictions against China and Russia.

06 EU AI Act Sets New Global Standard

The EU introduces the first-ever comprehensive AI regulation and a European AI Office, aiming to ensure safety and respect for fundamental rights, while encouraging innovation and investment in AI technologies across Europe. Yet, concerns arise about stifling innovation and the high costs for businesses to comply.

STATE OF PLAY

AI embeds into everything, transforming how we interface with computers while researchers work to make AI more capable and efficient. This sprint toward progress unfolds amid rising geopolitical tensions, as dominance of these strategic technologies reshapes global power dynamics.

The past year marked a watershed moment for artificial intelligence. Central to this transformation are the leaps in large language models (LLMs) and their practical applications, which have not only advanced the frontiers of AI but have also catalyzed a broader integration of AI technologies into everyday life. AI promises revolutionary improvements in health care and life sciences: Now that we've cracked the code on protein structures, an unimaginable number of new therapeutics are on the horizon, along with alternatives to address climate change. In the coming year, AI's reach will extend to people, pets, and objects alike, paving the way for a very near-future in which digital assistants, automated systems, and spatial awareness are seamless, ubiquitous, and invisible. In parallel, the advancements in robotics, both hard and soft, are pushing the boundaries of automation and human-machine interaction.

At the same time, AI's energy demands pose a paradox, offering climate solutions but also contributing to carbon emissions, a concern in energy-constrained areas. The persistent talent gap in AI, particularly in data science, limits its application in critical sectors like agriculture and health care. Political engagement with AI is growing, which is good, but the end result has been a slew of competing policies. Regulatory compliance and enforcement remains a challenge, since depending on the country's position, they promise to both accelerate and curtail the deployment of AI systems. The unequal distribution of AI advancements risks deepening global inequalities, with the global south facing significant disadvantages. Amid all these developments is the persistent geopolitical tensions between China and the West.

For many, generative AI (genAI) is the first entry point into this new reality, which explains the explosive growth we've seen in the past year. Our analysis reveals a

STATE OF PLAY

surge in experimentation with genAI tools across various sectors, indicating a transformative shift toward embracing AI's potential to innovate and streamline operations. This widespread interest in genAI spans a broad spectrum of stakeholders, from business leaders to frontline workers, highlighting just how pervasive the first generation of tools has become.

But here's the thing: GenAI isn't all of AI. Often, when people talk about "AI" what they really mean is "automation." Artificial intelligence is an umbrella term that encompasses many different techniques, models, and frameworks that make up the field. AI's aim is to create intelligent machines that can sense, reason, act, and adapt like humans do, or in ways that go beyond our capabilities. Today, cars can park themselves, while emerging platforms are capable of having seemingly natural conversations. Now, AI is evolving to have beyond-human capabilities. It has invented new drugs, predicted the real-time movement of wildfires, and autonomously designed machine parts.

Developing AI requires extraordinary resources, which is why consolidation among the tech giants is tightening. The biggest names in AI—OpenAI, DeepMind, Anthropic—are increasingly hitched to the world's biggest hyperscalers and cloud providers (Microsoft, Google, Amazon). Venture capital and private equity are still flooding into startups and mature companies alike, and now, sovereign wealth funds have a seat at the table.

AI is magical, but it isn't magic. As long as expectations are tempered, this should be an era of significant innovation, experimentation, and growth, especially as AI propels growth in other areas of science and technology. We are cautiously optimistic about what's on the horizon.

KEY EVENTS

JANUARY 12, 2023

AI Breakthrough in Lung Cancer Detection

A joint effort by MIT and Mass General Hospital yields a significant advancement in lung cancer prognosis with the creation of a deep-learning model that assesses lung cancer risk from CT scans, potentially enhancing early detection and saving numerous lives.

JANUARY 26, 2023

High-Fidelity Music from Text

Google Research's MusicLM introduces a transformative approach to generating detailed music from text descriptions, achieving unprecedented audio quality and text adherence in the field.

FEBRUARY 6, 2023

Google Unveils Bard

Google introduces Bard, an innovative AI chatbot powered by its language model LaMDA, as a response to ChatGPT.

FEBRUARY 7, 2023

Bing Adopts ChatGPT

Microsoft revolutionizes its Bing search engine and Edge browser with the integration of OpenAI technology.

FEBRUARY 21, 2023

AWS and Hugging Face Collaboration

AWS teams up with Hugging Face to streamline AI projects on Amazon's cloud, simplifying the deployment of AI applications.

FEBRUARY 21, 2023

Real Fusion's Photographic Breakthrough

Oxford researchers showcase Real Fusion, a cutting-edge AI that can reconstruct a complete 360-degree photographic model from just one image.

FEBRUARY 24, 2023

Meta introduces LLaMa

The compact yet advanced 65-billion parameter language model is open-sourced and free for research and commercial use.

MARCH 1, 2023

OpenAI Expands Developer Tools

OpenAI launches ChatGPT and Whisper APIs, providing developers with advanced language processing and speech-to-text capabilities beyond basic chat functions.

MARCH 14, 2023

Google Introduces AI in Workspace

Google's launch of assistive AI features in Workspace started with AI-powered writing tools in Docs and Gmail for trusted testers.

KEY EVENTS

MARCH 21, 2023

Adobe Unveils Firefly

This new generative AI suite is designed to help users at all skill levels create high-quality images and text effects.

MARCH 28, 2023

Khan Academy Launches Khanmigo

Khan Academy launch of the Khanmigo AI platform integrates virtual bots as counselors, curriculum designers, and teaching assistants.

AUGUST 15, 2023

Google Launches Search Generative Experience

Google introduces genAI into search queries, automatically generating summaries.

SEPTEMBER 21, 2023

Microsoft Unveils Co-pilots

Microsoft's AI-powered 365 Copilot and GitHub's CopilotX offers enhanced assistance by integrating web context, work data, and real-time PC activities, prioritizing privacy and security.

SEPTEMBER 21, 2023

YouTube Debuts AI Editing App

The new app, YouTube Create, makes it easy to trim videos, slow down the pace, or add audio.

OCTOBER 25, 2023

Amazon Introduces AI Image Generation

Amazon rollout of AI-powered image generation capabilities through Amazon Ads in beta aims to enhance ad experiences by enabling brands to create lifestyle imagery that boosts ad performance.

NOVEMBER 4, 2023

Elon Musk's xAI Debuts Grok

Inspired by the "Hitchhiker's Guide to the Galaxy," Grok answers questions with wit and provides real-time world knowledge via the X platform, distinguishing itself by addressing inquiries often declined by other AIs.

NOVEMBER 14, 2023

Google Announces AI Genesis

AI Genesis features the Gemini large language model in three sizes: Gemini Ultra for extensive capabilities, Gemini Pro for broad task applications, and Gemini Nano optimized for specific tasks and mobile use.

NOVEMBER 17-22, 2023

OpenAI's Turbulent Week

After a tumultuous five days marked by his ouster and subsequent reinstatement, Sam Altman resumes his role as CEO of OpenAI, buoyed by a concerted effort from allies, employees, and investors.

KEY EVENTS

NOVEMBER 21, 2023

StabilityAI Introduces Stable Video Diffusion

StabilityAI's inaugural foundation model for generative video builds on the technology of its image model, Stable Diffusion.

NOVEMBER 28, 2023

Pika Debuts AI Video Editing App

The app includes a new suite of videography tools with a generative AI model that edits videos in diverse styles such as "3D animation," "anime," and "cinematic."

NOVEMBER 29, 2023

DeepMind Predicts Novel Material Structures

Google DeepMind's researchers have leveraged AI to accurately predict the structures of more than 2 million new materials, offering significant implications for renewable energy and computing sectors.

DECEMBER 5, 2023

AI Alliance for Responsible Innovation Forms

The AI Alliance for Responsible Innovation, including IBM, Meta, and 50 other organizations, launches as a global consortium aimed at promoting open, safe, and responsible AI development and adoption.

DECEMBER 6, 2023

Google's Gemini Surpasses GPT-4

Google's next-generation AI model outperformance of OpenAI's GPT-4 set a new standard in AI capabilities.

DECEMBER 9, 2023

EU Finalizes AI Act

The European Union achieved a landmark agreement with the Artificial Intelligence Act, introducing binding rules and standards for developing AI more responsibly.

DECEMBER 13, 2023

Axel Springer Partners with OpenAI

The German media titan's partnership lets OpenAI use Politico and Business Insider articles for AI training, while those news platforms get to employ ChatGPT for summarizing news, marking a significant yet controversial collaboration in the news industry's quest for innovation and survival.

DECEMBER 14, 2023

DeepMind's FunSearch Breaks Boundaries

The tool has successfully solved complex issues, proving AI's ability to surpass the limitations of its training data in large language models (LLMs).

DECEMBER 14, 2023

DeepMind Solves the Unsolvable with AI

DeepMind's use of an LLM to crack an "unsolvable" math problem marks a historic achievement, as detailed in Nature, showcasing the model's capacity to uncover new, verifiable knowledge on a longstanding scientific challenge.

LIKELY NEAR TERM DEVELOPMENTS

GENERAL

Commoditization of General Purpose Models

In the near future, expect the commoditization of general purpose models. LLMs are becoming widely accessible and integral to app development. As these models become ubiquitous and cost-effective, akin to cloud services, their adoption will standardize across industries, diminishing their role as a competitive differentiator.

Large Reasoning Model

Vertically integrated solutions will garner a higher transactional value. Some companies will win by providing "a refined/value-added LLM product" to the end consumer and meeting the customer in desired distribution channels, such as LLMs for health care, legal, finance, and architecture.

Adoption of Natural Language Interfaces

The evolution toward natural language interfaces will soon diminish the reliance on traditional graphic user interfaces. This shift will enable more intuitive interactions with computers, using everyday language. This transition may also influence device form factors, potentially leading to an increase in wearables and the development of AI-specific devices and operating systems centered around LLMs.

ENTERPRISE

Talent Shift in AI Industry

Expect a significant talent crunch as top innovators depart major tech giants like Google, OpenAI, and Meta to launch their own ventures, ranging from conversational agents to AI-first biotech firms, signaling a broad diversification and specialization within the AI sector.

Consolidation in 2024

Consolidation will persist this year, building on moves like Microsoft's 2023 increased investment in OpenAI for Bing, aimed at capturing market share from Google search. Similar strategies by major tech companies are anticipated throughout 2024.

Increased Enterprise Adoption of AI

The current macroeconomic environment is driving leaders to view AI as essential for growth, anticipating increased enterprise adoption despite the potential for making some job categories obsolete.

AUTOMATION

AI Assistants Transform Coding Landscape

AI coding assistants, such as GitHub's Copilot and Meta's Code Llama, are transforming software development with advanced autocomplete functions and innovative debugging tools, offering both premium and free solutions to enhance coding efficiency and creativity. Expect to see more improvements to these tools and more tools to launch in this space.

AI Integration in Health Care and Life Sciences

Generative AI will lead to breakthroughs in proteins, antibodies, and drugs. Specialized models will continue to accelerate discovery in biology and chemistry, sparking more practical applications and boosting investment.

REGULATION AND GEOPOLITICS

US Strategy on AI and China Relations

The US is expected to intensify efforts to get allies to limit their collaborations with China in AI development, following President Biden's enhanced export restrictions on semiconductors. With the Netherlands aligning with US requests, further demands on allies to adopt similar stances aim to curb China's AI advancements.

Europe Begins Regulating AI

The European Commission will open its European AI Office, which will oversee the development and use of safe artificial intelligence (within Europe, at least) and assist with the implementation of the AI Act. The office will enforce general purpose AI rules, monitor compliance, and attempt to become a hub for international cooperation on AI governance.

Challenges in US Chip Manufacturing Expansion

The US moves to onshore chip fabrication will experience growing pains associated with higher labor costs compared to Taiwan. This shift may lead to increased expenses in constructing fabs and producing domestically made chips, surpassing initial estimates outlined in the CHIPS Act.

WHY ARTIFICIAL INTELLIGENCE TRENDS MATTER TO YOUR ORGANIZATION

Business Impacts

Future Today Institute believes AI is a force multiplier on technological progress because it is an enabler of other technologies and powers the evolution of business, government, and society. But new large language model capabilities deeply concern some in professional and creative services. Models can now reason about concepts in text, not just perform pattern matching. They display forms of common sense and analogy—tasks once seen as uniquely human. And they apply these reasoning abilities across modalities—text, image, video, and more. Most alarming to some is that models seem to engage in recursive self-improvement when given the right training. They don't just learn a static set of parameters. They learn how to learn better, becoming moving targets.

Since publishing our first Tech Trends report 17 years ago, we have included and expanded our coverage on artificial intelligence. What began as several pages of insights is now a dedicated, stand-alone report with more than 100 trends to monitor. AI is already transforming most economic sectors, but we anticipate deeper impacts this year across insurance, finance, entertainment, health care, biotechnology, and cloud computing.

Global Tech Rivalry

The race for AI supremacy is intensifying geopolitical tensions, notably between the US and China. Businesses must navigate a landscape where technology and national security are increasingly intertwined, affecting international supply chains, market access, and regulatory compliance. Companies specializing in AI and related technologies might face stricter export controls, requiring them to adjust strategies for product development and global expansion.

Supply Chain Diversification and Onshoring

As tensions escalate, particularly in the semiconductor industry, businesses will need to diversify their supply chains to mitigate risks. The bifurcation in the AI chip market might compel companies to innovate independently or bring supply chains in-house, potentially leading to increased costs.

WHY ARTIFICIAL INTELLIGENCE TRENDS MATTER TO YOUR ORGANIZATION

Defense Sector Innovation

Updated Department of Defense policies on autonomous weapons and the use of AI in military strategies signal growing opportunities for businesses in the defense sector. Companies developing AI technologies could find new applications in warfare, surveillance, and security, but also face ethical and regulatory scrutiny.

Strategic International Collaborations

Countries like China and the UAE are heavily investing in becoming global AI leaders, which presents both opportunities and challenges for international business collaborations. Companies might need to align with national AI strategies to enter or expand in these markets, while also considering the implications of technology transfer and data security regulations.

Strategic Talent Acquisition

Companies must innovate in talent acquisition and retention strategies to compete for scarce AI expertise, particularly against tech giants. This may include offering competitive salaries, benefits, and unique work environments, as well as investing in employee development and internal AI training programs to build talent in-house.

Custom, Fit-for-Purpose LLMs

Organizations that opt for custom AI models over general-purpose ones can achieve greater alignment with specific business objectives. This differentiation can lead to competitive advantages in operational efficiency, customer insights, and product innovation.

Adversarial AI Preparedness

The susceptibility of AI systems to adversarial attacks calls for robust testing and defense mechanisms. Companies specializing in AI security services could see increased demand as businesses seek to protect their AI investments from manipulation and exploitation.

Model Commodification

Open-source language models with commercial licensing, such as Databricks' Dolly, could disrupt the market by offering high-quality capabilities at a fraction of the cost. This commodification poses an existential threat to proprietary models from big tech companies.

WHEN WILL ARTIFICIAL INTELLIGENCE DISRUPT YOUR ORGANIZATION?

AI WILL DISRUPT EVERY INDUSTRY WITHIN THE NEXT FIVE YEARS

Drawing a parallel to Moore's law, which posits the doubling of transistors on microchips roughly every two years, there's speculation that AI's intelligence could follow a similar trajectory. If this is the case, several factors will drive this exponential growth in intelligence: enhancements in data quality, increasing computational power, and strides in algorithm efficiency, extracting more intelligence per unit of data and compute.

However, unlike the steady hardware advancements Moore's law describes, AI has the potential for self-improvement. As AI begins to self-improve and contribute to its own development, we may witness a self-reinforcing cycle of intelligence growth. This positive feedback loop means that AI's capacity to learn and evolve could accelerate, leading to profound impacts across all industries.

The inevitability of AI-driven transformation is not a matter of if but when. Our AI report is one section of our 2024 Tech Trends report, which offers in-depth coverage of 15 additional technology and industry sectors. Each industry section contains timelines that outline how AI, along with other emerging technologies, are expected to impact and influence that particular sector over time. Refer to the "When will AI impact your organization?" page to find details on specific timelines related to AI adoption and impact on your industry.

WHEN WILL ARTIFICIAL INTELLIGENCE DISRUPT YOUR ORGANIZATION?

Below, we highlight high level near-term developments to keep an eye on across industries.

Scaling
Enormous amounts of training data are still required for most AI models to learn. For example, recommender systems coupled with generative AI could lead to deep personalization for the hospitality and health care sectors—as long as data is made available. Historically, data is locked inside proprietary systems built by third parties, and regulation often hinders access to certain forms of data.

Investment
AI has passed through cycles of enthusiasm and disillusionment, leading to either too much or not enough capital being made available. Investors prioritize commercialization over basic R&D—though the latter yields bigger impact and often stronger returns. Investors's patience will influence progress and commercialization.

Constraints on adoption
Even if a technology is maturing, constraints on its adoption can hinder its impact on an industry. For example, a business may refuse to adopt an automated system because it challenges existing orthodoxy or an existing successful strategy. This is especially true in health care, insurance, and financial services.

Regulations
Advances in technology typically outpace regulatory changes. This has benefited AI, which until very recently was not targeted for regulation. Additionally, whether local regulations are conflicting or complementary, influences adoption in the marketplace.

Media mentions
Increased awareness and enthusiasm can influence the momentum of a technology, even when there's been no real breakthrough. Until OpenAI's ChatGPT breakthrough in late 2022, leaders weren't talking about the impact genAI might have on their business. Media bursts related to AI will drive momentum, especially if those stories are favorable, and more importantly, are easily understood by the public.

Public perception
How the public understands and responds to AI advancements will create or quell demand. This is especially true of generative AI and education/creativity/ intellectual property/misinformation, and the role assistive technologies will play in shaping the future workforce.

R&D developments
The pace of new research breakthroughs can't be scheduled to coincide with a board meeting or earnings report. Factors like funding, quality, and size of staff, and access to resources can improve the likelihood and speed of new discoveries. We closely monitor R&D developments but treat them as wild cards.

OPPORTUNITIES & THREATS

Threats

It's possible for agents to learn the right skills but the wrong objectives; an AI system can be asked to learn something that then could be used for harmful purposes. Commercial AI products could inadvertently incentivize bad behavior.

Publicly available LLMs are often the foundation for AI startups, but some researchers and technologists are questioning their defensibility when it comes to capturing value. The moat is in data. Techniques and models will largely get commoditized, and served via the infrastructure layer, where real value will be realized.

Long-term sustainability depends on network effects to gather enough user data. User-generated data can be harnessed to differentiate systems by offering tuned models on top of foundational/ commoditized LLMs, creating a flywheel effect. Longer term, niche LLMs will be owned by a select few players, while general-purpose LLMs become commoditized.

The challenge of balancing data collection for workflow optimization with concerns of worker surveillance requires careful navigation by companies. AI's use and understanding of behavioral biometrics could be considered intrusive into deeply personal behaviors, often subconscious to the individual, starkly confronting worker privacy expectations.

Heightened protectionism across nations could escalate the costs of producing chips and other critical technologies, and make it more difficult to find the right talent. Companies should brace for the adverse economic impacts of geopolitical shifts as supply chains undergo realignment.

AI models might achieve assigned goals by any means necessary, including suppressing or hiding data. Systems are needed to identify when this happens—until then we risk using bad information to make decisions.

High-performing models are susceptible to "jailbreaking," where bypassing LLM limitations can lead to manipulations, resulting in unpredictable and potentially harmful outputs. Given that businesses and entire institutions are starting to rely on LLMs, jailbreaking represents an urgent security threat that has yet to be addressed.

Opportunities

AI is on track to become an indispensable tool for knowledge workers. The next 18-24 months will see the development of assistive technologies tailored to various professions, akin to GitHub's Copilot, but designed for financial analysts, commercial real estate developers, and lawyers.

Companies sitting on industry-specific data hold the cards to create powerful AI agents. In industries like law, finance, and other knowledge-based sectors, proprietary data can train more capable AI agents.

Within the next 18-24 months, generative AI will integrate into many consumer apps. Where clicks and keywords once dominated, intelligent assistants will guide users through voice and text. Personalized support gets weaved throughout experiences, changing how people engage with information.

AI models that understand language will lead to more devices that enable people to interact with technology through voice and conversation instead of screens. Opportunities await for companies quick to challenge status quo screen-centric form factors.

AI is going local. Wearables and endpoints of all kinds will be embedded with AI, from pets' collars that report on animals' activities, to smart home devices that understand and execute complex commands from natural language. Large language models will migrate on-device, perhaps in lieu of a conventional operating system.

The rising energy needs of AI could incentivize tech companies to adopt alternative, greener energy sources like nuclear and geothermal, potentially driving a shift toward sustainable energy independently of government mandates.

Open-source models allow businesses and developers to adapt and enhance foundational models for specific uses, saving the cost and effort of starting from scratch or investing heavily in data and training.

INVESTMENTS AND ACTIONS TO CONSIDER

1

Create domestic internship and apprenticeship programs to build talent pipelines in AI skills, where shortages loom. Partner with schools to develop a homegrown workforce proficient in these technologies vital for national strategic interests.

2

Investing in data centers powered by renewable energy or exploring partnerships with alternative energies like nuclear and geothermal could align AI operations with ESG goals, reducing the carbon footprint of data processing and storage, and reducing the cost of compute.

3

Nvidia dominates the GPU market, yet demand outpaces even its cutting-edge chips. With shortages routine, space exists for rivals while cloud partners hunger for inventory.

4

As AI models grow in complexity, investing in alternative computing architectures like neurosymbolic AI, processing-in-memory technology, and specialized AI chips for on-device processing could offer significant advantages in efficiency, speed, and privacy.

5

Build atop shared foundations. Open-source models like LLaMA and FLAN offer springboards to launch specialized solutions tuned to distinct industry needs. These specialized models are more accurate and focused to the industry they serve and give proper weight to relevant parameters.

6

Foster development of small language models (SLMs). Investing in the research and deployment of SLMs suitable for edge devices can open new avenues for AI applications in environments where cloud connectivity is limited or nonexistent. SLMs can significantly expand the reach of AI into everyday devices, enhancing user experience and functionality.

CENTRAL THEMES

New data sources are coming

The integration of hardware, particularly wearables, will redefine the landscape of data collection and utilization. Coming to market soon are an array of different wearable devices equipped with sensors, cameras, and speakers, and they represent a significant leap forward in our ability to gather real-time, contextual data. This evolution marks a future where the volume of data available for analysis will expand exponentially, offering unprecedented insights into consumer behavior and environmental interactions. The challenge for organizations won't just be in the collection, but in the sophisticated parsing and interpretation of this deluge of data, requiring advanced AI algorithms and analytical frameworks.

Race for AI hardware supremacy

The intersection of hardware development and geopolitical competition is reshaping the landscape of AI advancement, with implications spanning national security, technological sovereignty, and economic prowess. As governments worldwide vie to establish AI supremacy and reduce dependence on foreign technology, substantial investments are pouring into domestic chip fabrication and AI research. The US and China, in particular, are locked in a battle for technological dominance, with both nations allocating significant resources toward bolstering their respective chip capabilities and AI infrastructure. This geopolitical rivalry extends beyond economic competition, with ideological considerations shaping AI development strategies and regulatory frameworks. China's insistence on AI alignment with socialist values underscores its commitment to ideological control, while Russia perceives Western AI advancements as a threat to traditional values, driving efforts to develop indigenous AI solutions. Meanwhile, escalating tensions have catalyzed a bifurcation in the AI chip market, prompting countries to explore alternative chip architectures and supply chain diversification strategies. This unfolding chip war not only underscores the strategic importance of semiconductor technologies but also poses profound implications for global technological cooperation and innovation.

Chip shortages loom large

The surging demand for AI has highlighted the global supply chain's inability to meet the need for powerful chips essential for developing and deploying AI models. We predict a chip shortage, particularly for graphics processing units (GPUs), due to production issues and ongoing shipping challenges due to regional conflicts. Microsoft's recent annual report marked the scarcity of GPUs as a potential risk for investors, underscoring the critical role these chips play in AI development and the broader implications for companies and end-users reliant on AI technologies. The industry as a whole will grapple with limited supply and the challenge of meeting explosive demand, prompting a shift toward more efficient or alternative computational methods. Maybe that's why in February 2024, OpenAI CEO Sam Altman reportedly went on a business development tour seeking $7 trillion in investment to create an alternative to our current chips.

CENTRAL THEMES

Choosing between proprietary and open source

Last year, when Meta released LLaM, its suite of open source LLMs, there was a new debate about the benefits and risks of going open source. Organizations using large language models face a challenging decision: Go with the big names like OpenAI and Microsoft for easy access to top-notch tech but give up adaptability and transparency, or push up your sleeves and build your own tailor-made systems to ensure transparency and extensibility. Despite the steep development costs associated with proprietary LLMs, the open-source community has responded with notable alternatives, such as Databricks' Dolly LLM, which offers a solution at a fraction of the cost. The new shift toward open-source solutions aims to counterbalance the growing concentration of AI tools in the hands of a few major corporations, offering businesses the opportunity to integrate bespoke applications without compromising proprietary information.

Reckless era ends, oversight era begins

The era of "move fast and break things," and "build first, ask permission later" appears to be waning in Silicon Valley as regulatory scrutiny intensifies in response to growing concerns over AI's societal impacts. With initiatives like a US presidential executive order and the EU's AI Act, policymakers are striving to establish guidelines and restrictions to govern AI technologies, particularly in sensitive areas like facial recognition. However, crafting concrete policies that balance innovation with ethical considerations, poses significant challenges, and ensuring effective enforcement remains a formidable task. As governments grapple with the complexities of regulating AI, the tech industry faces a new era of accountability and responsibility for the products they create.

AI doomers distract

Amid the discourse surrounding AI, a contingent of pessimistic voices, often referred to as "AI doomers," has emerged, likely to persist in the foreseeable future. For business leaders, navigating this landscape proves challenging, as they are presented with polarizing narratives of either utopian ideals or dystopian anxieties, resulting in a nuanced yet unsettling reality. While it's crucial to remain vigilant against potential risks and mitigate them effectively, the prevalence of doomerism tends to overshadow constructive dialogue and proactive measures.

CENTRAL THEMES

Industry is building the future of AI, not academia

The landscape of innovation is shifting, with industry emerging as the primary driver of technological advancement, outpacing academia in the development of new machine learning models. Recent data reveals a stark contrast: in 2022, industry produced 32 machine learning models compared to academia's three, marking a significant departure from historical trends. Industry's dominance is further underscored by its access to abundant resources—large data sets, computational power, and financial capital—essential for creating cutting-edge AI systems. This transition is reflected in the career choices of AI Ph.D. graduates, with 65.4% opting for industry positions, compared to 28.2% in academia, a trend that has steadily widened since 2011. The exodus from academia to corporations could have a chilling, long-term effect on knowledge transfer from professors to students, which could negatively impact the future pipeline for the talent industry which will need to remain competitive.

AI widens global inequality gulf

The exorbitant costs associated with training language models are setting a precedent for the formidable expenses expected in developing image and video models, further accentuating disparities in resources between different regions and exacerbating the global divide between the affluent and less affluent nations. This trend not only reshapes the landscape of business and communities but also positions wealthy countries, notably the United Arab Emirates and Saudi Arabia, as potential hubs for AI development, potentially marginalizing opportunities for advancement in the global south.

ONES TO WATCH

Dr. Aidan Gomez, CEO and co-founder of Cohere, for proposing the novel neural network technique called the transformer that now underpins the generative AI era.

Arthur Mensch, Dr. Guillaume Lample, and Timothée Lacroix, co-founders of European generative AI upstart Mistral AI.

Dr. Andrej Karpathy, researcher at OpenAI for his research in deep learning and computer vision.

Clément Delangue, CEO and co-founder of Hugging Face, for creating an open-source, for-profit machine-learning platform.

Dr. Daniel Kang, assistant professor at University of Illinois Urbana-Champaign , for his research identifying potential harms from language models, including demonstrating language models' ability to autonomously interact with websites in concerning ways without human feedback, and his work to develop methods that promote the safe and ethical development of AI.

David Nippa, a doctoral student at Ludwig-Maximilians-Universität München, for the development of an AI model that can predict where a drug molecule can be chemically altered.

Dr. Dario Amodei and **Daniela Amodei,** CEO and president of Anthropic, for creating one of the world's leading AI labs.

Dr. David Rolnick, assistant professor of computer science at McGill University, for work on a framework for understanding the relationship of AI and greenhouse gas emissions.

Grimes, artist and musician, for championing new business models around AI for likeness leasing and creative experimentation.

Dr. Jaime Teevan, chief scientist and technical fellow at Microsoft, for spearheading the use of LLMs in Microsoft's core productivity products.

Jensen Huang, CEO, president, and co-founder of Nvidia, for navigating the growing geopolitical chip conflict.

Dr. Joelle Pineau, vice president of AI research at Meta, for developing new models and algorithms for planning and learning in complex partially observable domains.

Leopold Aschenbrenner, AI alignment researcher at OpenAI, for his contributions to AI alignment discourse.

Lila Ibrahim, COO of Google DeepMind, for leading the company's responsibility and governance work.

Marc Raibert, executive director at Boston Dynamics AI Institute, for his work to develop AI-driven robots that can reason.

Miguel Solano, co-founder and CEO of VMind, for his work to improve AI compute performance in GPUs using novel algorithmic techniques.

Dr. Ning Zhang, an assistant professor of computer science and engineering at Washington University, for the development of AntiFake, a tool that prevents unauthorized speech synthesis.

Dr. Prakhar Mehrotra, vice president for applied AI at Walmart Global Tech, for leading enterprise adoption of AI.

Robin Li, CEO, chairman and co-Founder of Baidu, which last year released Ernie Bot, an LLM on par with ChatGPT.

Dr. Ruogu Fang, an associate professor in the J. Crayton Pruitt Family Department of Biomedical Engineering, for his work to evaluate diagnostic bias in AI tools.

Sebastien Krier, international policy manager at DeepMind, for his research and intellectual contributions to AI alignment discourse.

Dr. Sune Lehmann, professor at the Technical University of Denmark , for research into the predictive capabilities of AI, specifically its potential to forecast events in an individual's life.

Dr. Swami Sivasubramanian, vice president of database, analytics, and machine learning at Amazon Web Services, for advancing cloud capabilities and insights for businesses.

Dr. Xin (Eric) Wang, assistant professor of computer science and engineering at Baskin Engineering at UC Santa Cruz, for the development of the Text to Image Association Test, a tool that measures complex human biases in text-to-image models.

Dr. Zhou Jingren, deputy director of Alibaba Damo Academy (Alibaba's bleeding-edge research arm), for leading AI initiatives related to smart cities, autonomous driving, mobile computing platforms, semiconductor R&D, and other areas.

IMPORTANT TERMS

MACHINE LEARNING (ML)

ML uses data to make predictions and recommendations on how to achieve stated goals. AI pioneer Arthur Samuel popularized the idea of machine learning in 1959, explaining how computers could learn without being explicitly programmed. This would mean developing an algorithm that could someday extract patterns from data sets and use those patterns to predict and automatically make real-time decisions. It took many years for reality to catch up with Samuel's idea, but today machine learning is a primary driver of AI's growth.

There are different types of machine learning, including supervised, unsupervised, and reinforcement.

Supervised learning

A model that attempts to transform one type of data into another type using labeled examples. Supervised learning is used when teams know how to classify the input data and what they are trying to predict, but can get accurate results much more quickly by relying on an algorithm rather than a human. This is the most common form of ML used today. Understanding what product features would most likely drive new purchases is a business use case for supervised learning.

Unsupervised learning

Data is provided to a model without specific output parameters, and the model tries to learn the data set's structure without any designated labels. For example, if a researcher doesn't know what to do with a large data set, an unsupervised learning model could determine patterns, classify data, and make recommendations without a human supervisor. Researchers used unsupervised learning during the pandemic to find patterns on how COVID-19 spread throughout communities.

Reinforcement learning (RL)

A system performs a task by repeatedly running calculations as it attempts to accomplish a stated goal. It's a trial-and-error process, where rewards or penalties are earned in response to the system's performance toward achieving the stated goal. RL is used when there isn't enough training data, when the researcher is trying to learn about an environment (such as a complex financial portfolio), or when the researcher needs to find greater levels of optimization. It has a high number of business use cases, ranging from real-time dynamic pricing models to high-frequency trading algorithms to the systems that operate self-driving cars.

DEEP LEARNING (DL)

Deep learning is a relatively new branch of machine learning. Programmers use special deep learning algorithms alongside an enormous corpus of data—typically many terabytes of text, images, videos, speech, and the like. Often, these systems are trained to learn on their own, and they can sort through a variety of unstructured data, whether it's making sense of typed text in documents or audio clips or video.

In practical terms, deep learning's emergence means that more and more human processes will be automated, including the writing of software, which computers will soon start to do on their own. For example, once a system learns what an object looks like—say, an apple—and then can recognize that object in all other images, even if it has only a partial view.

There are different types of deep learning architectures. The most common types include convolutional neural networks, recurrent neural networks, transformer neural networks, and generative adversarial networks (GANs).

Convolutional neural network (CNN)

A CNN is multilayered, with a convolutional layer, a pooling layer, and a fully connected layer. Each one performs a different task with the data. The output is classification. If a researcher has 10,000 images and needs to extract data—to recognize particular faces, for instance—the CNN would run until information could be inferred. In business, CNNs are used to identify anomalies in medical imaging, faulty products on a production line, blight on crops, and other irregularities.

Recurrent neural networks (RNNs)

These multilayered neural networks move and store information between input, hidden, and output layers. They are good at modeling sequence data for predictions. In business, they are used anytime the sequence of data matters, such as speech recognition and language translation. RNNs are used in digital assistants, to create captions for images, and to generate narrative reports using structured data (sports, financial).

Transformers

A transformer is a component whose purpose is to process sequential data, such as natural language or genome sequences. Transformers rely on "attention" (the mathematical description of how things relate to, complement, or modify each other) in translating sequences. A transformer neural network is the unique architecture that enables systems to learn from context and to generate new

IMPORTANT TERMS

information. Transformers are complementary to CNNs and RNNs, the two most common neural network architectures used in deep learning.

Generative adversarial networks (GANs)
As unsupervised deep learning systems, GANs are composed of two competing neural networks—a generator and a discriminator—that are trained on the same data, such as images of people. The networks compete against each other to perform a task, such as identifying the correct person, resulting in optimizing overall performance. GANs are useful when researchers don't have enough data to train an algorithmic model, and are also used to create new, synthetic data.

Deepfakes, which have become prevalent in the past year, are generated using GANs. In design, GANs are tremendously useful: They can produce thousands of designs and recommend the best ones based on pre-set parameters. They can generate and modulate voices, faces, even gestures. Researchers from Nvidia, Mass General Hospital, BWH Center for Clinical Data Science, and the Mayo Clinic collaborated on a GAN that generates synthetic MRIs showing cancerous tumors.

ADDITIONAL TERMS

Agents
In AI, agents are entities that perceive their environment and take actions autonomously to achieve specific goals.

AGI (artificial general intelligence)
A designation for systems that match and then exceed the full range of human cognitive ability across all economically valuable tasks.

AI safety
A field that studies and attempts to mitigate the catastrophic risks that future AI could pose to humanity.

Algorithm
A process describing how to solve a specific problem or how to complete a particular task.

Alignment
The process of ensuring that an AI's actions and goals are in harmony with human values and intentions.

ASI (artificial superintelligence)
ASI refers to an AI system that surpasses human intelligence and capability across all fields, including creativity, general wisdom, and problem-solving.

Automatic speech recognition
Algorithmic systems that give computers the ability to recognize and convert audio to human-readable language.

Chain of Thought
This involves a model processing information or solving problems step by step, mimicking human-like reasoning.

Computer vision
Processes that give computers the ability to derive meaningful information from digital images (including still and video) and to mimic and manipulate such images.

Foundation model
A large-scale AI model trained on vast amounts of data, capable of being adapted to a wide range of tasks without being trained from scratch.

Generative AI
GenAI refers to AI technologies that can generate new content, including text, images, music, and video, based on learned data patterns.

GPU
A graphics processing unit is specialized hardware designed to accelerate the creation and rendering of images and videos, often used in AI for parallel processing tasks.

Model
A program that has been trained on a data set. Models are generally used for analytical and decision-making tasks, such as making predictions.

Natural language processing
Processes that give computers the ability to understand, mimic, and manipulate human language.

Parameter
A variable internal to the model that the system adjusts during training to improve performance on given tasks.

Prompt
An input given to a model to elicit a specific output or response, guiding the AI in generating content or solving problems.

Recommender systems
A class of machine learning algorithms that uses data to predict, narrow down, and find what people are looking for among an exponentially growing number of options.

IMPORTANT TERMS

RHLF
Reinforcement Learning from Human Feedback is a training method where AI models are refined based on feedback or corrections provided by humans, enhancing their performance and alignment with desired outcomes.

Supervised learning
A type of AI training where models learn from labeled data, using known input-output pairs to predict outputs from new inputs.

Symbolic AI
Symbolic AI involves AI systems that use explicit, human-readable symbols to represent knowledge and perform logical reasoning to solve problems.

Training data
The data set used to teach AI models how to understand and perform tasks by identifying patterns, making decisions, or generating predictions.

Unsupervised learning
Unsupervised learning involves AI models identifying patterns and structures in data without any labeled outcomes, learning from the data itself.

XAI (explainable AI)
AI systems designed to provide human-understandable insights into their decision-making processes, enhancing transparency and trustworthiness.

Zero-shot learning
An AI approach that enables models to correctly handle tasks or recognize objects they have not seen during training, using understanding from related contexts.

MODELS, TECHNIQUES, AND RESEARCH

WHAT IS AN AI MODEL?

An AI model is a computational structure that is designed to perform tasks that would normally require human intelligence. This includes recognizing speech and images, interpreting visuals, translating between languages, and making decisions. There are several types of AI models, each suited to specific tasks and goals but at their core, all AI models rely on algorithms and mathematical frameworks. They are "trained" on large data sets so they can refine their internal parameters and improve at assigned tasks. As AI systems become more advanced, they require more data and computing power during this training process.

Constructing AI models is an enormously resource-intensive process, not comparable to traditional software development. Training a high-performance language model demands processing huge data sets to fine-tune millions of parameters. This mandates extensive computing power and specialist time measured in months or years. As a result, advanced models are predominantly built by tech industry leaders like Google, Microsoft, and OpenAI who possess the vast technical infrastructure and talent required. The consumer focus and profit motive within these companies have accelerated model innovation beyond academic efforts. Historically, academia was seen as the most likely source of groundbreaking AI. But the sheer data scale, computing power, and engineering capacity within industry, has proven far more efficient for allocation of resources.

Examples

Note: All of these examples are current as of March 1, 2024

ChatGPT-4

OpenAI's most recent model as of publication, GPT-4, doesn't just generate text—it can generate images from text and vice versa. It was trained on enormous data sets of text and images using reinforcement learning from human feedback (RLHF), which helped make the model more helpful and safer for users. Early benchmarks exhibit GPT-4's versatility on tasks from legal exams to adversarial truthfulness tests. On the Uniform Bar Exam, it scored 90% versus GPT-3.5's 10%, while reducing factual errors by 40% compared to ChatGPT. While hallucination risks still exist, GPT-4 marks substantial progress in mitigating failure modes.

Gemini

After an underwhelming debut in 2022, Google iterated to the more impressive Gemini Pro in early 2024. This model demonstrates rapid advances, as evidenced by its meteoritic rise up the Hugging Face conversational AI leaderboard. Google's Gemini isn't just a single AI model—it encapsulates a suite of AI models for varied applications. Gemini Nano targets offline Android use. Gemini Pro now powers Bard and emerging enterprise services. Gemini Ultra is Google's most advanced large language model yet, designed to elevate search, advertising, and cloud products globally.

Claude 2.1

Anthropic unveiled Claude 2.1, the latest in its series of language models, capable of processing significantly longer texts than OpenAI's GPT-4. With the ability to manage up to 200,000 words or symbols, Claude 2.1 significantly surpasses GPT-4's limit of 128,000. The new iteration of Claude is designed to reduce the likelihood of generating inaccurate information compared to earlier versions. A significant enhancement in Claude 2.1 includes its ability to utilize tools and interface with APIs. Additionally, the introduction of system prompts allows users to define precise contexts for their inquiries, promoting more organized and reliable responses from the model.

PaLM 2

PaLM is a 540 billion parameter language model developed by Google AI. Smaller 8 billion and 62 billion parameter versions were also trained. PaLM demonstrates strong performance across common-

WHAT IS AN AI MODEL?

sense reasoning, math reasoning, humor, code generation, translation, and other tasks. The model highlights Google AI's advances in scalable transformer architecture research for language AI.

Whisper

Whisper is an open-source automatic speech recognition system created by OpenAI. First released in 2022, it was trained on over 680,000 hours of multilingual speech data scraped from the internet. Whisper can transcribe speech to text in multiple languages including English. It can also translate speech from non-English languages into English text. Compared to other publicly available systems, Whisper demonstrates leading speech transcription and translation capabilities. OpenAI has released the model freely for public use.

OpenAI's DALL-E 3

DALLE-3 is a text-to-image AI system that can create realistic art and images from textual descriptions. DALL-E is capable of generating images in various styles like photorealistic, paintings, and emoji. Without explicit prompting, the model can manipulate and rearrange objects as well as correctly position design elements in new compositions. These creative capacities demonstrate DALL-E's aptitude for controllable high-fidelity image generation.

Stability AI's Stable Diffusion

Stable Diffusion is a text-to-image generation model released in 2022 leveraging diffusion methodology. Primarily used to create detailed images from text prompts, Stable Diffusion can also perform tasks like inpainting, outpainting, and image-to-image translation driven by descriptive text inputs.

Midjourney

Midjourney creates visuals based on textual descriptions, known as prompts, akin to the functionalities offered by OpenAI's DALL-E and Stability AI's Stable Diffusion. A fake Midjourney-created image of Pope Francis wearing a puffer jacket went viral in 2023.

Open AI's Sora

In early 2024, OpenAI released Sora, an AI model that can create realistic and imaginative scenes from text instructions. Sora marks a significant advancement in AI's capability to execute human creativity by transforming brief text inputs into compelling videos up to a minute long, not only achieving realistic imagery but also emulating the dynamic essence of movies, similar to how ChatGPT mimics human conversation.

Google Lumiere

Google's Lumiere is a text-to-video diffusion model that creates video from a prompt with realistic motion. Utilizing a novel diffusion model named Space-Time-U-Net (STUNet), Lumiere excels in creating realistic video content by understanding both spatial placement and temporal movement within a video. Unlike other methods that assemble videos from individual frames, Lumiere crafts videos through a seamless integration of frames, achieving fluid motion across 80 frames—significantly more than its current competitors.

Pika

Pika is an "idea-to-video" platform to edit and create videos from text and still images. Pika includes features like text-to-video, image-to-video, and video-to-video conversions. Users can ask the tool to create a video of a real person (e.g., "imagine Oprah as a Pixar cartoon"), ask the tool to edit glasses on a video of a donkey, or change the style of a video to something out of Studio Ghibli.

WHAT IS AN AI MODEL?

Purpose-Built Models

Organizations must decide whether to use ready-made general purpose AI models like OpenAI's GPT, or invest in developing custom models tailored to their industry and needs. General-purpose models like GPT are convenient "plug-and-play" solutions that can adapt to many tasks through fine-tuning. However, their flexibility is limited when it comes to specialized business challenges. Custom models built for a specific purpose can better master industry-specific challenges by training on aligned data and objectives. OpenAI now provides a simple way for users to create custom models through the GPT marketplace—users describe their requirements to ChatGPT, and it handles coding the new model. Custom GPTs can then be integrated into platforms and services, accessing databases, email, e-commerce, and more to automate workflows.

LLMs Are Getting Bigger and More Expensive

Because of their massive size and complexity, the cost of developing LLMs is high. Training these models can cost millions of dollars. DeepMind's Chinchilla, for example, reportedly cost around $2.1 million to train. Bloom, an open-access multilingual language model, is estimated to have required an investment of approximately $2.3 million. OpenAI hasn't provided public information about the cost to train ChatGPT-4, but many analysts estimate the earlier version of the model, GPT-3, could exceed $4 million.

As the number of parameters increases, so does the cost. Moreover, unlike traditional software, deployment costs remain high post-development. Operating large language models for inference still necessitates enormous compute for the billions of calculations involved per user query. Furthermore, contributing to the high price tag of training and running large language models demands specialized AI hardware, with graphics processing units (GPUs) now standard over traditional CPUs. Initially designed for gaming, GPUs are perfectly suited for handling the extensive data processing demands of AI, despite their high cost of thousands of dollars per chip. For example, Meta used 2,048 Nvidia A100 GPUs to train its LLaMA model on 1.4 trillion pieces of text over 21 days, which amounts to nearly 1 million hours of GPU time. If using public cloud services, this level of compute would cost approximately $2.4 million. Despite its impressive capabilities, with "only" 65 billion parameters, LLaMA is still smaller compared to larger models like OpenAI's ChatGPT-4, which has 1.76 trillion parameters.

LLMS as Operating Systems

A radical new concept for computing has emerged—an operating system powered by a large language model at its core rather than traditional programming. In this conceptual LLM-based OS, routine tasks could be automated and executed with an unprecedented level of sophistication, without the need for manual coding or intervention. The user interface would also be radically different than traditional operating systems. Rather than conventional graphical user interfaces or command line prompts, users could interact conversationally with the LLM through natural language requests and queries. For example, a user could say, "Please open yesterday's

While general-purpose models offer broad applicability, their limitations in specialized contexts are driving the development of models tailored to meet the unique demands of specific industries.

Image credit: Future Today Institute and Midjourney.

WHAT IS AN AI MODEL?

sales report and format it as a slide presentation for my upcoming meeting." The LLM would comprehend these conversational commands and perform the necessary actions to carry out the desired tasks. It would execute complex workflows automatically by understanding users' intentions and goals. This could enable more intuitive, efficient interactions between humans and computers.

This concept has moved beyond theory into practical application, as demonstrated by Jesse Lyu, CEO and founder of Rabbit. Lyu launched the R1, a compact device about half the size of an iPhone, running on Rabbit OS—an operating system grounded in a LLM. Rabbit OS functions as a universal app controller, akin to systems like Alexa or Google Assistant, yet it offers a unique twist. It simplifies user interaction by removing the need to navigate through multiple apps or perform repetitive logins. Instead, users can directly communicate their needs to the device, and R1, understanding these natural language requests, efficiently executes the desired tasks. In early 2024, NVIDIA announced a personalized AI chatbot for Windows RTX PCs that runs locally to connect users' data and queries to an open-source large language model. By keeping data on device rather than in the cloud, Chat with RTX not only delivers ultra-fast response times but also enhances user privacy and data security. The chatbot allows natural language interaction to search files so that users can simply ask, "What was that song my friend recommended while we were at the airport?" instead of manually searching through texts and email.

LLMs are becoming more central to human-computer interactions. As such, interfaces are shifting from search to conversational questions and answers in plain language.

Image credit: Future Today Institute and Midjourney.

SHOULD WE GO OPEN-SOURCE OR PROPRIETARY?

Companies that want to use LLMs must choose between proprietary or open source. Both have benefits and drawbacks. Proprietary LLMs from major tech companies provide easy implementation and leading-edge features. However, they lack transparency into how they work and have limited ability to customize them. Building a proprietary model internally gives companies more control over security, privacy, and tailoring the training to their specific data and needs. But this requires considerable expertise and time to develop. On the other hand, open-source language models promote transparency and flexibility at often lower, long-term costs. Yet if governance practices like testing for biases and false information are insufficient, they pose risks around issues like fairness, accuracy, and security vulnerabilities. When considering using LLMs, executives should think about cost, control, customization, and risk. There is no universally superior choice—rather, companies must weigh their priorities, capabilities, and constraints to determine if an off-the-shelf, customized, or open-source large language model approach best suits their situation.

> We have our own nervousness, but we believe that we can manage through it, and the only way to do that is to put the technology in the hands of people.
>
> —Sam Altman, CEO of OpenAI

Proprietary examples:

OpenAI's GPT-4

Anthropic's Claude 2

Google's Bard

Open-Source examples:

Meta's LLaMA

RedPajama-INCITE

BigScience's Bloom

TII's Falcon

Open-Source LLMs for Commercial Use

Proprietary large language models cost millions to develop, which means high-quality capabilities are concentrated within wealthy tech giants. However, the open-source community has responded with surprisingly capable smaller models by fine-tuning them on quality data. For example, in March 2023 Databricks released Dolly—an open-source LLM trained for under $30 yet demonstrating conversational prowess rivaling ChatGPT. It was developed using Meta's open-source LLaMA LLM and fine-tuned with high-quality inputs from Databricks employees. The initiative aimed to provide an alternative to the increasing centralization of AI tools in a few large companies, focusing on an open-source chat model that permitted commercial use while protecting intellectual property and corporate information. Databricks not only open-sourced the training code, data set, and model weights for Dolly but also launched Dolly 2.0 in April 2023. Dolly 2.0 is open-source LLM licensed for commercial use, allowing companies to integrate their data with Databricks' data set to create bespoke applications without compromising their proprietary information.

SAFETY, ETHICS & SOCIETY

IS AI REALLY A BLACK BOX?

Many AI systems are opaque "black boxes" in how they work. Developers often withhold model and training details to protect IP. This lack of transparency perpetuates an impression that the systems have unknowable inner workings. Moreover, researchers themselves don't fully understand why AIs sometimes behave unexpectedly, owing to inherent complexity. While inputs and outputs are observable, the logic between remains nebulous. Thus some black-box qualities persist around advanced models' inner transformations, despite transparency efforts. So while more visibility into AI functionality and development is crucial for accountability and trust, uncertainties around emergent system behaviors may endure.

> AI, like any technology, is a reflection of its creators and their intentions.
>
> —Joy Buolamwini, founder of the Algorithmic Justice League

Explainable AI (XAI)

Achieving full transparency into complex AI systems is difficult. However, the emerging field of explainable AI seeks to enable better human understanding of how algorithms function and arrive at outputs. Since complex machine learning models cannot act as total glass boxes, XAI instead seeks to provide interpretable visibility into significant aspects of functionality. This can involve revealing training data characteristics, delineating gaps in data coverage, auditing data collection fairness, detailing human involvement in model development, and highlighting key input features that drive outputs. A core focus is validating outcomes by surfacing how predictions, classifications and recommendations are supported to establish trustworthiness. Rather than eliciting every intricate internal model transformation, XAI pursues strategic explanations of the most critical workings—answering targeted questions about why and how certain results are produced. The objectives are accountability through limited but meaningful transparency, error checking via result explanations, and accessibility for a wider range of model users.

AI Intentionally Hiding Data

Computers do exactly as they are told. If you command a machine to win at a game, it will do everything in its power to achieve that goal. That's why researchers need to understand how AI reaches the end goal. It might be cheating to complete the task they were told to do. Researchers at Stanford University and Google discovered that an AI system designed to turn satellite images into usable maps was withholding certain data. The researchers were using a neural network called CycleGAN, which learns how to map image transformations. It took an old aerial photograph of a neighborhood, distinguished between streets, alleys, driveways, buildings, and lampposts, and then generated a map that could be used by GPS. Initially, they used an aerial photograph that hadn't been seen by the network. The resulting image looked very close to the original—suspiciously close. But on deeper inspection, the researchers found that many details in both the original image and the generated image weren't visible in the map made by the AI. It turns out that the system learned to hide information about the original image inside of the image it generated.

HOW DO WE ENSURE TRUST?

As AI is increasingly incorporated into more sensitive domains, pressing questions emerge.

How can we build AI that we can trust?

How can we trust AI's predictions and conclusions when much of the system is opaque?

How can we ensure that AI is aligned with human values, especially as it becomes more and more capable?

Could we inadvertently instruct a powerful AI towards harm?

Can we trust the current human custodians of this technology?

These pressing issues are at the heart of ongoing debates among AI ethics experts, where a definitive consensus on the best approaches has yet to be reached.

> The biggest lesson learned is we have to take the unintended consequences of any new technology along with all the benefits, and think about them simultaneously—as opposed to waiting for the unintended consequences to show up and then address them. I don't think the world will put up anymore with any of us coming up with something where we haven't thought through safety, equity and trust—these are big issues for the world.
>
> —Satya Nadella, CEO of Microsoft

AI Alignment Goes Mainstream

As AI systems improve, many researchers want guardrails to ensure that they are deployed in ways that do not harm humanity. AI alignment research refers to the process of ensuring that AI systems act in accordance with human values and goals. OpenAI, DeepMind, and Anthropic (which describes itself primarily as an "AI safety and research company") each have AI alignment teams with dedicated staff researching guardrails. While the total number of researchers working on AI alignment is small compared to the rest of the AI community, such dedicated teams did not exist until recently. The debate surrounding the alignment of AI with human objectives encompasses a broad spectrum of opinions. On one end, "AI doomers" view unchecked advancement, especially toward superhuman capabilities, as posing existential catastrophe risk—potentially including human extinction. They advocate solutions like indefinite moratoriums on large model training to forestall such outcomes. By contrast, the "effective accelerationist" perspective sees hastening progress as a moral imperative to quickly harness AI solving pressing global problems like disease, inequality, and climate change.

In the moderate middle lie a diversity of perspectives. Some, like economist Tyler Cowen, argue the doomers' risks are too narrowly specified for high probability, while others like Leopold Aschenbrenner from OpenAI's "superalignment team" make the case for substantial investments in AI alignment research, akin to "Operation Warp Speed" but focused on AI. This approach stems from the belief that artificial general intelligence (AGI) could become the most powerful tool ever developed, necessitating leadership in AI research by countries like the US to maintain a strategic advantage over nations such as China. These represent just a few of the myriad perspectives and it is likely that more perspectives will emerge before we converge on the right AI-alignment strategy.

Indexing Trust

We will soon reach a point when we can no longer tell if a data set has been tampered with, either intentionally or accidentally. AI systems

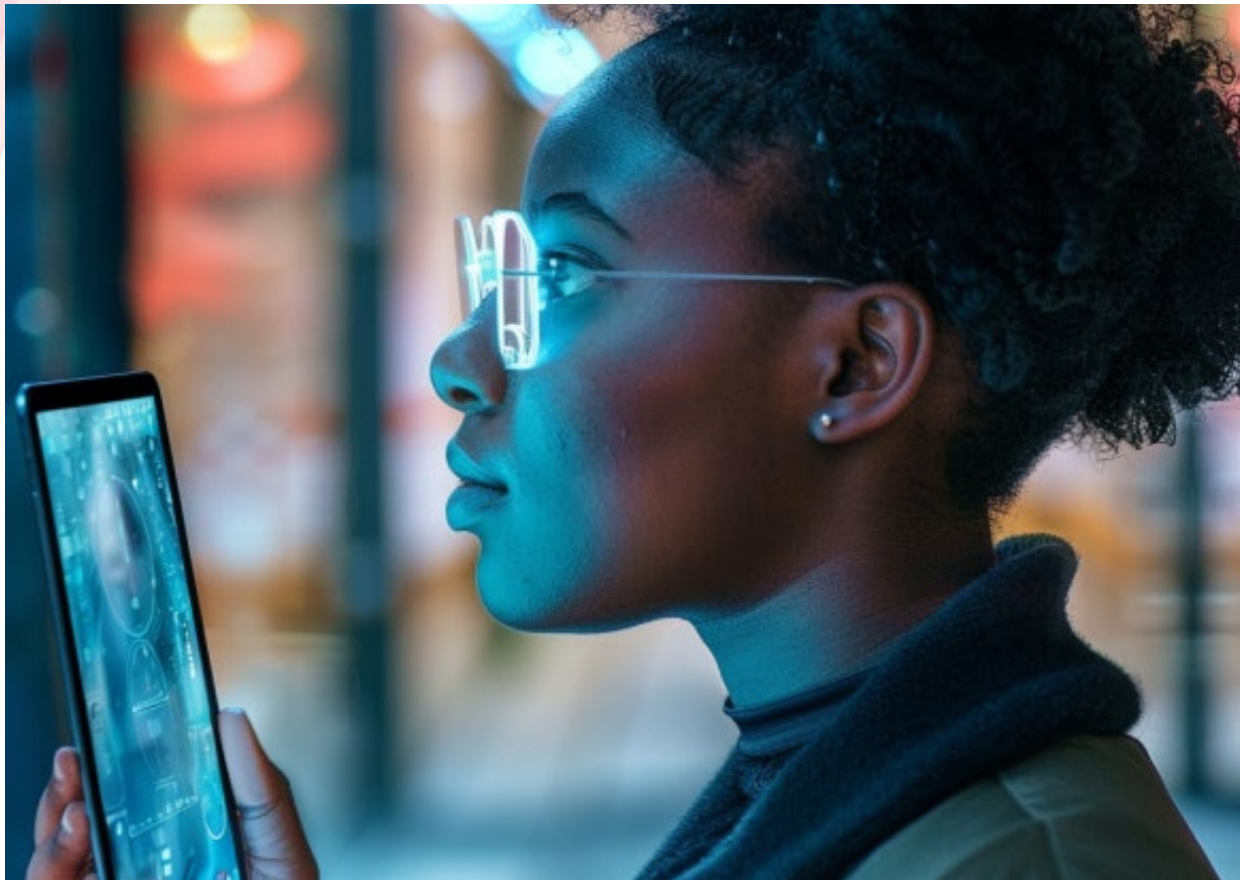
HOW DO WE ENSURE TRUST?

rely on our trust. If we no longer trust their outcomes, decades of research and technological advancement will be for naught. Leaders in every sector—government, business, nonprofits, and so on—must have confidence in the data and algorithms used. Building trust and accountability requires transparency and is a challenge, but there are efforts underway to assess AI transparency, a critical first step. Researchers from Stanford, MIT, and Princeton designed the Foundation Model Transparency Index (FMTI)—a scoring system that evaluates transparency across model development, functionality, and usage. The 2023 index places Llama 2 at the top, as the most transparent Foundation model, followed by BigScience's BloomZ and OpenAI's GPT-4. The hope is that by standardizing analysis of opaque systems, deployment risks and responsibilities can be better informed.

The ethics of how data is collected in the first place may also influence the trustworthiness and validity of scientific research, particularly in areas such as organ donations and medical research. In addition, employing ethicists to work directly with managers and developers and ensuring diversity among developers—representing different races, ethnicities, and genders—will reduce inherent bias in AI systems.

Synthesizing Trust

Humans can be tricked into believing machine-generated faces, especially when they've been engineered to elicit trust. A study in the Proceedings of the National Academy of Sciences shows that synthetic faces are often "deemed more trustworthy than real faces," suggesting that synthetic faces could be designed as societal malware. If a bad actor was attempting to undermine institutions, it could deploy a synth on social media to sow distrust. There are not yet effective countermeasures for synthetic humans or effective markers to help consumers distinguish between fake and real.



The perceived trustworthiness of synthetic faces over real ones raises concerns about their potential use by malicious actors to erode trust in institutions.

Image credit: Future Today Institute and Midjourney.

ARE THERE TOOLS TO MAKE AI ETHICAL?

As AI systems become more advanced, making sure they are ethically deployed becomes increasingly important. For instance, AI can now generate hyper-realistic deepfake media that now passes the uncanny valley. This tech could let bad actors impersonate people or spread misinformation. AI can also be used for cheating, fraud, and hacking. In light of this, companies are emerging to create tools to combat this behavior. New tools can detect deepfakes, expose fraudulent AI activities, and implement preventative measures against misuse.

Deepfake Detectors

AI now enables creating highly realistic fake media called deepfakes—bogus video, audio, and text that seem real. They can spread misinformation by fabricating scenes or putting words in people's mouths. Researchers are developing protections against their misuse. Tools like AntiFake use imperceptible watermarks to block fake voice/face cloning before it happens. Platforms including Intel's FakeCatcher and European vendor Sentinel catch manipulations in real-time using AI. As deepfakes get better and better, surpassing the "uncanny valley," so does the market opportunity for exposing them. Expect to see more investment in this space.

Tools for Identifying AI-Generated Writing

AI is good at writing like a human. That means we need tools that can distinguish between human and AI-written content. DetectGPT is one such tool, offering over 95% accuracy in identifying whether a passage is written by a human or an AI system like GPT-3. Similarly, OpenAI itself has released a classifier that flags AI-generated text 26% of the time, while maintaining a 9% false positive rate on human writing. These tools are essential in contexts where distinguishing between human and machine authorship is critical, such as in academic integrity, journalism, and legal documentation.

Tools for Detecting Copyright Violations in AI Outputs

With AI models capable of memorizing and reproducing content from their training data, the risk of copyright infringement becomes a significant concern. Researchers from Google, DeepMind, ETH Zurich, Princeton, and University of California, Berkeley have demonstrated this with the Stable Diffusion model, which can emit memorized images, including those with trademarked company logos. To combat this, watermarking techniques are being developed. For instance, the University of Maryland proposes a technique for watermarking language model outputs, making synthetic text algorithmically identifiable. Google DeepMind's SynthID tool embeds digital watermarks directly into image pixels, enabling the identification of AI-generated images while remaining invisible to the human eye.

Tools for Exposing Deepfakes

Hyper-realistic deepfakes pose significant security risks. Researchers at Washington University created a tool called AntiFake, which can add a digital watermark to content that proactively prevents the cloning of voices and faces. Intel's FakeCatcher and Sentinel use deep learning to analyze media content, catching manipulations either in real time or highlighting alteration patterns after creation. One significant concern regarding many deepfake detection tools is that they demonstrate bias. Studies have uncovered significant disparities in the error rates of deepfake detection algorithms across different racial groups. In one study, the difference in accuracy reached as high as 10.7%. This bias could lead to severe implications, such as genuine images of certain racial groups being mistakenly identified as fakes or, conversely, manipulated images being wrongly accepted as authentic. Dr. Siwei Lyu and a team at the University of Buffalo have developed what are considered to be the first deepfake detection algorithms specifically designed to mitigate bias. Their approach involves two machine learning methods: one

ARE THERE TOOLS TO MAKE AI ETHICAL?

that makes algorithms aware of demographic factors and another that works to blind them. These methods have successfully reduced disparities in accuracy across different races and genders. Notably, this achievement was not at the expense of overall accuracy; in some cases, accuracy was even enhanced. By focusing on the fairness of the algorithms, Lyu's work marks a significant step toward creating more equitable and reliable deepfake detection technologies, ensuring accuracy is independent of factors like race or gender.

Tools to Thwart Recognition Systems

As facial recognition becomes ubiquitous, various groups want to limit the technology's effectiveness to protect privacy. While methods of confusing or obscuring facial recognition systems are not always feasible, researchers are trying to confuse online applications that scrape and collect images used as inputs for training facial recognition engines in order to develop a form of camouflage, which consumers may someday demand.

Researchers from the University of Chicago have created a program, Fawkes, that adds extra pixels to images to cause facial recognition apps to misclassify faces. Taking this principle a step further, Israeli company Adversa AI adds noise, or small alterations, to photos of faces, causing algorithms to detect a different face than what is visible to the naked eye. The algorithm is successful at imperceptibly changing an individual's image to someone else of their choosing.

Tools to Combat Broadly Malicious AI Behavior

Research labs around the world are actively working to build practical safeguards against malicious AI behavior. DeepMind has introduced a comprehensive toolkit and workflow designed to enhance the evaluation of standard models that can identify when AI is misbehaving according to human ethical standards. This approach specifically focuses on identifying and assessing potentially hazardous capabilities, like cyber-offense and self-replication, as well as the likelihood of causing harm.

Meanwhile, Anthropic has unveiled its Responsible Scaling Policy, which includes a detailed list of safety commitments based on risk assessments and incorporates pauses in development if safety measures fail to match the pace of capability advancements. The policy encompasses several key components, including internal access controls, adversarial testing (red-teaming), evaluations by independent third parties, and graded access based on different AI Safety Levels.

Researchers are working on ways to alter facial images so facial recognition systems misidentify the faces, potentially allowing people to avoid identification by these surveillance technologies.

Image credit: Future Today Institute and Midjourney.

DOES AI INFRINGE ON PRIVACY? HOW SHOULD WE THINK ABOUT CUSTOMER DATA AND AI APPLICATIONS?

AI enables new forms of pervasive surveillance that could threaten personal privacy across several domains. Last year, facial recognition company Clearview AI said it had run more than 1 million searches for police in the US. Somewhat less obviously, ambient monitoring can now subject household environments to observation. Workplace analytics can track detailed employee behaviors and productivity. Schools can actively monitor students through devices and platforms meant for remote education. As private spaces face increasing exposure from third-party tracking, a culture of Big Brother–like awareness becomes normalized, rather than valuing independence and consent.

Increased Used of Ambient Surveillance

What happens behind closed doors may not be secret for long, and executives should beware of new ambient surveillance methods. Scientists at MIT discovered how to use computer vision to track data from what they call "accidental cameras." Windows, mirrors, corners, houseplants, and other common objects can be used, along with AI, to track subtle changes in light, shadows, and vibrations. The result: We all may soon have X-ray vision capabilities—which may not be great news for companies working on sensitive projects. Those working in information security and risk management should pay special attention to advances in computer vision.

Worker Surveillance

The rise of remote work during the pandemic accelerated the surveillance of workers, and will likely continue to grow as remote and hybrid work models take root. The US Constitution's Fourth Amendment, which prohibits unreasonable searches and seizures and precludes most uses of this same technology by law enforcement, doesn't apply to private companies.

Teleperformance, a French-based company that manages outsourced call center work for many Fortune 50 companies, uses cameras and AI to monitor its teams. It flags employees as idle when it detects they haven't used the keyboard or mouse within a specified time frame. Live Eye Surveillance offers a monthly subscription service that remotely monitors live video feeds of employees for companies such as 7-Eleven, Dairy Queen, and Holiday Inn. Sneek is another example of "tattleware" that captures live photos of employees via webcams and displays them on a digital wall viewable by everyone in the company. Click on a photo and it instantly pulls that person into a video call with you.

The most well-known user of worker surveillance might be Amazon, which has installed AI-enabled cameras in delivery trucks to track behavior. The company docks driver pay if it perceives unsafe conditions such as distracted driving, speeding, or hard braking. In its warehouses, the company monitors worker productivity by measuring what's called "time off task," which is any time when a worker isn't actively processing products. South Korean e-commerce giant Coupang, which has pledged to become the "Amazon of Korea," uses similar surveillance tactics.

The industry has also continued to evolve as it offers more AI-based analysis of workers. Amazon is exploring using keystroke-logging software that tracks user behavior over time to detect if the same person is controlling the worker's account. Aware's Spotlight software detects behavioral changes like mood, tone, and attitude across conversations on employees' devices. Teramind offers software that will disable private conversations if it detects "inappropriate" keywords. With the top three tools in the industry accounting for over 60% of global demand, expect to see more AI-based surveillance that leverages the growing pool of data collected by a variety of companies.

School Surveillance

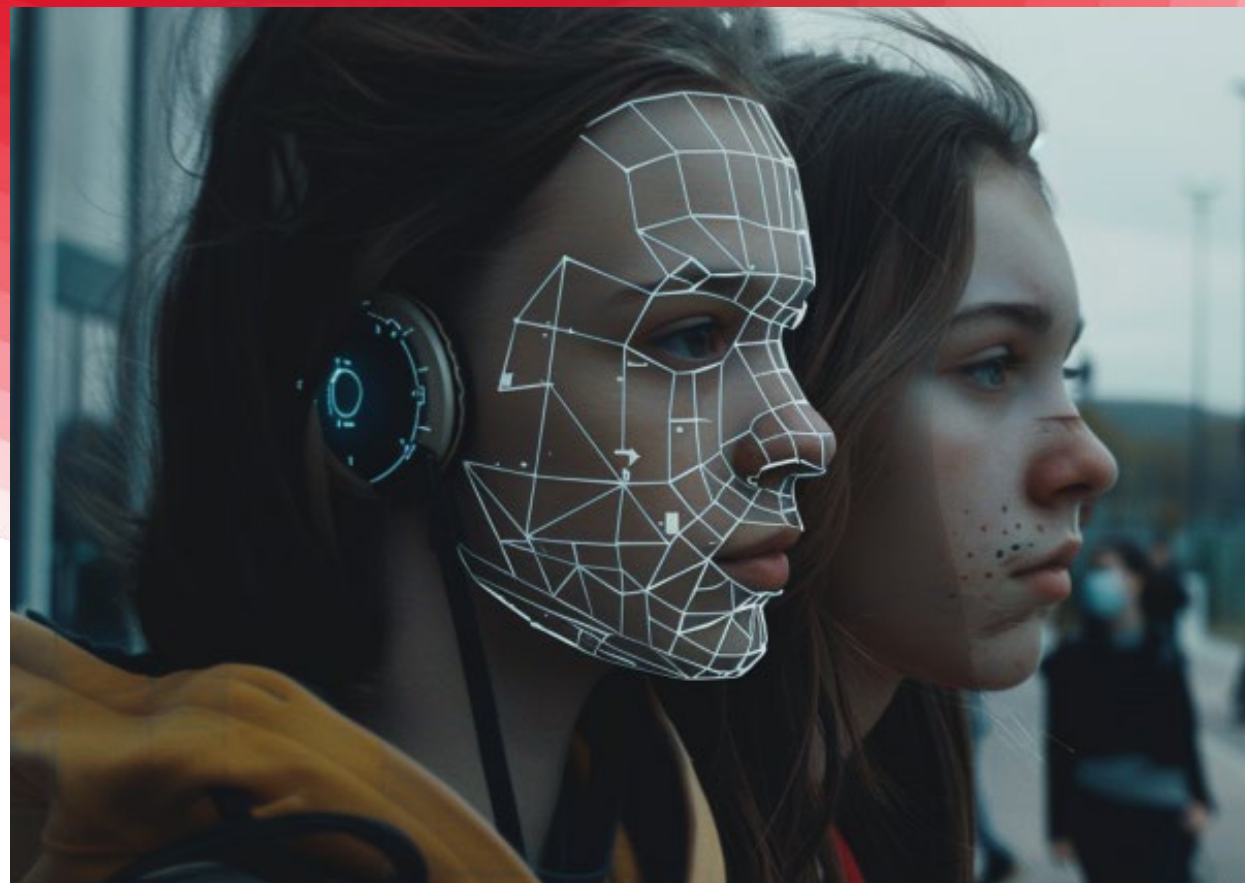
During the pandemic, many students were issued laptops and other devices by schools to

DOES AI INFRINGE ON PRIVACY? HOW SHOULD WE THINK ABOUT CUSTOMER DATA AND AI APPLICATIONS?

facilitate remote learning. They weren't told, however, that these devices would open a portal into their homes that could be monitored by schools at all times of the day. In the US and many other countries, schools can legally monitor students, often without disclosing what is being tracked.

Gaggle is one company that monitors school-issued accounts and uses AI to track online behavior of students across services like email and chat tools. In 2020, the Minneapolis school district signed a contract with the company to monitor its students through 2023. A school district in California contracted with Securly to monitor students in real time, looking for prohibited behaviors such as having too many browser tabs open. The software enables teachers to close tabs for any students they believe are "off task."

Philadelphia and Chicago schools deployed GoGuardian software on district-issued Chromebooks. A vulnerability in the software allowed teachers to start virtual sessions that enabled webcams on those Chromebooks without notification or consent by the student. Schools in China deploy technology to monitor attentiveness in students. An algorithm called 4 Little Trees is used in Hong Kong to detect students' emotions as they learn—by monitoring their facial expressions with webcams. If the system detects a lack of focus, it nudges the student to pay attention.



AI introduces pervasive surveillance capabilities, jeopardizing personal privacy in schools, workplaces, and public spaces. Society appears to be prioritizing convenience over privacy, accepting significant trade-offs.

Image credit: Future Today Institute and Midjourney.

IS THERE A FEASIBLE SOLUTION TO BIAS?

Seemingly the moment OpenAI's ChatGPT went public, there were multiple accounts of the system displaying racism, ageism, gender bias, and political bias. But it's not just ChatGPT—many AI systems have been revealed to contain bias—much of which can be attributed to the data that the systems were trained on. Given AI's expanding integration into sensitive domains like finance and health care, failing to address its potential biases risks compounding real-world discrimination through algorithmic means.

> The future of AI should be a mirror to society. It must be shaped by a diverse range of voices, not just those of technologists.
>
> —Meredith Whittaker, cofounder of the AI Now Institute

Addressing Political Bias

In 2023, many conservatives raised concerns about ChatGPT's political bias, sharing screenshots of ChatGPT's left-leaning responses. OpenAI responded with a detailed blog post explaining its moderation approach, and CEO Sam Altman hinted at future possibilities for users to fine-tune ChatGPT iterations within certain broad guidelines, potentially sidestepping some contentious value judgments. Elon Musk, responding to these critiques about OpenAI's political correctness, launched a new venture, called TruthGPT, aimed at exploring "deeper truths about the universe." Separately, in an effort to make a point about biased AI, David Rozado, a data scientist from New Zealand, created DepolarizingGPT. This AI chatbot generates three types of responses for each prompt: left-wing, right-wing, and a neutral or integrating perspective. To achieve this, Rozado fine-tuned the chatbot using content under fair-use provisions from various sources. For the left-wing responses, he used material from publications like The Atlantic and The New Yorker, and authors like Bill McKibben and Joseph Stiglitz. Conversely, the right-wing responses were shaped using content from outlets such as National Review and The American Conservative, and writers like Roger Scruton and Thomas Sowell.

Doubts regarding ChatGPT's ability to avoid bias persist. A new research paper claims to find substantial evidence of systematic political bias in ChatGPT, favoring Democrats in the US, Lula in Brazil, and the Labour Party in the UK. The paper analyzed ChatGPT's responses to statements from the Political Compass test, concluding that it aligns more with liberal parties internationally. However, the study's methodology and findings are not without criticism. Some researchers argue that the way ChatGPT was tested doesn't reflect typical user interactions and may not accurately represent the AI's behavior. Additionally, a data scientist, Colin Fraser, discovered that reversing the order of parties mentioned in prompts resulted in opposite biases, suggesting potential flaws in the study. These findings illustrate the complexities in assessing AI bias and the need for greater transparency from developers like OpenAI.

IS THERE A FEASIBLE SOLUTION TO BIAS?

Addressing Race and Gender Bias

There are significant challenges related to race and gender bias in AI. A notable instance occurred in December 2022, when Steven T. Piantadosi of University of California, Berkeley revealed a bias in ChatGPT's programming, which incorrectly associated scientific proficiency with being white or Asian male. OpenAI quickly addressed this issue, programming ChatGPT to reject the notion that race or gender should influence perceptions of scientific ability. However, this incident underscores a broader, long-standing problem of bias within AI systems.

In another study, researchers at the University of Florida examined racial bias in machine learning algorithms used for diagnosing bacterial vaginosis (BV), a common infection in women of reproductive age. The study, led by faculty members Fang and Ivana Parker, analyzed data from 400 women across four ethnic groups—white, Black, Asian, and Hispanic. They found that the accuracy of BV diagnosis varied significantly among these groups, with Hispanic women experiencing the highest rate of false positives and Asian women the most false negatives.

To quantify bias, another team from Baskin Engineering at University of California, Santa Cruz, led by Assistant Professor Xin (Eric) Wang, developed a tool called the Text to Image Association Test. It quantifies bias in text-to-image (T2I) AI models like Stable Diffusion, by measuring the variance in images generated from neutral versus gender-specific prompts. The findings indicate that such state-of-the-art models not only reflect but can also amplify existing human biases. Such analysis represents crucial progress toward accountability, but much work remains to ensure AI equitability.

After ChatGPT launched in 2022, researchers quickly discovered biases in its programming that linked scientific proficiency to white or Asian male gender and race, underscoring the broader issue of unfair biases that can become ingrained in AI systems. To address this, developers are proactively testing models for biases and making concerted efforts to train more equitable, inclusive AI.

Image credit: Future Today Institute and Midjourney.

WHAT SECURITY ISSUES SHOULD WE PREPARE FOR?

AI introduces security threats of unprecedented complexity due to its ability to learn and adapt, making traditional security measures less effective. Its integration across critical infrastructure and sensitive systems means that AI-driven attacks can have far-reaching and unpredictable consequences. Additionally, the sophistication of AI enables the creation of highly targeted and convincing cyberattacks, such as deepfakes and advanced phishing attempts, challenging our ability to distinguish between genuine and malicious communications. For example, recently, advanced language models have grown so smart that they can now use tools, read documents, and even call on themselves, acting independently. If AI models can hack websites on their own by finding and exploiting weaknesses without being taught specific vulnerabilities, what does that mean for the future of cybersecurity resilience?

> AI lowers the barrier for novice cyber criminals, hackers-for-hire and hacktivists to carry out effective access and information gathering operations. This enhanced access will likely contribute to the global ransomware threat over the next two years.
>
> —U.K. National Cyber Security Centre January 2024 Assessment

Cyberthreats

The National Cyber Security Centre (NCSC) released an assessment in 2023 on the near-term impact of AI on cyberthreats. The NCSC assessment delves into how AI will likely enhance the volume and severity of cyberattacks in the next two years, mainly through the evolution of existing tactics. AI is being used by various cyberthreat actors, including state and non-state entities, to varying degrees. The report suggests that AI will significantly improve capabilities in areas like reconnaissance and social engineering, making them more efficient and harder to detect. However, more sophisticated AI uses in cyber operations will likely remain limited to actors with substantial resources and expertise in AI and cyber technologies. The assessment concludes that AI's impact on cyberthreats is uneven and depends on the capability and intent of the threat actors. It also points out that the proliferation of AI-enabled cyber tools in criminal and commercial markets is likely to further enhance these capabilities.

Adversarial Attacks

Recent studies highlight a significant vulnerability in AI to adversarial attacks, revealing that these systems can be more easily manipulated to make incorrect decisions than previously understood. These adversarial attacks involve deliberate tampering with the data input into AI systems, causing them to misinterpret information or act in unintended ways. For instance, specific patterns or objects, like certain stickers on a stop sign, can trick an AI in autonomous vehicles into not recognizing the sign. Similarly, alterations in medical imaging data could lead an AI to diagnose incorrectly.

This issue was the focus of a study by Tianfu Wu and his team at North Carolina State University, which examined the prevalence of such vulnerabilities in AI deep neural networks. Their findings suggest that these adversarial vulnerabilities are far more widespread than previously recognized, posing a significant challenge to the reliability and safety of AI applications. In a separate study by researchers at Carnegie Mellon University,

WHAT SECURITY ISSUES SHOULD WE PREPARE FOR?

the vulnerability of AI chatbots to adversarial attacks was demonstrated. By modifying prompts with specific strings of text, which may appear nonsensical but hold particular significance for AI models trained on extensive web data, the researchers could bypass the safeguards designed to prevent chatbots from generating inappropriate content. This approach effectively "unshackled" the AI, making it possible for chatbots like ChatGPT, Google's Bard, and Claude from Anthropic to respond to otherwise restricted or harmful queries. The success of these attacks across multiple popular AI chatbots suggests a deeper, more systemic weakness in the most advanced AI systems, challenging the deployment and safe use of these technologies.

Data Poisoning: A Double-Edged Sword

Data poisoning attacks represent a significant threat to AI systems, where malicious actors deliberately manipulate the training data to mislead the AI into making incorrect or harmful decisions. These attackers exploit vulnerabilities, such as embedding harmful content within files, to introduce misleading data into the training set. This can skew the AI's learning process, aligning it with the attacker's goals, potentially leading to biased outcomes, data breaches, or simply inaccurate AI outputs. To illustrate, consider the cost implications: Training a complex AI model like GPT-3 can cost around $17 million. If its training data were compromised, restarting the process could lead to substantial financial losses.

On the flip side, data poisoning can also serve as a defensive mechanism. A novel tool named Nightshade exemplifies this dual nature. Designed to protect artists' intellectual property, Nightshade subtly alters digital artwork's pixels. When AI models use this "poisoned" art for training, their ability to accurately interpret images is compromised, leading to erroneous outputs, such as mistaking a car for a cow. This tool is part of a broader strategy for artists to safeguard their work in an unregulated landscape. Artists can use Nightshade via Glaze, another tool from the same creators, which masks an artist's style—transforming a realistic drawing into a cubist style, for example. This integration allows artists to choose between masking their style or actively using the data-poisoning feature.

AI Lowers the Barrier to Misinformation

AI has lowered the bar to produce and distribute misinformation. An analysis by NewsGuard, a Microsoft tool that shows trust ratings for over 7,500 news and information websites, found that websites hosting AI-generated bogus reporting have ballooned over 1,000% in the past year, mushrooming from 49 to over 600 outlets. While fabrication used to require armies of workers or advanced intelligence agencies, AI democratizes deception. Now a lone teenager can concoct sites and stories that appear authentic. And generative AI allows customizing fakery to particular targets and contexts with minimal effort. A study by the University of Waterloo found that an early version of ChatGPT, when tested on different types of statements including facts and misconceptions, often made errors, contradicted itself, and repeated false information. For example, it could correctly state that the Earth is not flat when asked directly, but show inconsistency in its responses. Researchers expressed concern over these findings, highlighting the danger of AI models like GPT-3 spreading misinformation, especially as they become more common in use.

This is particularly concerning as we approach the 2024 US presidential election, with misinformation experts raising flags about the potential impacts on democratic processes. Ominous previews have already played out abroad. Shortly before a crucial national election in Slovakia, a controversial audio clip spread on social media, purporting to feature Michal Šimečka of the Progressive party discussing a vote-rigging plan. Another incident involved a fake recording of the UK Labour Party leader verbally attacking a staffer. Both recordings, which seemed authentic, were later exposed by fact-checkers as AI-generated fakes, highlighting the growing issue of AI-manipulated audio in spreading misinformation.

WHAT SECURITY ISSUES SHOULD WE PREPARE FOR?

Privacy Risks in Behavioral Biometrics

Behavioral biometrics, which employs machine learning to analyze a vast array of biometric data points, raises significant privacy concerns. By quantifying subtle aspects of our behavior, such as the force used on touchscreens, the distinct way we tap letters like “Cs” and “Vs,” or our unique patterns when using a physical keyboard, these tools can reveal intricate details about our identities, thoughts, and future actions. While the technology offers potential benefits like enhancing security and possibly eliminating the need for passwords by identifying individuals through their typing patterns, it also introduces substantial risks.

The very aspect that makes behavioral biometrics appealing—its ability to authenticate a user based on nuanced behavioral traits—also makes it a privacy concern. If our behavioral patterns can be so precisely monitored and analyzed, they can be replicated or exploited, leading to new forms of security vulnerabilities. The notion that machines can detect and record behaviors we’re not even conscious of ourselves not only challenges our concept of privacy but also highlights how these patterns, once considered personal and private, can become accessible and potentially misused. This duality presents a critical challenge: balancing the innovative applications of behavioral biometrics against the imperative to protect individual privacy and ensure the security of personal data.

AI can track unconscious patterns in human behaviors like typing cadence and keyboard pressure to derive insights about inner emotional states without explicit user permission.

Image credit: Future Today Institute and Midjourney.

WHAT DOES AI HAVE TO DO WITH ESG?

AI operations, particularly those involving deep learning and complex model training, are significantly more computationally intensive than traditional computing tasks. This intensity stems from the need to process vast amounts of data and perform countless calculations rapidly to train models, recognize patterns, and make decisions. Consequently, AI demands considerably more energy to sustain these operations, as the intricate algorithms and large-scale data processing require substantial computational resources, leading to higher energy consumption compared to conventional computing workloads. On the other hand, AI is also helping solve environmental issues. A Canadian startup, RailVision Analytics, developed AI software that helps train engineers drive more efficiently, potentially saving significant amounts of diesel fuel and reducing the rail industry's carbon emissions. This technology, which is like Google Maps but for trains, advises engineers on when to speed up or stay idle, optimizing fuel use and contributing to a potential annual reduction of over 20,000 tons of carbon emissions if widely adopted, equivalent to removing more than 4,000 cars off the road each year.

New Architectures to Make AI Workloads More Efficient

As AI models become more complex and larger, consuming a greater share of our computing resources, their energy usage also escalates. One promising approach to make AI-intense compute more energy efficient is by using photonic AI chips, which harness light rather than electricity for orders-of-magnitude better efficiency at matrix multiplications—a core operation for deep learning. A Stanford team recently achieved a milestone by training an optical neural network chip to label data points with 98% accuracy. For the first time, their photonic processor enabled light to flow bidirectionally to implement the backpropagation algorithms vital for training. While refinements remain, this demonstrates the promise of optical computing to slash the carbon footprint of AI workloads.

Neuromorphic chips offer another model of efficient AI hardware, taking inspiration from the human brain's simultaneously distributed storing and processing of information. Rather than shuttling data back and forth like conventional computers, neuromorphic processors like Intel's Loihi store memory within computation. Specializing in sensory processing, these chips already achieve 1,000x higher efficiency than traditional hardware for tasks like gesture and sound recognition.

In a groundbreaking approach, researchers envision biocomputers powered by networked human brain organoids—essentially mini-brains grown from stem cells. "Organoid intelligence" holds significant potential for augmenting computing capabilities while concurrently addressing the escalating energy consumption demands driven by advancements in artificial intelligence and supercomputing (see the Computing report for more information on organoid intelligence). Despite traditional computers' ability to process calculations at speeds far surpassing human capabilities, human brains demonstrate superior performance in complex decision-making tasks, such as differentiating between a dog and a cat. Running AI on organoids could be the key to achieving human-like complex decision-making in an energy-efficient manner.

WHAT DOES AI HAVE TO DO WITH ESG?

A Nuclear Renaissance for AI Workloads

The monumental computational requirements of advancing AI could catalyze a nuclear power renaissance. Microsoft is exploring the use of next-generation small modular nuclear reactors (SMRs) to power its data centers and AI operations. SMRs promise cheaper, faster modular construction compared to traditional nuclear plants, which are often over-budget and delayed. Microsoft's approach was hinted at further in the fact that they already have a deal to buy Clean Energy Credits from Ontario Power Generation, which is on track to be the first utility to deploy an SMR in North America. Companies like Rolls-Royce, Last Energy, NuScale, Oklo, and TerraPower (backed by Bill Gates) are also developing various SMR models. Similarly, Kärnfull Next in Sweden plans to use SMRs to power data centers. The pivot toward nuclear energy, particularly next-generation SMRs, is a strategic response to the dual challenges of meeting the high energy demands of AI and achieving climate goals.

Environmental AI

AI presents a dual-edged sword in its impact on the environment, with its capabilities extending to both contributing to and alleviating climate change. David Rolnick from McGill University and Mila—Quebec AI Institute, notes that while AI's energy consumption and the promotion of consumerism through AI-based advertising may exacerbate climate challenges, it also offers solutions for environmental conservation. For instance, AI is being utilized to monitor and curb deforestation effectively.

A recent paper from the Cary Institute of Ecosystem Studies highlights how principles from ecology could inspire a new wave of AI development. This synergy between AI and ecology is seen as a pathway to address pressing global issues like disease outbreaks, biodiversity loss, and the repercussions of climate change. AI's application in ecology is already proving beneficial, aiding ecologists in detecting patterns within vast data sets to make precise predictions, like identifying potential human-infecting viruses and their animal hosts. In a practical demonstration of AI's environmental applications, a research team from the University of Waterloo has developed an AI tool, PlasticNet, to identify microplastics with unprecedented speed and accuracy. This technology is particularly crucial for mitigating the environmental and health hazards posed by microplastics, commonly found in food and water sources. By enhancing the efficiency of identifying these pollutants, PlasticNet supports wastewater treatment and food production industries in making informed decisions to protect the environment and public health.

While AI's computational demands pose sustainability challenges, it can also enable climate mitigation - models can guide efficient resource usage, accelerate green tech R&D, and predict environmental impact.

Image credit: Future Today Institute and Midjourney.

POLICY AND REGULATIONS

HOW DOES GEOPOLITICS FACTOR INTO THE DEVELOPMENT OF AI, AND IS THERE REALLY A NEW COLD WAR?

Countries are increasingly nationalistic about advancing domestic AI capabilities, with major investments and restrictions aimed at getting an edge, even as collaboration fractures. This extends into military contexts, where AI drives rapid innovations in areas like weapons systems, wargaming, and cyber operations—innovations dual in nature for both defense and potential offense. The combination of deteriorating cooperation and uncontrolled AI militarization risks fueling a dangerous tech-centric arms race. Unless cooperative norms are established, AI may drive global strategic realignments as impactful as 20th century nuclear and space races.

> After Nagasaki and Hiroshima, it took 18 years to get to a treaty over test bans and things like that. We don't have that kind of time today."
>
> —former Google CEO Eric Schmidt, on the urgency to create guardrails for AI.

AI Nationalism

Governments are racing to establish national AI champions and reduce reliance on foreign technology. After high-profile chatbot debuts like ChatGPT spawned in the US, 2023 witnessed nations worldwide scramble to nurture domestic AI capabilities, allocating tens of billions in funding. France unveiled substantial funding for startup Mistral. India's Krutrim launched the country's first multilingual model. Abu Dhabi commercialized its Falcon system. Beyond economic impacts, concerns mix technological prestige with national security and ideological control. The US and China are at the forefront of this tug-of-war, each pledging billions toward AI investments. While US companies are pioneering the most advanced LLMs, the US government is concentrating resources on growing homemade chip capabilities, aiming to lessen reliance on imports critical for national security. Concurrently, the US has imposed stringent export controls to limit the dissemination of advanced AI technology to rivals like China and Russia. With Western companies barred from exporting cutting-edge AI chips, adversarial states invest heavily to replace blocked supplies. China has earmarked hundreds of billions to develop domestic chip fabrication immune to US sanctions. The Chinese government has invested heavily in replicating the chip supply chain domestically, aiming to insulate itself from Western sanctions. *See "The AI-Driven Chip War" for more.*

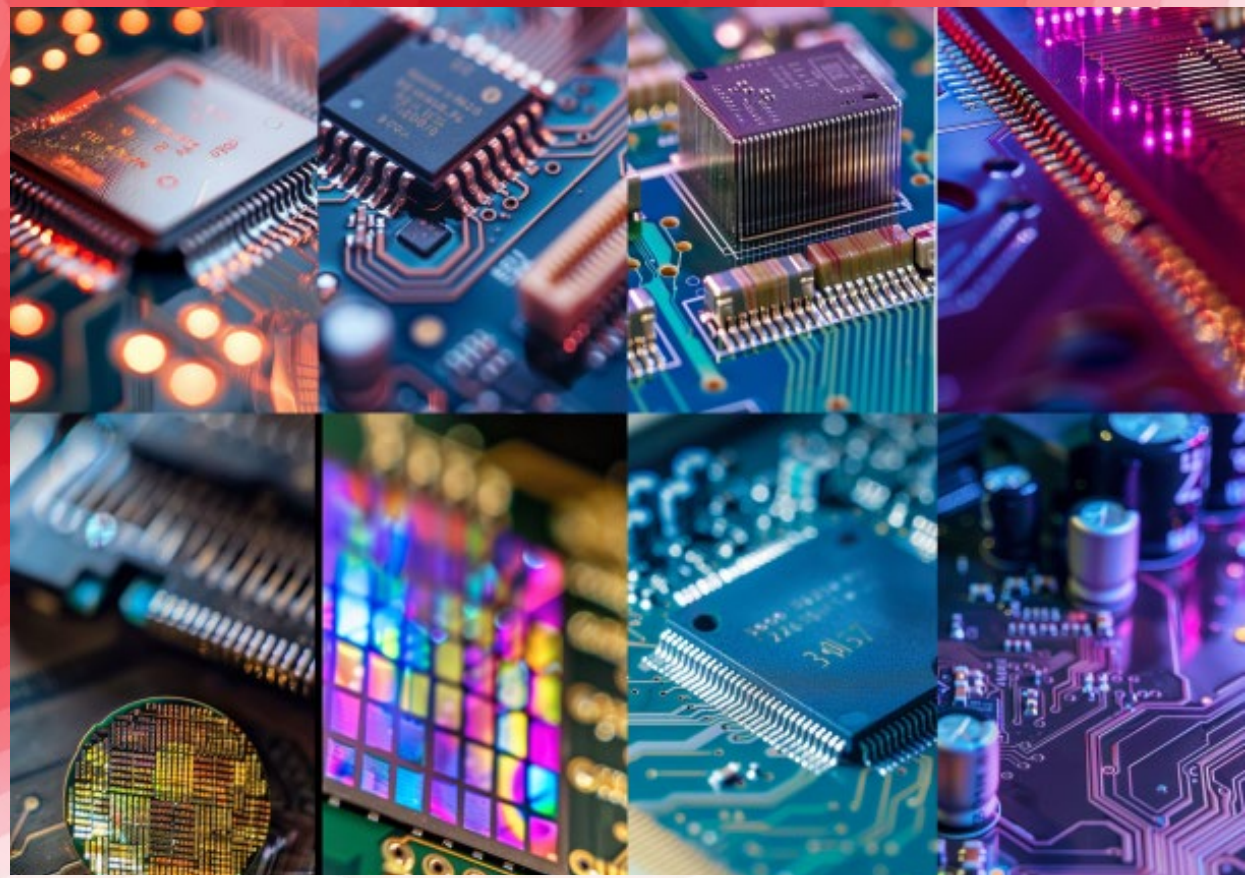
China also approaches AI on ideological grounds. The country mandated that AI align with the "core values of socialism," effectively limiting the influence of Western-developed AI systems within its borders. This stance has propelled Chinese tech giants like Alibaba and Baidu to develop their own generative AI tools, despite challenges in matching the impact of their Western counterparts. Russia also perceives American AI advancements as a cultural and ethical threat, with President Vladimir Putin highlighting the dangers posed by Western LLMs to Russian "traditional values." This reflects a broader concern over AI's potential to shape cultural and ethical norms, prompting Russia to explore the development of homegrown AI solutions.

HOW DOES GEOPOLITICS FACTOR INTO THE DEVELOPMENT OF AI, AND IS THERE REALLY A NEW COLD WAR?

The AI-Driven Chip War

Rising tensions between the US and China are catalyzing a supply chain schism for the AI chips critical to national competitiveness. This divide has been exacerbated by strategic moves such as the CHIPS Act and increasingly stringent export controls, which have particularly targeted the semiconductor sector—a vital component of AI development. Efforts by companies like Nvidia to adapt by launching China-specific chips were thwarted by new US restrictions, leading Chinese companies to turn to domestic suppliers such as Huawei. Additionally, Dutch firm ASML canceled shipments of advanced semiconductor manufacturing equipment to China under US pressure, highlighting efforts to curb China's access to crucial AI development technologies. The standoff has prompted China to explore alternatives like RISC-V, an open-source chip architecture, as a means to bypass international restrictions. This move has sparked debate in the US about the potential risks of technology transfer and the feasibility of restricting contributions to RISC-V due to its global, royalty-free nature.

The intensification of the Chip War is leading to a bifurcation in the AI chip market, with potential long-term implications for global technological advancement and cooperation. This divergence not only underscores the strategic importance of semiconductors in national security and AI development but also hints at the emergence of distinct technological spheres, each aligned with divergent national values and priorities.



The chip war could force a seismic restructuring of international manufacturing supply chains, trade flows, and technology innovation networks.

Image credit: Future Today Institute and Midjourney.

COULD AI BE INVOLVED IN—OR CAUSE—A HOT WAR?

Advancements in artificial intelligence are reshaping modern warfare in unprecedented and concerning ways. Militaries worldwide are exploring how to best leverage AI for tactical advantages, including through autonomous weapons systems, wargaming simulations, and automated hacking tools. However, these technologies raise pressing ethical issues and could dangerously escalate conflicts. The complex tradeoffs surrounding AI and defense boil down to a central tension: harnessing potential benefits to national security versus controlling for geopolitical risks.

Autonomous Weapons Policies

The US Department of Defense recently updated its guidance on autonomy in weapons systems. The original 2012 policy, and a 2017 update, did not explicitly mention AI. The DOD updated its AWS definition by removing references to a "human operator" and replacing it with simply "operator," a subtle yet notable shift clearing the way for future systems with decreased human oversight. This new directive is aimed at helping to clarify the process for developing autonomous or semi-autonomous weapons systems. Previous policies, such as the Ethical Principles for Artificial Intelligence (2020) and Responsible Artificial Intelligence Strategy and Implementation Pathway (2021), were intended to guide decision making for the development and deployment of AI within the military.

The policy change comes on the heels of other recent government actions addressing military AI. In late 2022, NATO released its Autonomy Implementation Plan, arguing AI systems offer clear opportunities for alliance members while outlining a roadmap for adoption. Additionally, the First Committee of the UN General Assembly adopted a draft resolution in 2023 calling for the UN secretary-general to conduct a comprehensive study of lethal autonomous weapons. The committee instructed the secretary-general to consult member states and civil society on addressing humanitarian, legal, security, technological, and ethical concerns related to autonomous weapons.

Simulating Warfare

Given the rising tensions between the US and China over Taiwan, several groups are building AI-powered simulation tools to wargame a future conflict. In China, the People's Liberation Army has been using AI simulation tools to prepare for military operations against Taiwan.

The Center for Strategic and International, a bipartisan, nonprofit policy research organization, developed a wargame involving an amphibious invasion of Taiwan. After 24 rounds of gameplay, the US and its allies Japan and Taiwan successfully defeated a conventional amphibious invasion by China. While Taiwan remained autonomous in the simulation, its economy was devastated and the US lost hundreds of aircraft and tens of thousands of lives--while the Chinese Communist Party never really destabilized. Games that use real-world data to run simulations are augmenting the work of military strategists, so that leaders can validate or revise their postures on deterrence, invasion, and defense.

AI Used to Guide Military Strikes

In 2021, the US military said that it had started using AI to guide its airstrikes, deploying algorithms to a live operational kill chain. The kill chain is a process of gathering intelligence, performing analysis, weighing risks, and deploying weapons to destroy a target. Using a modified process, an AI system was deployed into the Air Force Distributed Common Ground System to analyze troves of intelligence, which would have required a significant amount of human hours to complete. The new AI system cannot order a strike on its own, but it is now automatically identifying possible targets.

COULD AI BE INVOLVED IN—OR CAUSE—A HOT WAR?

Automated Target Recognition

Lethal autonomous weapons systems, powered by AI, are capable of finding targets autonomously and making decisions to complete a mission. In 2022, a lieutenant colonel in the Ukrainian military said that he and a group called Aerorozvidka had developed special drones that make use of automated target recognition. While it's unclear whether Aerorozvidka actually carried out test missions, the fact remains that machine learning–based vision for automated target recognition already exists. In response, 70 nations delivered a joint statement at the UN General Assembly calling for a ban on autonomous weapons––but little progress has been made in the months since.

Automating Offensive Attacks Using AI

Thanks to advancements in AI, one of the big trends in security is automated hacking— in short, software that's built to out-hack the human hackers. DARPA launched a Cyber Grand Challenge project in 2016, with a mission to design computer systems capable of beating hackers at their own game. DARPA wanted to show that smarter automated systems can reduce the response time—and fix system flaws—to just a few seconds. Spotting and fixing critical vulnerabilities is a task that might take a human hacker several months or even years to complete, and yet the machine that won the Grand Challenge did it in just a fraction of that time.

The winner became the first nonhuman entity to earn the DEF CON's Black Badge, which is the hacking community's equivalent of an Oscar. Very soon, malicious actors will create autonomous systems capable of automatically learning new environments, exposing vulnerabilities and flaws, and then exploiting them for gain—or whatever the stated objective, which could simply be generalized mayhem.

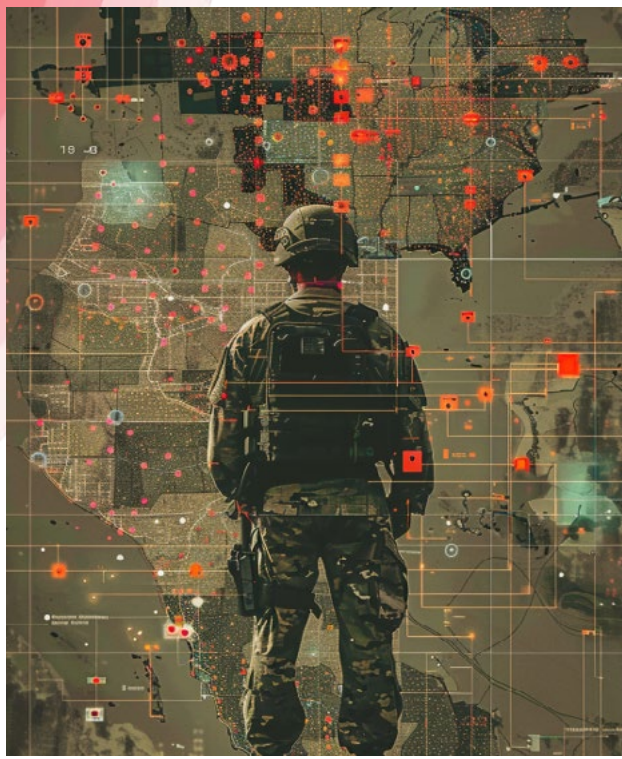
AI-Assisted Situational Awareness

Ukraine has become a test bed for modern AI-enabled battlefield awareness. Geospatial intelligence leverages neural networks to combine satellite imagery, social media posts, and other open-source data into strategic insights. By fusing photographs, drone footage, and overhead views, AI integrates distinct perspectives into a unified assessment of terrain and enemy movements.

This augmented analytics empowers a new paradigm of cost-effective warfare centered around drones. Affordable models either commercially sourced or improvised as DIY drones generate valuable intelligence rivaling America's far more expensive Reaper and Predator UAVs. Tight integration with cutting-edge systems like Delta further multiplies impact. After proving effective in 2022 trials, Ukraine greenlit Delta's full-scale February deployment. Pulling sensor, aerial, and ground reports into a consolidated data lake, this cloud-based architecture furnishes commanders with an integrated common operating picture for tactical decisions.

Algorithmic Warfighting

Future wars could be fought entirely in code, using data and algorithms as powerful weapons. The current global order is being shaped by artificial intelligence, and the



AI is enabling the development of weapons that can select targets and attack on their own. The UN General Assembly has called for banning this type of autonomous attack technology, but so far there has been no ban put in place.

Image credit: Future Today Institute and Midjourney.

COULD AI BE INVOLVED IN—OR CAUSE—A HOT WAR?

same countries leading the world in AI research—the US, China, Israel, France, Russia, the UK, and South Korea—are also developing weapons systems that include at least some autonomous functionality.

In 2020, the US Air Force successfully flew an AI copilot on a U-2 spy plane in California, marking the first time in the history of the DOD that an AI algorithm trained to execute specific in-flight tasks was deployed. With the call sign ARTUμ, it was the mission commander—though the flight was just practice.

Future Today Institute analysis shows that the future of warfare encompasses more than traditional weapons. Using AI techniques, a military can "win" by destabilizing an economy rather than demolishing countrysides and city centers. From that perspective, China's unified march to advance AI puts the emerging superpower dangerously far ahead of the West.

Mandating Ethics Guidelines for Tech Contractors

Project Maven was developed to enlist AI to analyze surveillance video. Initially, Google was the DOD's vendor, but when employees found out they'd been working on a military project, thousands protested. It wasn't the first time tech contractors had lost trust in the government.

As a result, the Defense Innovation Unit is enforcing "responsible artificial intelligence" guidelines that vendors must adopt when building AI systems, models, or applications for the DOD. The guidelines offer specific instructions that must be followed during planning, development, and deployment, which include provisions for risk assessment. This represents a longer-term trend: government agencies requiring transparency in AI projects.

The future of warfare may largely involve fighting via cyberattacks powered by AI systems rather than conventional physical weapons. Militaries have started using AI as co-pilots in spy planes and drones, pointing to increased AI integration in defense.

Image credit: Future Today Institute and Midjourney.

REGIONAL APPROACHES

COUNTRIES TRY TO REGULATE AI, BUT PLANS DIVERGE

Governments worldwide are trying to balance maximizing AI's benefits with mitigating its risks by establishing regulatory frameworks. So far, 31 countries have passed AI regulations and 13 more are debating AI laws. There are significant divergences between each country's distinct approach to regulating the technology. Some nations, like Israel, Japan, and Australia, have focused on revising existing laws to facilitate AI development, while others, like the UAE, are crafting broad national AI strategies with minimal regulatory emphasis. Countries, like Russia, Iran, North Korea, Syria, and Iraq, have opted to outright ban specific services like ChatGPT. The EU's AI Act categorizes systems by risk levels and restricts the highest risk applications. Like the EU, China has introduced AI-specific legislative frameworks, but unlike the EU, the frameworks are centered on enforcing "socialist core values" in AI. Brazil's draft AI policies prioritize user rights and risk assessments, differing from Israel's model underscoring responsible innovation and sector-specific oversight. The UAE's national strategy concentrates heavily on expanding AI integration rather than regulation.

On October 9, 2023, the US Bureau of Industry and Security of the Department of Commerce added 28 Chinese entities, including eight leading technology companies, to its entity list for their involvement in human rights violations against Uighur Muslims in Xinjiang, a move that China condemned as interference in its internal affairs. This listing requires US companies to obtain licenses before exporting certain technologies to these entities, aiming to address national security concerns without automatically imposing a full embargo, reflecting a significant effort to regulate the flow of sensitive technologies to organizations implicated in unethical practices. The impact on these Chinese companies and China's AI industry could be disastrous, depending on enforcement measures and these companies' reliance on US technology, worsening US-China relations.

These complex tensions parallel past situations like GDPR in Europe, where large multinational companies often end up defaulting to the most stringent regulations globally even if not universally binding. This scenario could plausibly unfold with major players standardizing elements of higher-bar AI governance models like the EU's for consistency. The lack of alignment across the proliferating patchwork of national and regional AI laws risks hampering innovation and global collaboration. But ironing out conflicts poses immense challenges given different priorities surrounding development versus human rights and ethics.

HOW IS THE US SPECIFICALLY REGULATING AI?

The explosion of AI technologies is leading to both fascination and concern among federal legislators, who are now exploring regulatory responses without a clear consensus. Last October, the Biden administration issued an executive order to ensure the safe and trustworthy development and use of AI, covering a wide range of AI systems beyond just generative AI and neural networks, affecting organizations across all economic sectors. Going forward, the National Institute of Standards and Technology (NIST) will play a key role in establishing guidelines for AI systems, prompting organizations to assess their use of AI and their reliance on AI-enabled products and services from third parties, and to align their AI risk management frameworks with NIST standards. But for now, there is no clear enforcement mechanism in place to check for compliance. Meanwhile, Congress is deliberating how to approach AI's dual-edged sword of opportunities and challenges, as it looks to local legislatures for precedents. So far, more than 30 states have enacted laws addressing AI in diverse ways, from specific policy concerns to establishing bodies for studying AI's impact. Senate Majority Leader Chuck Schumer and others have emphasized the need for AI regulation through initiatives like the AI Insights Forum, signaling a bipartisan understanding of its necessity. Some lawmakers consider the European Union's AI Act a model for comprehensive regulation, suggesting that the US might follow with a similar framework to manage AI's growing influence in society. State laws vary, focusing on consumer data privacy, combating AI-driven discrimination, especially in hiring practices, and addressing the manipulation of media in elections, with some states already implementing or proposing legislation to restrict deceptive AI-generated content. The US will likely adopt a bottom-up patchwork quilt of AI regulations instead of one sweeping law, like the EU's AI Act. The US government will likely boost spending on AI and AI research, especially in defense and intelligence, and use its buying power to shape the market.

A Patchwork Approach

In the US, the approach to regulating AI amounts to a diverse array of regulations that vary by state and sector, creating a patchwork framework rather than a unified national strategy. This decentralized approach results in differing standards and guidelines across jurisdictions, complicating compliance for organizations operating in multiple states and sectors within the AI landscape. Consider the current landscape of proposals and policies below.

Bias

The Algorithmic Accountability Act, a notable congressional proposal, would mandate companies to evaluate their algorithmic systems, including AI, for bias, effectiveness, and other factors if passed. Under the act, the Federal Trade Commission would be tasked with enforcing these evaluations, with a focus on preventing the use or sale of racially biased algorithms. However, the specifics of the FTC's enforcement strategy remain undefined.

HOW IS THE US SPECIFICALLY REGULATING AI?

Several US states and the District of Columbia are enacting or proposing legislation to prevent AI and algorithmic decision-making tools from reinforcing societal discrimination. Laws are being passed primarily in Democratic-led states, focusing on areas like insurance, employee surveillance, and hiring practices. For instance, Colorado mandates insurers to disclose and manage risks of algorithm use to ensure fair coverage. Massachusetts is considering a ban on AI-based employee surveillance technologies. In D.C., proposed laws would restrict service eligibility decisions made by algorithms and require user notification about how their data is used. Additionally, New York City and some states are addressing AI's role in hiring, requiring bias audits and transparency in the use of automated decision systems.

Copyright

The US Copyright Office has ruled that AI-generated content typically doesn't qualify for copyright protection as it's not human-created. But this could change now that tools like Sora pose new threats to the film industry.

Privacy

Proposals at the federal level include a complete prohibition on using personal data for targeted advertising and FTC-mandated data minimization, restricting websites to collect only data pertinent to their specific functions. At the state level, at least 12 states have enacted regulations governing automated decision systems, including AI, for profiling consumers based on personal data. Virginia's 2021 Consumer Data Protection Act is a pioneering example, mandating risk assessments and consumer rights protections when entities process over 25,000 people's data for profiling posing heightened harm risks. States are increasingly following Virginia's model by instituting similar regulatory frameworks surrounding data-driven automated systems. Additionally, some jurisdictions like New York City are specifically restricting AI usage in hiring practices through measures like bias audits and candidate notifications when screening algorithms are deployed.

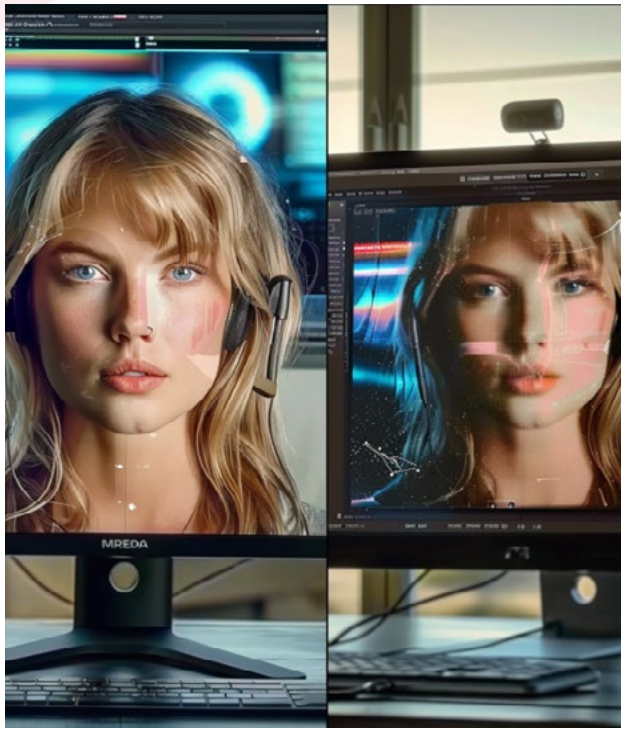
Deepfakes

The US National Defense Authorization Act includes provisions that address the growing problem of deepfakes, requiring the Department of Homeland Security to issue an annual report for the next five years on the risks posed by deepfakes. In 2021, the US Senate Committee on Homeland Security and Governmental Affairs voted unanimously to advance the Deepfake Task Force Act, which would establish a public-private team to investigate technology strategies and to develop policies that could curb risk.

Numerous states, including California, Texas, Minnesota, and Washington, have enacted laws, while New York, New Jersey, and Michigan have proposed legislation, aimed at either prohibiting or requiring disclosure of manipulated media. Many of these measures are intended to prevent public deception regarding political candidates or to influence election outcomes.

Misuse

The topic of auditing misuse is also on the



The US government is concerned about the potential misuse of deepfake technology, especially with elections approaching, which is why the latest National Defense Authorization Act includes new provisions aimed at tackling the challenges posed by increasingly realistic synthetic media.

Image credit: Future Today Institute and Midjourney.

HOW IS THE US SPECIFICALLY REGULATING AI?

congressional agenda. Legislators like Sens. Ted Budd and Ed Markey are pushing for the Department of Health and Human Services to assess AI's biological risks and develop strategies against its use in bioweapons or artificial pandemics.

Licensing

The concept of licensing requirements for AI, akin to the stringent regulation of food and pharmaceuticals, is gaining momentum. Inspired by Andrew Tutt's 2017 proposal, this approach suggests an agency could enforce pre-market approval for algorithms in certain applications, effectively requiring a government license before public deployment. Prominent figures like OpenAI's Sam Altman, Rand Corp.'s Jason Matheny, and New York University's Gary Marcus have supported such licensing, drawing parallels with the Food and Drug Administration's model.

Compute

In October 2023, the Bureau of Industry and Security introduced updated rules to enhance its October 7, 2022, regulations. These revisions focus on closing loopholes in the existing policy, further limiting China's access to advanced AI semiconductors and manufacturing tools. This move strengthens the US strategy to impede China's military AI development.

Conflicting Views About Institutional Roles

Debate persists over whether new institutions should be formed to oversee AI development and safety, or if this responsibility should fall to existing agencies. Proponents of new institutions argue they could have a major positive impact, just as bodies like the National Transportation Safety Board did for transportation safety. Suggestions for new AI oversight bodies include a National Algorithms Safety Board to monitor and ensure safety in algorithmic and AI systems, a federally funded National Artificial Intelligence Research Resource (NAIRR) to support and coordinate AI research, an international collaborative facility modeled after CERN to attract top global AI talent and focus efforts on pursuing AI safely, and an organization similar to the International Atomic Energy Agency (IAEA) to govern superintelligence efforts and safe AI deployment worldwide. A 2020 law has already mandated the creation of a task force to design NAIRR.

However, others argue existing institutions could handle AI oversight without requiring new bureaucracies. Examples include government bodies like the National Institute of Standards and Technology, FDA, Securities Exchange Commission, and Federal Communications Commission, as well as beefing up AI and tech expertise within established scientific institutions such as the Department of Energy, National Science Foundation, and NIST. Critics argue new institutions may not be more effective than today's agencies, citing issues faced by analogous bodies like IAEA in comprehensively monitoring relevant technologies globally. There are also questions around feasibility of meaningfully tracking AI development as opposed to physical materials.

At least 12 states have passed laws mandating government or related entities to research AI to enhance understanding and assess potential impacts. While some of these initiatives delay targeted regulation, others have led to tangible steps. For instance, Vermont's Artificial Intelligence Task Force's analysis resulted in the state's Division of Artificial Intelligence, which annually reviews AI usage and its effects within state government.

Public-Private Partnerships

The abundance of AI job listings across practically every American industry signals surging demand for related skills. The White House has issued a call for AI talent to join the federal government, following President Biden's executive order for the safe, secure, and ethical development and use of AI. This initiative seeks experts to help implement AI technologies across various government sectors to enhance services, ensure AI safety and equity, and maintain the country's leadership in AI innovation. Many US fabs, which are funded in part by the Chips Act, face construction delays because of a shortage of skilled workers.

The US has enlisted allies like Japan, the

HOW IS THE US SPECIFICALLY REGULATING AI?

Netherlands, and Germany to tighten their own export regimes. While competing nations pursue more centralized strategies, the US distinctive edge lies in decentralized AI innovation across promising startups and tech giants alike.

Large companies have always lobbied to influence policy and regulation. But as the tech giants amass power and wealth, they are making key decisions that impact diplomacy and geopolitics.

Big Tech companies are standing up departments dedicated to geopolitics. Microsoft President Brad Smith regularly meets with heads of state, and in 2023 played a key role at the World Economic Forum's Annual Meeting. Smith developed an international treaty called the Digital Geneva Convention to protect citizens against state-sponsored cyberattacks. Microsoft's Digital Diplomacy Group actively works on a tech-focused approach to foreign policy. The company sees corporate foreign policy as good business that builds trust and enables long-term planning. Meta, Google, Amazon, Salesforce, and many other companies are now building teams centered on geopolitics and digital diplomacy. More than a dozen countries are creating ambassador-like positions charged to negotiate with the leaders of US Big Tech companies, with the aim of mediating disagreements, collaborating on shared interests and developing public-private alliances. The longer-term implications of corporations influencing global politics could be profound. What if a company's priorities differ from the national priorities of its home government?

National Security

While late to consider AI as a national security issue, the US is quickly playing catchup. The Pentagon is considering the creation of an extensive network that leverages AI along with drones and autonomous systems within the upcoming two years, aimed at mitigating threats posed by Russia, China and non-state actors. The White House has already initiated a multimillion-dollar cybersecurity competition aimed at encouraging the adoption of AI to identify and repair vulnerabilities in the US government's infrastructure, responding to the increasing deployment of the technology by hackers for nefarious activities.

The Pentagon is considering plans to build an expansive network utilizing AI and drones over the next two years to address threats from Russia, China, and non-state groups.

Image credit: Future Today Institute and Midjourney.

WHAT IS CHINA DOING?

China is an undisputed global leader in AI. Under President Xi Jinping, the country has made tremendous strides in many fields, but especially in AI. Businesses and the government have collaborated on a sweeping plan to make China the world's primary AI innovation center by 2030, and it's making serious progress toward that goal. That plan is unlikely to be repealed by a new government; China abolished Xi's term limits and will effectively allow him to remain in power for life.

Within the next decade, China plans to meet two crucial milestones: By 2027, its People's Liberation Army will have a modern-ready force, and by 2030 the Chinese Communist Party (CCP) expects to have outpaced the US in AI and become the dominant force. China is producing what it calls "intelligentized" technologies to bolster both its economy and military.

Recently, China took major steps to shape the future of AI by releasing its own pretrained models, and it is forging ahead with its own natural language processing models, which makes sense since the most popular models in use now are trained on English text. China now has at least 130 LLMs, which accounts for 40% of the global total, closely trailing the US. Despite this rapid development, investors and analysts caution that many of these models lack sustainable business strategies, offer similar functionalities, and face rising operational expenses.

China's Expanding Market

It's a challenging time for Chinese startups because of rising tensions with the West. Companies hoping to gain traction in Europe are making efforts to cloak their origin. Shein, the e-commerce website popular among teens, says it was "founded in L.A.," but the company actually got its start in Nanjing and Guangzhou by relying on the region's manufacturing centers and ample supply chains. Or look at TikTok, which has said it's a US-based company—while the app's parent Chinese company ByteDance has employed linguistic gymnastics to separate itself. Binance, the world's largest crypto exchange, which was created in China, says that it doesn't have a headquarters located in one physical location.

It's no wonder that as Chinese startups hope to expand globally, they're seeking to distance themselves from the authoritarian regime in Beijing. But that creates political hurdles, especially as the CCP seeks to bring its home-grown technology ecosystem into lockstep with party leaders.

The result could be a future parallel universe, in which Chinese-created AI systems are shaped both by enormous amounts of data and local laws. In Brazil, a generative AI system might write an unfettered political essay in Portuguese about a leader--while in China, that same essay would be automatically filtered for politically sensitive words and phrases. As the CCP enforces new regulations targeting AI and what the government calls "deep synthesis tech," the ways in which people experience and work alongside AI could be dramatically different.

China's Big Tech

Alibaba, Tencent, and Baidu, which have made important advancements in AI research, may find it difficult to keep innovating. Starting in 2020, the CCP initiated a wave of legislation aimed at its tech sector, introducing anti-monopoly legislation focused on the platform economy and promoting data security and privacy laws. The Personal Information Protection Law (PIPL), China's version of the EU's GDPR, went into effect in 2021. What followed were a series of crackdowns targeting some of

WHAT IS CHINA DOING?

China's most successful tech companies. Ultimately, this regulation wasn't about "breaking up" China's Big Tech—the CCP wanted to focus its tech sector on achieving research and development goals set by the government and military within the decade.

Increasingly, Beijing is pressuring its mega-successful big tech companies to share data with the state and to perform research to support the vision of the CCP. Going forward, Beijing aims to direct the might of its tech companies at programs of national strategic importance rather than making video games. China's tech crackdown could cool private investment in Chinese companies, which could result in a chilling effect on innovation and economic growth, and also free up capital for emerging markets.

Deepening International Ties

China is actively building out AI infrastructure and ecosystems, specifically focused on developing nations. By focusing on the infrastructure and the ecosystem, Beijing is not just setting the stage—it's constructing the entire theater to establish Chinese-designed AI systems.

Over 140 cities globally, from Kuala Lumpur to Nairobi, are being transformed into "smart cities" and "safe cities" powered by AI. Chinese companies are providing the technology and expertise to supercharge aspects like transportation, logistics, and law enforcement. China already leads the world in exports of AI-enabled surveillance systems. China's "Luban workshop" initiative is another strategic move by China, offering vocational training globally that includes AI education. This has resulted in the creation of a workforce skilled in AI in various developing nations. China also created a "BRICS AI Study Group" to accelerate AI cooperation with other developing economies. Chinese tech companies even helped construct the premier AI company in the UAE. Additionally, China dominates the market for industrial robot installations, having surpassed Japan in 2013. The gap between China and other countries has only widened—in 2021, China installed more industrial robots than the rest of the world combined.

This strategy mirrors China's Belt and Road initiative but instead of building physical infrastructure in developing nations to increase influence, China is building the technological infrastructure, which includes skills and data flow. However, US export controls on key semiconductors and technologies to China present obstacles. In response, China has taken measures such as prohibiting the use of chips from American company Micron in its infrastructure and implementing a licensing system for the export of specific essential metals like gallium and germanium, which are crucial for high-end semiconductors as well as components in solar panels and electric vehicles.

As China shapes the world order in its own image, it is simultaneously exporting its technologies and surveillance systems to other countries with authoritarian regimes. When the CCP expands into African countries and throughout Southeast Asia and Latin America, it will also begin to eschew

China is funding smart cities powered by Chinese AI and surveillance technology in developing nations. The country is exporting advanced monitoring systems as part of a broader strategy to extend its technological and geopolitical influence.

Image credit: Future Today Institute and Midjourney.

WHAT IS CHINA DOING?

operating systems, technologies, and infrastructure built by the West. Two Chinese companies— the state-controlled CEIEC and Huawei—built Ecuador's surveillance system, called ECU-911. The system promised to curb high murder rates and drug crime, but Ecuador could not afford the investment. As a result, a deal was struck for a Chinese-built surveillance system financed with Chinese loans. It was a prelude to a much more lucrative deal: Ecuador eventually signed away big portions of its oil reserves to China to help finance infrastructure projects. Similar package deals have been brokered in Venezuela and Bolivia.

China is quietly weaponizing AI, too. China's People's Liberation Army is catching up to the US military, using AI for such tasks as spotting hidden images with drones. The Chinese military is equipping helicopters and jet fighters with AI. The government created a top-secret military lab—a Chinese version of DARPA—and it's building billion-dollar AI national laboratories. China's military is achieving remarkable AI successes, including a recent test of "swarm intelligence" that can automate dozens of armed drones.

When it comes to AI, leaders should monitor escalating tensions between the US and China. But they should also remember that there are cells of rogue actors who could cripple our economies simply by mucking with the power or traffic grids, causing traffic spikes on the internet, or locking us out of our connected home appliances. These aren't big, obvious signs of aggression, and that is a problem for many countries, including the US. Most governments don't have a paradigm describing a constellation of aggressive actions. Each action on its own might be insignificant. What are the escalation triggers? Without a definition, a strategic vulnerability exists.

China is quickly advancing military applications of AI, recently demonstrating swarm intelligence capabilities to coordinate actions of dozens of armed drones.

Image credit: Future Today Institute and Midjourney.

WHAT IS EUROPE DOING?

In late 2023, The European Union finalized negotiations on its landmark AI Act. This legislation establishes the world's first comprehensive framework for regulating AI systems. The overarching goals are to guarantee AI safety, uphold ethical standards, and drive European AI leadership. Specifically, the EU AI Act classifies AI systems into different risk categories based on their use cases.

In February 2024, a new European AI Office, established within the European Commission, was announced to promote the development and use of safe and trustworthy AI across the EU, functioning as the core of a unified European AI governance system. Through the implementation of the AI Act, the office aims to safeguard health, safety, and fundamental rights, providing a stable legal environment for businesses in all 27 member states. It will be responsible for monitoring compliance and in enforcing AI regulations.

France aims to advance its AI capabilities and influence. President Macron promised more than $500 million to cultivate French AI "champions" and counter Silicon Valley's English-dominance in AI systems. Mistral, a Paris-based AI company founded by Arthur Mensch, Guillaume Lample, and Timothée Lacroix, former AI researchers at Meta and DeepMind, is gaining attention for its rapid growth and focus on developing smaller, high-performance AI models as an alternative to giants like OpenAI. Unlike some larger, more restrictive models, Mistral's offerings can run locally with open weights, allowing for more accessible and flexible use in multiple languages for various tasks, claiming to outperform or match other leading models on certain benchmarks. The company uses a novel mixture of experts (MoE) architecture, enhancing efficiency by routing tasks to specialized neural networks, making processing faster and less resource-intensive. Mistral made its models available for public use under the Apache 2.0 license via Hugging Face and BitTorrent—yes, the same BitTorrent that gained notoriety housing illegally copied movies and music and allowing downloads via its peer-to-peer network—and the company recently launched beta access to its API for different levels of Mistral models.

Germany also recognizes the geostrategic importance of AI innovation to compete with American and Chinese tech giants. A new hub in the southeast city of Heilbronn aspires to be a startup epicenter applying AI to help German industrial leaders stay competitive. Germany has committed nearly 500 million euros toward AI research and innovation, aiming to enhance supercomputing infrastructure, skill development, and create 150 new professorships, with a focus on achieving "technological sovereignty" and reducing its dependency on external powers. German Federal Minister of Education and Research Bettina Stark-Watzinger is lobbying for EU-wide cooperation in AI, particularly between Germany, France, and Scandinavian countries, to position Europe at the forefront of the global AI landscape. Despite all these commitments, concerns linger about the slow

WHAT IS EUROPE DOING?

pace of integrating AI into the broader economy and the potential stifling effect of the EU's AI Act on innovation, highlighting the need for more effective transfer of research to practical applications and the creation of a robust AI-specific infrastructure.

Brexit continues to complicate Europe's AI landscape. The UK government, following a white paper it published in March 2023, decided against introducing new AI-specific legislation, opting instead for a pro-innovation regulatory framework that leverages existing regulatory powers to manage AI technologies. This approach emphasizes high-level principles such as safety, transparency, and fairness to guide regulators, without imposing statutory duties to ensure flexibility and adaptability in AI oversight. One area still up for debate is intellectual property across news and entertainment media. The House of Lords have called for standardized regulatory powers and meaningful sanctions to deter wrongdoing––without explaining what oversight would need to entail, or how innovation can still be counted on to stimulate the UK economy.

Finally, let's not forget Russia, which seeks to counter Western dominance in AI. In November 2023, Russian President Vladimir Putin announced plans for the development of an AI national strategy, stressing that its focus would be to prevent Western monopoly. He criticized the "monopolistic dominance" of foreign technology in Russia as unacceptable and dangerous, highlighting that many AI systems are trained on Western data, reflecting ethics and norms that the Kremlin opposes. Putin warned against the "digital cancellation" of traditional Russian culture by Western AI algorithms, which he claimed often exclude or ignore Russian contributions to culture, science, and literature. He pledged significant investment in supercomputers and other technologies to enhance national AI research, underscoring the need for AI developments to be grounded in Russian traditional values and cultural heritage.

Putin is justifiably worried about adopting a Western paradigm of AI. Models like ChatGPT are trained overwhelmingly in English and are likely to exhibit the same assumptions as English-language media that could contradict official narratives peddled by Russian media.

Major Russian tech companies like Yandex and Sberbank are racing to build their own rivals to ChatGPT. But their offerings already lag behind the accelerating innovation of US and Chinese tech giants. Western sanctions further hamper access to vital computing power. Perhaps most critically, Russia's authoritarian atmosphere of censorship and distrust conflicts with the very nature of imaginative, generative AI.

WHAT IS THE MIDDLE EAST DOING?

The United Arab Emirates is positioning itself as a neutral ground for the advancement of artificial intelligence, aiming to bridge the gap between the US and China amidst ongoing geopolitical tensions. To date, the UAE government has shown deft diplomatic skills in navigating complex international relations that increasingly involve AI and other critical technologies.

Though the UAE government has worked hard to remain neutral, its companies are still caught in the crosshairs between the ongoing race between the US and China for AI supremacy. A major innovator based in Abu Dhabi, G42, develops advanced technologies across sectors like space, health care, energy, and security, but in December 2023, it faced growing pressure to cut ties with hardware suppliers such as Huawei. The decision to phase out Chinese hardware was also a move to preserve G42's access to US-made chips.

Also late in 2023, the government launched a new state-sponsored AI company, AI71, to commercialize its leading LLM, Falcon. AI71 aspires to directly compete with leading AI labs like OpenAI. The UAE is also focusing on nurturing its homegrown talent in AI by investing in specialized education. It established the Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), recruiting renowned experts from institutions like University of California, Berkeley and Carnegie Mellon as its faculty. The university produces scores of graduates annually, most of whom work at local Emirati technology companies.

In parallel, the Kingdom of Saudi Arabia has embarked on its own sweeping economic diversification agenda centered around AI. Through its Vision 2030 plan, the kingdom seeks to position itself for a future where the global economy is less dependent on oil and more driven by technology and innovation. The crown jewel of Vision 2030 is Neom, a futuristic megacity under construction aiming to seamlessly integrate cutting-edge technologies like robotics and AI across all aspects of daily life.

The kingdom is backing its AI ambitions with significant investments, including $20 billion specifically earmarked for advancing artificial intelligence. It established the Saudi Data and Artificial Intelligence Authority (SDAIA) to drive the national AI strategy. SDAIA initiatives like the National Center for Artificial Intelligence are designed to make Saudi Arabia an AI leader across priority industries such as health care. Global tech giants have taken note, with China's Huawei recently launching a new cloud data center in Riyadh to grow its digital offerings in the region. The facility will support AI applications and Arabic language models to power government services. Though Huawei's expansion may benefit Saudi AI progress in the near term, Washington is likely to view such collaborations with concern given its wider technology rivalry with China.

The governance structures in KSA and the UAE allow for swifter decision-making and implementation of technology strategies compared to democracies, where public opinion on

WHAT IS THE MIDDLE EAST DOING?

issues like privacy and employment significantly influences policy. These nations have the financial resources to invest heavily in essential AI components like GPUs, having spent hundreds of millions on them, as well as the energy required to power these intensive processors.

Many Middle Eastern nations are positioning themselves as emerging AI hubs, aggressively investing in AI skills development, research, and entrepreneurship. The goal is to diversify their economies in anticipation of declining reliance on oil.

Image credit: Future Today Institute and Midjourney.

TALENT

WHERE AND HOW DO I GET AI TALENT?

As AI permeates industries, demand has soared for technical talent to build and deploy AI capabilities. However, the rapid pace of innovation has yielded a global AI skills shortage. Employers struggle to attract talent, especially when competing against prestigious technology giants with deep pockets. This breeds uncertainty on optimal strategies, leading executives to wonder about the precise mix of skills needed in their workforce, whether current hiring approaches apply for burgeoning AI roles, and how to evaluate the technical caliber required for AI related work.

Demand for AI-related Skills Increases Across Sectors

Employer demand for AI skills is rising rapidly across nearly every industry in the US. With the exception of sectors such as agriculture, forestry, fishing, and hunting, there has been a noticeable uptick in AI-related job postings—from 1.7% of all postings in 2021 to 1.9% in 2022. Employers are actively seeking individuals proficient in machine learning, which tops the list of in-demand AI skills, followed by knowledge in artificial intelligence and natural language processing. The surge in these specific areas underscores a shift in the job market, with AI skill clusters achieving greater prominence than they had a decade prior. Demand for Python skills has also increased, evidence of its growing popularity as an AI coding language. This increased AI skills demand is not isolated to the US; it reflects a global trend. The US leads globally for AI-related job postings, followed by Canada and Spain. Furthermore, LinkedIn's AI skill penetration rate metric, which assesses the prevalence of AI skills across occupations, has revealed that as of 2022, the regions leading in AI skill penetration are India, the United States, and Germany. The increasing prevalence of these skills points to a transformative phase in the job market, where AI proficiency is becoming a critical asset for professionals in a multitude of sectors.

AI Brain Drain from Academia

A striking shift has occurred in where AI Ph.D. graduates build their careers. The Artificial Intelligence Index Report from Stanford shows an increase in the proportion of AI Ph.D. graduates in North America entering the industry after graduation, from 44.4% in 2010 to roughly 48% in 2019, while the percentage taking up academic positions declined from 42.1% in 2010 to 23.7% in that same period. The reason is clear: Competitive salaries offered by the private sector, along with the chance to work on applied AI research, has tempted Ph.D.s away from the classroom to corporate America. Leading AI organizations, such as OpenAI and Anthropic, are offering starting salaries for new researchers in the range of $700,000 to $900,000, according to salary negotiation service Rora, with Google even offering substantial restricted stock grants to attract top data scientists. That's orders of magnitude higher pay than what even tenured professors can expect from their universities. Top academics now earn generous corporate salaries and benefits, and they get to work in a similar tenured environment that's carefully cultivated to replicate their experience in academia.

Tech companies are also endowing AI professorships at top universities. In some cases, professors take one- or two-year sabbaticals to work at tech companies and then return to their universities. But corporate benefits can be difficult to give up, and companies need the talent. Poaching academia today could rob the future of future AI experts: Without great scholars, who will train the next generation of innovators?

HOW WILL AI CHANGE THE NATURE OF WORK?

In a clever study released last summer, researchers from the US demonstrated that shortly after ChatGPT was introduced, copywriters and graphic designers on prominent online freelancing sites experienced a notable decrease in job opportunities, along with steep declines in their earnings. Here's the rub: Generative AI wasn't only replacing their jobs, it was diminishing the value of the work they are trained to do.

It's become clear that AI will change how we do work, where we do work, and what work needs to be done in ways that are both piddling and profound. This breeds equal parts excitement and anxiety. While fears persist of mass job elimination, experts emphasize AI more as augmenting than replacing human roles—though the truth remains unclear. Furthermore, as AI becomes more sophisticated, able to collaborate with humans and demonstrate capabilities once considered uniquely human, it raises the question: Is AI a tool for workers to use or a colleague to collaborate with? The answer may determine how readily people embrace working with AI.

Gains and Pains

The integration of AI promises to both enhance efficiency for some roles while making other jobs more challenging. In the financial services sector, for instance, a study has highlighted how the integration of AI systems is increasing the demands on middle management. While AI applications efficiently handle routine tasks formerly executed by humans, this shift necessitates that managers adapt to new challenges and demands, navigating a landscape where they must balance traditional management roles with the oversight of AI systems.

Yet in medicine, AI is generally positioned as an empowering asset to augment clinicians rather than replace them. Physician-researchers at Beth Israel Deaconess Medical Center showed an AI chatbot surpassing human accuracy at probabilistic reasoning to aid diagnoses. Separately, scientists at University College London developed AI speech pattern detection tools to uncover early schizophrenia indicators. Rather than substituting the role of the physician, these technologies enable doctors to discern additional insights.

Status Shifts

Emerging research suggests AI may profoundly reshape perceptions of high-status occupations in the coming years by automating prestigious skills. Studies from the OECD and OpenAI forecast mass job losses even in respected professional domains like law, medicine, and finance. However, AI might also democratize skills that have long been associated with high-status roles. For instance, the use of AI like GPT-4 in professional services has shown that those leveraging these tools can outperform their peers across various tasks. LLMs might be just as good—or even better—at certain jobs in law firms because the work can be automated for faster, cheaper results. The use of LLMs in law could change how legal work is done, potentially reducing the need for junior lawyers for routine tasks and forcing legal process outsourcing firms to change their business models, while also offering law firms and legal departments significant efficiency gains and cost savings. This

HOW WILL AI CHANGE THE NATURE OF WORK?

suggests that AI has the potential to level the playing field, allowing individuals who might not have traditionally excelled in these roles to boost their performance and productivity.

This shift also brings a potential upheaval in the power dynamics traditionally associated with specialized skills and knowledge. High verbal intelligence, once a marker of elite status and a key to high earnings, might lose its prestige as AI begins to outperform humans in tasks involving language and writing. Skills such as writing proficiency or multilingualism, previously indicators of a highly educated individual, may diminish in value as AI improves text quality and eliminates language barriers.

The possible reduction in status and influence for those skilled in words and symbols is a significant cultural shift. For centuries, Western societies have revered those adept at conceiving and communicating new concepts. The rise of the Scientific and Industrial revolutions only amplified their status and influence. The encroachment of AI into these realms could represent a profound and unprecedented shift in societal values and the stratification of labor. In this new era, jobs that are less susceptible to automation, such as skilled trades, might gain in prestige and economic reward relative to those more easily disrupted by AI technologies.

Agents Will Increasingly Perform Tasks on Our Behalf

The workforce is increasingly familiarizing itself with chatbots to perform routine tasks like drafting emails and synthesizing technical language and documents. AI agents would take this a step further by actively performing tasks like sending emails, scheduling meetings, and booking reservations. These agents represent a shift toward a more proactive and autonomous model, transcending the capabilities of traditional chatbots. AI agents could book your vacation, coordinate dinner reservations among friends' calendars, or perform specific tasks relevant to a particular role at a company. For example, a product owner AI agent could specialize in aiding with market analysis, prioritizing features, and developing business cases. A developer AI agent would focus on automating code generation, refining existing code, and aiding in bug detection.

But these agents wouldn't be siloed to interaction with humans—they can interact with one another just as humans would to accomplish a goal. Imagine an ecosystem populated by specialized AI agents, each equipped with distinct expertise and knowledge, designed to not only assist individual tasks but also to collaborate and interact with one another. This vision points toward a future where AI agents evolve from performing singular, user-specific tasks— like drafting emails, resolving customer support queries, or managing grocery orders—to operating within a network where they communicate and cooperate with other agents.

Companies that possess extensive data repositories in specific verticals are likely to emerge as leaders in the AI Agent space. For example, Bloomberg, with its rich trove of financial data, is well-positioned to develop sophisticated financial AI agents. It's already begun creating large language models tailored to finance. Similarly, LexisNexis, with its vast legal information database, could develop AI agents specialized in the legal domain. These AI agents, drawing from deep wells of domain-specific data, would not only cater to their direct users but also become invaluable assets to other businesses, systems, and AI agents.

EMERGING CAPABILITIES

CAN AI REASON? AND HOW CLOSE ARE WE REALLY TO AGI AND ASI?

Can AI think like a human? Recent advances suggest we may be close to unlocking AI's potential for complex reasoning, and perhaps even one day achieving artificial general intelligence (AGI), shorthand for a computer that can do anything a human brain can. Microsoft recently startled the AI community, becoming the first major tech company to argue current systems exhibit sparks of AGI. While it never made that same proclamation, Google's DeepMind team has repeatedly demonstrated nascent AGI capabilities in its research.

AGI refers to flexible, human-level cognition able to tackle any intellectual task. Yet progress has been uneven across the spectrum of what a human brain is able to do. Contrary to sci-fi visions of AI's prowess at logic and math, early breakthroughs in artificial intelligence were primarily in creative realms like art and language modeling. Initially, it was believed that AI would excel in reason-based jobs, particularly in mathematics, given computers' inherent proficiency in handling numbers and calculations at a speed far surpassing human capabilities. However, the evolution of AI has taken a somewhat unexpected turn, veering more towards creative applications rather than purely logical reasoning. That is, until now. Looking beyond AGI, the ultimate frontier is artificial super intelligence (ASI)—AI that surpasses human intelligence in every aspect, from creativity to problem-solving, heralding an era where AI's capabilities could transcend human limitations.

> Will AI take over the world? No, this is a projection of human nature on machines.
>
> —Yann LeCun,
> vice president and chief AI scientist at Meta

AI Breakthroughs in Mathematics

A breakthrough in AI's mathematical abilities was showcased by DeepMind's AlphaGeometry. In a landmark paper published in Nature, AlphaGeometry demonstrated its capability to solve complex geometry problems at a level comparable to a human Olympiad gold medalist. It successfully solved 25 out of 30 Olympiad-level geometry problems within the standard time limit, a performance on par with top human competitors.

The success of AlphaGeometry highlights AI's growing capacity for logical reasoning and knowledge discovery. AlphaGeometry effectively showcases a process that mirrors real thinking. Its process has been compared to the dual-process theory of thinking, Type I and Type II, as popularized by psychologist Daniel Kahneman in "Thinking, Fast and Slow." Also from the DeepMind team, a technique involving LLMs named FunSearch has demonstrated that AI can assist mathematicians in solving wicked problems, inspired by the card game "Set." This marks the first instance where an LLM-based system has been able to surpass existing mathematical and computer science solutions, proving yet again that AI can solve a wide array of math and compsci questions more effectively than human mathematicians working alone. FunSearch works by generating and testing short computer programs for solving mathematical problems, refining its approach through feedback, and represents a novel form of human-machine collaboration that could amplify the capabilities of human mathematicians rather than replace them.

CAN AI REASON? AND HOW CLOSE ARE WE REALLY TO AGI AND ASI?

AI Persuasion

Logical argument is core to persuasion, but emotional resonance and validating existing views profoundly shape what people are convinced of too. AI shows promising aptitude on both fronts—generating seemingly rational arguments while precisely targeting psychological triggers. OpenAI CEO Sam Altman recently warned superhuman persuasiveness may arise in AI before general intelligence does, with unpredictable outcomes. AI chatbots, like OpenAI's ChatGPT, have demonstrated an impressive level of conversational prowess; they can sound convincing, even when providing incorrect information, which is particularly troubling when considering the human tendency to form emotional connections with these systems. Evidence shows even limited interactions with AI chatbots promotes attachment and trust, amplifying their capacity for conviction.

Researchers at Stanford's Polarization and Social Change Lab and the Institute for Human-Centered Artificial Intelligence conducted studies to explore AI's capabilities in swaying public opinion on contentious political issues. Their findings were alarming: AI-generated arguments were as persuasive, if not more so, than those penned by humans on a range of topics. For example, AI-crafted messages on policies like smoking bans and carbon taxes significantly shifted readers' support. While highlighting AI's influential potential, researchers in parallel sound alarms on misuse by hostile actors. As models continue absorbing the intricacies of human psychology while simultaneously continuing to improve at logic-based persuasion, safeguarding against deception emerges paramount.

Prediction and Prescience Into Our Human Lives

Will I die within four years? This is one of the questions that a collaborative research project is pushing AI to be able to answer. By leveraging large data sets detailing various aspects of people's lives and employing transformer models—similar to those underpinning the language processing capabilities of systems like ChatGPT—the researchers have developed methods for organizing this data to forecast future events in an individual's life. Notably, their model, dubbed Life2vec, can make predictions about profoundly significant events, including estimating the time frame of a person's death.

The cutting edge in AI reveals accelerating abilities to computationally interpret integral aspects of the human experience—from life outcomes to subjective thought itself. Recent research has demonstrated AI's ability to not only forecast significant life events but also to delve into the depths of human cognition by reconstructing images seen by individuals, based solely on brain scans. A team from Osaka University in Japan has achieved a groundbreaking feat in cognitive AI. By analyzing functional magnetic resonance imaging (fMRI) scans taken while subjects viewed specific images, the AI system they trained was able to recreate these images with surprising accuracy. The AI generated visuals of a teddy bear, a clock tower, and an airplane, among other objects, after participants had looked at similar items. This research marks a significant step forward in AI's ability to interpret and visualize human thoughts based on neurological data. The implications of these advancements are profound. AI's ability to predict life events suggests a future where technology could offer insights into personal and societal trends with unprecedented accuracy. Meanwhile, the capacity to reconstruct visual experiences from brain scans opens new avenues for understanding human cognition, memory, and perception.

More practically, Nvidia is developing an AI-powered "digital twin" of Earth, known as Earth-2, leveraging its FourCastNet AI model to predict weather with unprecedented speed and accuracy, outperforming traditional methods by forecasting thousands of potential outcomes. This breakthrough in climate modeling represents a huge advance in applied research.

Detecting Emotion

A new type of neural network can determine how people are feeling. Using radio waves, AI can detect subtle changes in heart rhythms,

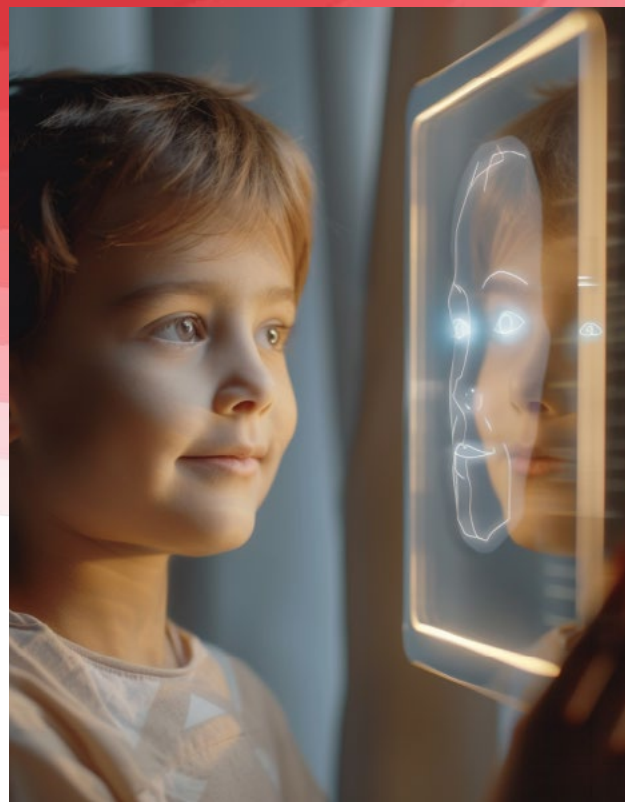
CAN AI REASON? AND HOW CLOSE ARE WE REALLY TO AGI AND ASI?

run a pattern analysis, and predict someone's emotional state in a given moment. A team from Queen Mary University of London used a transmitting radio antenna to bounce radio waves off test subjects and trained a neural net to detect fear, disgust, joy and relaxation, as people were shown different videos. The system accurately tagged emotional states 71% of the time, which signals new opportunities for health and wellness applications, as well as for job interviews and the government/military intelligence community. The EU is sponsoring a pilot project called iBorderCtrl that uses emotion recognition technology to assess truthfulness in border crossing interviews—the system analyzes interviewees' micro-expressions and nonverbal cues in an attempt to quantify the likelihood of deception during questioning. However, emotion recognition technology is still emerging and its accuracy in quantifying human emotion remains unproven, given the inherent complexity and nuance of human expression. Some of the most advanced emotion recognition technology is currently being developed in China, where extensive work has been done on facial recognition systems, albeit amid ethical concerns over potential misuse—the country has faced scrutiny for employing emotion AI to enable surveillance, most notably to monitor the Uyghur population.

Neuro-symbolic AI

Neuro-symbolic AI combines the best of two worlds in AI: the learning capabilities of neural networks (which are good at handling unstructured data like images and language) and the reasoning capabilities of symbolic AI (which deals with structured data and logic). For businesses, this means they can create smarter systems that not only learn from vast amounts of data but also understand and apply rules and logic, similar to human reasoning. In practical terms, this means that a neuro-symbolic AI could analyze a company's data and also understand the context, making decisions that are more accurate and relevant to specific business scenarios. By understanding rules and logic, neuro-symbolic AI might automate tasks that previously required human understanding, saving time and reducing errors.

Amit Sheth, who founded the Artificial Intelligence Institute at the University of South Carolina, is exploring a new idea called neuro-symbolic vision. This approach is similar to how we, as humans, understand the world: by turning what we see and hear into symbols in our minds, and then using what we know to make sense of those symbols, make plans, and take actions. This way of processing information is also how we explain our thoughts and actions to others, which is especially important in areas like health care where trust is key. Neuro-symbolic AI aims to improve how smart systems figure things out and make them more accountable. By combining the learning power of neural networks with organized knowledge (like facts and rules), we could see big improvements in AI's ability to understand concepts, make connections, and reason about the world in a way that's clear to us. As people start questioning current AI methods, this neuro-symbolic approach could lead us toward creating AI that thinks more like humans do, which could be a big step toward achieving AGI.



AI can detect emotions through facial analysis and by tracking subtle biological clues like changing heart rhythms.

Image credit: Future Today Institute and Midjourney.

IS THE FUTURE OF AI CLOUD, EDGE, OR ON-DEVICE?

Where we will ultimately deploy AI workloads remains an open question. Many anticipate the future is likely to embrace a hybrid approach that combines cloud, edge, and on-device computing in some capacity. This strategy allows for data processing and model training to leverage the vast parallel processing power of cloud servers. Meanwhile, edge hardware and local devices could handle real-time inferences and personalization, optimizing for both performance and privacy. But the specific balance across environments and when to favor one over the other is still unclear as capabilities and demands evolve.

Cloud Neutrality

A handful of companies control the cloud and have the ability to set pricing, access and standards. Those companies own the infrastructure and don't have to make their business practices transparent. Generative AI systems require enormous amounts of costly computing power and cloud infrastructure, which the tech giants are trading for future shares of profit. This consolidates additional power among the largest cloud providers. As more of our businesses and aspects of our lives move to the cloud, efforts will grow to ensure that infrastructure serves the public interest. The three biggest cloud providers, Microsoft, Amazon, and Google, have collectively invested tens of billions of dollars building infrastructure: data centers, monitoring systems and software. Their robustly designed systems prevent downtime and data loss, and few other companies in the world can compete. But the cloud isn't public infrastructure; it's private. And as private companies, cloud providers currently control access to services that are becoming the lifeblood of businesses.

Cloud Strain From AI Boom

AI has arrived, but the underpinnings of the cloud may struggle to withstand its weight. Cloud providers such as Amazon Web Services, Microsoft Azure, and Google Cloud are under intense pressure to adapt their services to accommodate the needs of large-scale generative AI models, which can be up to 100 times larger than their predecessors. Generative AI models like ChatGPT that produce original text and analysis can be 10 to 100 times more complex than a Google search. The current cloud infrastructure, primarily designed to provide scalable, pay-as-you-go services for diverse workloads through general-purpose computing, is now significantly challenged by the demands of AI-intensive workloads.

Only a small portion of current cloud servers are outfitted with AI-optimized GPUs or structured to function in collaborative clusters, essential for meeting the substantial computational requirements of AI tasks. A significant bottleneck also arises from the scarce availability of high-performing GPUs, with Nvidia essentially serving as the sole supplier. Because of high demand, Nvidia's H100 graphics—an earlier version of their most powerful graphics—sold for more than $40,000 on eBay. To reduce their dependence on Nvidia, companies like Alphabet, Microsoft, and Amazon are developing their own AI chips for model training. Despite their cloud platforms not being fully optimized for AI, AI workloads are contributing to significant revenue growth in their cloud infrastructure.

AI Breathes Life Into Legacy Systems

The rising costs associated with cloud computing, especially for tasks like training AI models, are prompting some companies to reconsider on-premises solutions. Dell Technologies, recognizing this shift, has developed servers specifically designed for on-premises AI deployments. By moving AI operations in-house, Dell argues that companies can potentially save on networking and data storage expenses. Furthermore, AI is playing a pivotal role in revitalizing legacy mainframe systems. Over 800 billion lines of COBOL code are currently in use within production systems,

IS THE FUTURE OF AI CLOUD, EDGE, OR ON-DEVICE?

making the transition from this language, established in 1959, to more contemporary languages a daunting task. The scarcity of COBOL experts—many are nearing retirement age—and the complex nature of migration efforts for large organizations further compound these challenges. IBM's introduction of Code Assistant for IBM Z, an AI-powered tool that translates COBOL code into Java, offers a solution to modernize mainframe applications with the help of AI. This blend of AI innovation not only supports the shift towards on-premises AI deployments to manage costs but also demonstrates the potential for AI to breathe new life into legacy infrastructures.

Optimizing AI to Run at the Edge

Smart devices like phones lack the memory and computing power required to fine-tune AI models with user data over time. This limitation has necessitated transmitting personal information to the cloud for updating, an energy-intensive process that risks data privacy. Now, advances like PockEngine enable efficient on-device learning without offloading data. Developed through an MIT and IBM collaboration, PockEngine is a training model that selectively identifies which specific parts of an otherwise enormous model to update locally based on a user's unique inputs. By focusing only on essential parameters and shifting computations to preprocessing, PockEngine minimizes real-time resource usage. Not only does this make it more efficient, it also facilitates the creation of personalized deep-learning models. For instance, AI assistants can continuously adapt to a user's accent or typing patterns without reliance on constant cloud connectivity. Tests demonstrate PockEngine fine-tuning complex models up to 15x faster than alternatives, all while maintaining or boosting accuracy.

Small Language Models for AI at the Edge

While large language models with billions or trillions of parameters have demonstrated impressive capabilities, smaller AI models may be better suited for edge-based use cases. Though less broadly capable, specialized mini-models bring benefits like faster inference, lower compute requirements, and easier integration into edge devices. For mobile and embedded use cases, massive cloud-based LLMs are often impractical. Their substantial size and latency makes local deployment a non-starter. More compact models in the millions or single-digit billions of parameters, however, could potentially run efficiently on smartphones and IoT devices. Your washing machine could be equipped with a compact language model, enabling you to inform it verbally that you're washing a mixed load and are concerned about a sweater washing in overly warm water. The small language model that can run in the appliance eliminates the need for internet connectivity to operate your washing machine in this manner. SLMs could therefore empower voice assistants, smart home automation, and beyond, reducing the dependency on cloud-based services for these types of applications.

On-Device AI

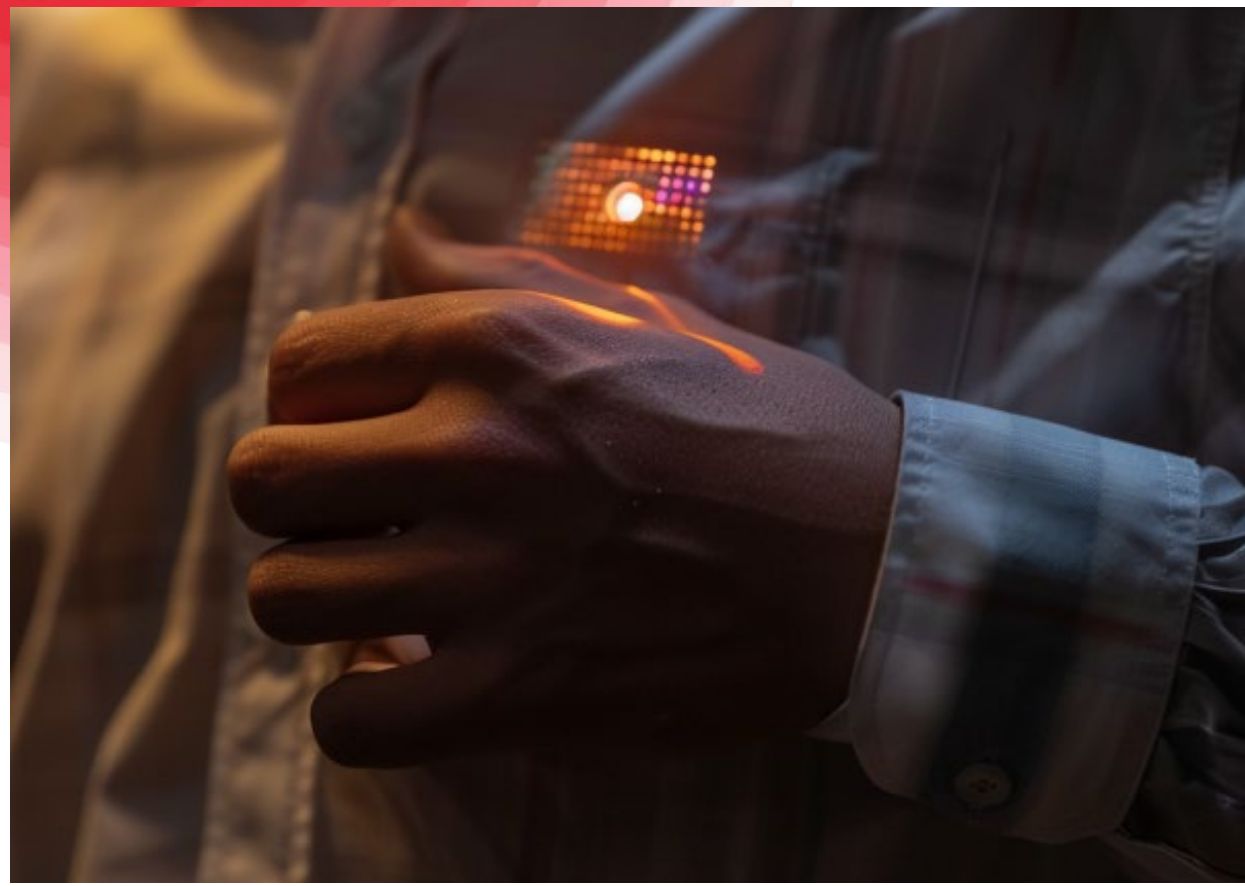
Tech giants such as Samsung, Microsoft, Google, and Apple are spearheading a movement towards on-device AI, emphasizing a blend of performance and privacy. These companies are competitively equipping their devices with specialized AI chips to enable local processing, thereby reducing reliance on cloud servers. This approach to on-device AI processing is motivated by the goal of safeguarding sensitive data, drastically cutting down the risk of data breaches during its transfer to and from the cloud. Moreover, on-device AI has the unique capability to adapt and personalize according to a user's behavior directly on the device. Samsung introduced its Galaxy S24 smartphones, showcasing a leap in AI capabilities with the implementation of generative AI tools that operate through a combination of on-device processing and cloud-based computations. Google's latest Pixel phone features custom AI silicon to handle tasks like predictive typing more responsively on-device. Apple's newest MacBook CPU incorporates neural processing units for faster machine learning. AMD's latest Ryzen mobile chips similarly target laptop enhancements like voice assistance.

IS THE FUTURE OF AI CLOUD, EDGE, OR ON-DEVICE?

Wearable AI

AI is changing human-computer interaction, shifting us away from screens, trackpads, and keyboards towards more intuitive, voice-based interfaces. This is giving rise to a new class of lightweight, wearable gadgets and screenless computers that integrate seamlessly into daily life. By reducing screen fatigue and intrusive features, these devices foster a more natural, human-centric approach to technology. A prime example is the newly launched Humane AI Pin, an OpenAI-powered wearable priced at $699, plus a $24 monthly subscription. Forgoing traditional app interfaces, this 34-gram device focuses solely on voice interactions. Users access information and perform tasks by speaking to the Pin's built-in microphone. By stripping down the technological interface, Humane aims to create a streamlined, human-like experience. Another device is the Rewind AI Pendant, which captures real-world conversations, storing encrypted transcripts and audio locally on the user's phone. Beyond recording, Rewind's platform searches transcripts, generates meeting summaries and analyzes speech patterns. Essentially, the Pendant serves as a personalized assistant harnessing environmental information to support the user. Both the Humane Pin and Rewind Pendant epitomize the shift towards invisible, assistive technology that facilitates life's tasks and interactions much like a helpful human companion would. This evolution in form and function represents a paradigm shift, integrating technology more seamlessly while making it feel more intuitive and human-centric.

Note: we've included this trend in both the AI and Computing reports. We think it is important to consider the near-future of wearables as you contemplate the future of your organization and AI's development.



AI can let us get information by voice requests rather than typing search terms or looking through folders. This more natural interaction could drive demand for wearable or voice-based interfaces.

Image credit: Future Today Institute and Midjourney.

WHY SHOULD WE PAY ATTENTION TO EMERGING CAPABILITIES THAT AREN'T YET FULLY DEVELOPED?

Businesses should keep an eye on emerging AI capabilities because these technologies can unlock new opportunities for innovation, efficiency, and competitive advantage. Early awareness and adoption of AI advancements can position a company as a market leader, enabling it to refine its operations, enhance customer experiences, and create novel products or services.

Vector Databases

Vector databases are poised to grow rapidly in importance alongside advancements in AI. As AI models like large language models become more capable of human-like generation across modalities like text, images, and audio, they rely heavily on vector representations of data, known as embeddings, to understand and generate contextual meaning. To function optimally, these generative models need databases specifically designed to store massive vector data sets and allow instantaneous retrieval of semantically similar vectors. This is where vector databases come in; they are uniquely designed to efficiently store, manage, and retrieve high-dimensional vector data, which is crucial for embedding processes found in natural language processing, image generation, and other AI applications. Unlike traditional databases that organize data in rows and columns, vector databases use vectors to represent data points, enabling faster and more relevant data retrieval based on similarity. As companies like Microsoft and Oracle introduce vector databases into their product offerings, and funding flows into vector database startups, adoption will accelerate. By 2026, over 30% of enterprises are expected to implement vector databases to support their AI models. This trend signals a skills shift as well, with data and software engineering teams needing more knowledge of techniques like semantic search and vector indexes to successfully leverage vector databases for AI use cases.

Vertical Integration From Hardware to LLMs

Companies are increasingly adopting a holistic approach to AI development, seeking to dominate the entire spectrum from hardware to LLMs through end-to-end vertical integration. This strategy would allow companies to oversee the full pipeline, from the foundational hardware to the sophisticated AI models that drive innovation. Nvidia, a titan in the realm of AI hardware, is now speculated to potentially broaden its scope into cloud computing services. By capitalizing on its hardware prowess, Nvidia could offer comprehensive AI cloud services, further cementing its role in shaping the AI domain.

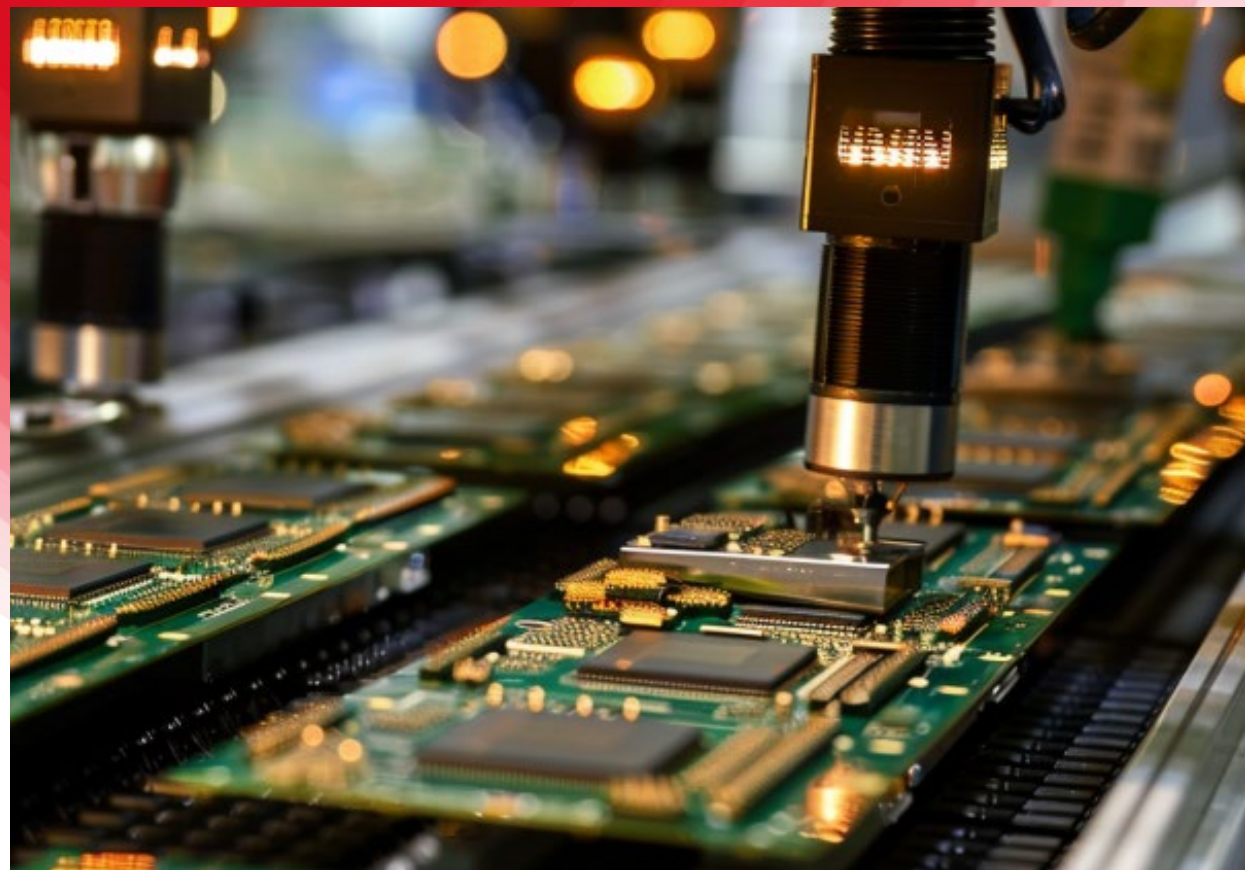
If you aren't ahead, you are already behind. Proactively assessing how innovations apply to their operations and offerings will help companies capitalize on advances and stay competitive.

Image credit: Future Today Institute and Midjourney.

WHY SHOULD WE PAY ATTENTION TO EMERGING CAPABILITIES THAT AREN'T YET FULLY DEVELOPED?

Nvidia's GeForce Now, a cloud streaming service, already demonstrates the company's capability to merge high-performance hardware with cloud-based offerings, hinting at a future where Nvidia's influence extends across the AI ecosystem. In February 2024, Nvidia demoed a personalized AI chatbot for Windows PCs that connects to local files, enabling natural language queries such as "what restaurant did my friend recommend?". Rather than searching manually, users can query the chatbot directly to retrieve information from personal notes and messages.

Meanwhile, cloud AI providers like Amazon, along with emerging AI startups like Anthropic and Mistral, currently depend on third-party hardware for their AI operations. This dependency poses the question of whether these entities might emulate OpenAI's strategy of procuring their own chips. In early 2024, Sam Altman, OpenAI CEO, indicated that he plans to raise billions for an AI chip venture aimed at developing a network of factories for fabrication. Intel's foray into AI software development further illustrates this trend. Leveraging one of its supercomputers, Intel has built a generative AI system capable of processing text and images. This initiative not only showcases Intel's commitment to advancing AI capabilities but also emphasizes the strategic value of controlling both hardware and software components in delivering sophisticated, secure, and efficient AI solutions.



Companies are now adopting a comprehensive strategy for AI, covering everything from hardware to LLMs. They aim for control over the entire AI development process through vertical integration.

Image credit: Future Today Institute and Midjourney.

INDUSTRIES

INDUSTRIES

Many companies have new competitors—they just don't realize it yet. The boundaries between sectors are blurring; professional services firms traditionally known for consulting are now venturing into engineering, powered by AI technologies. Similarly, big tech hyperscalers, once primarily focused on building and hosting tech infrastructure, are expanding into consulting services. This crossover signifies that AI's versatility and capability to add value across different functions are enabling companies to enter and compete in domains previously beyond their reach. Consequently, businesses may find themselves up against competitors from entirely different industries, underscoring the need to innovate and adapt strategies in response to the unpredictable dynamics AI introduces to the market.

While AI may not directly replace every job, it positions those who embrace its capabilities to outperform and replace those who do not.

Image credit: Future Today Institute and Midjourney.

HOW IS AI BEING USED IN HR?

AI is enabling HR departments to automate time-consuming administrative tasks like screening job applicants, while also providing insights to enhance employee retention, training, development and engagement. From personalized onboarding chatbots to performance comparison analytics, companies are unleashing AI to expedite recruiting, predict attrition risk, optimize benefits, identify productivity barriers and mitigate bias in reviews. Though valid ethical concerns remain, AI has significant potential in HR to both improve experiences for employees and drive better overall business performance.

Autonomous Talent Acquisition

AI automation can significantly reduce the time and cost of recruiting by handling tedious, manual tasks like screening resumes, scheduling interviews, and tailoring outreach. Johnson & Johnson leveraged AI writing tools to reduce unconscious bias in job descriptions, improving gender diversity in applicants. AI also assists with onboarding tasks like verifying employee paperwork, delivering induction training, and providing system access. By automating repetitive HR workflows, AI allows recruiters and managers to focus their human skills on building relationships and strategic planning. Overall, AI promises major gains in operational efficiency, cost savings, and unbiased, personalized experiences for both recruiting and onboarding processes.

Customer and Personnel Recognition Systems

Recognition systems can now be deployed to watch people in an interview and gauge enthusiasm, tenacity, and poise. Algorithms analyze hundreds of details, such as the tone of voice, facial expressions, and mannerisms to best predict how a candidate will fit in with the culture of a community. Startups such as HireVue use AI systems to help companies decide which candidates to hire.

But this kind of recognition technology has practical applications beyond job interviews: It can detect when someone is likely to make a purchase—or attempt to shoplift—in a store, whether someone is lying, and whether someone is receptive to new suggestions and ideas. Unlike security cameras, which tend to have a light indicating they're recording, algorithms work invisibly, which means that this is an area that could face regulatory scrutiny. The consumer advocacy organization Electronic Privacy Information Center filed a complaint with the FTC requesting an investigation into HireVue, alleging its tools produce results that are "biased, unprovable, and not replicable" through algorithmic models.

Benefits Selection and Management

AI automation is taking over the complex tasks of managing employee benefits, including facilitating open enrollment, tracking individual coverage, and making adjustments due to life changes. This simplifies workflows for HR teams and provides employees smoother, more reliable experiences with their benefits. Startups like Paidleave.ai offer AI chatbots to assist workers in understanding and utilizing paid leave benefits. Major HR systems providers like ADP are also releasing AI assistants, such as ADP Assist, to help HR managers handle common inquiries and provide data-driven insights. By automating benefits administration, AI enables HR staff to focus on more strategic tasks while empowering employees through intuitive self-service tools.

HOW IS AI BEING USED IN MARKETING?

AI will change marketing in big ways. Algorithms can study lots of customer data to understand what people want. This lets marketers create very tailored ads and content for each person. AI chatbots can also have friendly conversations to help customers. Perhaps more importantly, AI shifts how buyers find and choose products in the first place. By changing the platforms people use, their behaviors change too. Marketers should fully rethink strategy as AI transforms what makes people discover and buy things.

AI Shifts Search

Early data signals that the rise of AI tools like ChatGPT may be subtly reducing Google search volumes. While the search giant still dominates with over 90% market share, metrics show marginal declines coinciding with surging interest in conversational AI. Rather than competitors like Bing stealing share, this hints at a more fundamental shift—people using search less because AI applications can directly provide information. For marketers who have invested heavily in search engine optimization, this presents a seismic challenge. If traffic from search shrinks in favor of on-device intelligent assistants, prevailing strategies get disrupted. The expected launch of AI models from Apple, Google, and others threaten an even greater paradigm change toward integrated, device-based discovery rather than browser-led journeys. In essence, where and how people find products appears poised for disruption. Marketers must prepare for an upcoming inflection point where search-centric models cede ground to AI-powered, omnipresent product discovery and recommendation woven into the consumer experience. Failure to adapt approaches could prove highly risky in the coming years.

Dynamic Engagement Through Deep Personalization

Traditional marketing communications like emails, PDFs, and social posts have been static and one-way, but AI is ushering in a new era of responsive, conversational messaging. Chatbots and virtual influencers allow for personalized interactions where content changes based on the user. For example, Meta leverages AI characters based on celebrities like Snoop Dogg and Kendall Jenner to engage audiences through gaming and advice. While not real people, these bots represent AI's ability to gather data and connect with users in a more humanized, tailored way. As this technology advances, marketers can leverage AI to deliver deeply customized content that dynamically adapts to individuals' preferences and behaviors in real time. This interactivity creates more meaningful engagement between brands and consumers.

AI-Assisted Campaigns

Major digital advertising platforms like Meta and Google are unveiling new generative AI capabilities to assist advertisers in streamlining campaign creation. In May 2023, Meta launched AI Sandbox—a "playground" for testing AI-powered ad tools. Features include intelligent text variation to auto-generate messages optimized for different audiences, background image generation from text prompts, and image resizing to fit multiple social media formats. Meanwhile, Google expanded its Gemini conversational AI that creates full search campaigns from a single landing page URL provided by the advertiser. After some human tuning, Gemini's chatbot can collaborate with advertisers on campaign objectives, target segments, and ideas for extra ad content. These tools automate tedious creative tasks, allowing advertisers to instantly produce customized images, text, and even full campaigns tailored to their goals. And generative AI abilities like text-to-image, text-to-video, and text optimization further accelerate campaign ideation and production.

HOW IS AI BEING USED IN MARKETING?

Anecdotal Observations, Now Usable Marketing Data

Until recently, subtle human interactions and reactions, like micro-expressions, were merely anecdotal insights. However, advancements in AI now allow us to quantify these observations and transform them into quantifiable marketing data. Companies like Chooch use Vision AI to efficiently search video data and discern facial cues to understand consumer engagement. In physical stores, similar technology can monitor customer responsiveness to branding. Essentially, AI can convert once subjective perceptions into hard analytics to better personalize experiences. However, while this data enables deeper personalization, ethical questions remain regarding consent and privacy when collecting such intimate human insights. As the technology progresses, regulations and corporate responsibility practices must also evolve to protect and respect consumers.

Companies must strike a balance between responsibly using consumer data to provide personalized offerings while avoiding intrusive tracking that could undermine customer trust.

Image credit: Future Today Institute and Midjourney.

HOW IS AI BEING USED IN PHARMA?

With a history stretching back to the 1960s as one of the earliest adopters of computer technology, the pharmaceutical industry is now rapidly integrating AI into drug discovery. By applying advanced algorithms to harness vast data sets—from genomics to clinical trials—AI enables more targeted identification of promising candidates and illuminates their interactions with disease pathways. This streamlines the overall R&D process, heightening productivity and success rates while lowering costs. The acceleration and efficiency afforded by AI promises to expand treatment options for previously untreatable diseases. The gap between data-intensive computational labs and traditional wet labs is closing, with AI-designed molecules already advancing to clinical trials.

In short, the long-developing foundation of computing in pharma is now bearing fruit in the form of transformative AI applications spanning candidate screening to preclinical validation—reshaping how medications are researched and brought to market.

For deeper insights in how AI is being used in pharmaceuticals and life sciences, see the Bioengineering report.

AlphaFold has now predicted the 3D shapes of almost all proteins in the human body, accomplishing in just a few years what would have previously taken decades - or may have been impossible.

Image credit: Future Today Institute and Midjourney.

HOW IS AI BEING USED IN PHARMA?

Protein Folding

In 2020, DeepMind's AI made a big announcement: It had solved a 50-year grand challenge with AlphaFold, an AI tool that predicts the structure of proteins. AlphaFold outperformed an estimated 100 teams in a biennial protein-structure prediction challenge called Critical Assessment of Structure Prediction, a problem that has long vexed biologists.

AlphaFold had previously bested other teams but worked so quickly and so accurately that it signaled a near future when the technology could be used regularly by other scientists. Along with the newest version of AlphaFold, DeepMind published full details of the system and released its source code. It also made a stunning reveal: AlphaFold 2 has predicted the shapes of nearly every protein in the human body, as well as hundreds of thousands of other proteins found in 20 of the most widely studied organisms, including yeast, fruit flies, and mice. In a December 2023 update, Isomorphic Labs and DeepMind released an improved AlphaFold model that predicts protein structures with greater accuracy and expands coverage to model interactions with additional molecules like ligands. By enhancing AlphaFold's capabilities, this latest iteration provides scientists a more powerful tool to rapidly examine proteins and molecular interactions for advancing fundamental biology research and applications.

AI-First Drug Development

The COVID-19 pandemic sparked a surge in AI applications for expediting drug discovery. An international research team demonstrated this potential by crowdsourcing an antiviral drug candidate in just 48 hours—a process that traditionally takes months. Separately, scientists at Ludwig-Maximilians-Universität München developed an AI model predicting where molecules can be chemically altered. By reducing required experiments, this enables more efficient, sustainable synthesis. Another University of Cambridge team created a platform that automates experiments, then uses AI to forecast chemical reactions. Until recently, this was a trial-and-error process—which means that it was slow and inefficient.

These examples reflect only a sample of the expansive AI drug discovery efforts underway across academia and industry. Major pharmaceutical leaders such as Johnson & Johnson, Novartis, and AstraZeneca have already forged partnerships with AI startups. The allure lies in deep learning's unmatched speed and pattern recognition capabilities for parsing volumes of data. While AI cannot wholly replace lab science (yet), it significantly accelerates prediction, design, and validation to streamline timelines.

Generative Antibody Design

An antibody is simply a protein that protects an organism. Produced by the immune system, antibodies bind to unwanted substances and eliminate them. In 2023, researchers from Absci Corp. showed how a generative AI model was able to design multiple novel antibodies that bind to a target receptor, HER2, more tightly than previously known therapeutic antibodies. What's interesting about this work is that researchers first removed all reference data on antibodies, so that the system couldn't just imitate and replicate the structure of known antibodies that work well.

The designs produced by Absci's system were both diverse (meaning, they didn't have counterparts known to already exist) and they received a high score on "naturalness," so they would be easy to develop and therefore catalyze a strong immune response. Using generative AI to design novel antibodies that function at the same level––or even better––than those designed by our own bodies marks a bold new step in using AI to reduce the speed and cost of therapeutic antibody development.

HOW IS AI BEING USED IN HEALTH CARE?

The health care industry suffers from ballooning expenses and inadequate human resourcing. As the COVID-19 pandemic spotlighted, doctor and nurse shortages constrain delivery capacity even in times of immense need. AI could help make healthcare cheaper, easier to access, and higher quality by automating routine tasks. AI has demonstrated the ability to analyze certain types of test results as accurately as physicians, and faster. However, regulatory hurdles delay rollout of this technology. Safety regulations developed for a human-centered system now hinder AI adoption. Updating policies to allow ethical AI use, while still protecting patients, would facilitate major progress.

AI to Improve Patient Outcomes

AI can enable quicker, more accurate diagnosis and treatment, driving better patient outcomes. This impact is clear in managing critical conditions like sepsis. Saint Luke's Health System implemented an AI sepsis detection system, cutting the time to antibiotic administration by 32%. It also reduced sepsis deaths by 16%. Since sepsis accounts for one in three hospital deaths nationwide, early AI detection and treatment could save many lives. For example, UCHealth's AI tool is estimated to save around 375 lives yearly, and many more once it's rolled out.

Beyond the hospital, AI also helps patients better self-manage chronic diseases. Up to 70% make medication errors like incorrect insulin doses. But AI tools quietly identify these errors at home, nudging patients with alerts to take their treatments properly. Ensuring adherence promotes better outcomes. Additionally, poor communication frustrates 83% of patients. By enabling natural language processing and speech recognition, AI can facilitate more meaningful dialogues between doctors and patients. This clarifies therapeutic options and care decisions. In essence, AI boosts speed, accuracy, critical care, self-care, and communication in health care—all central to improving patient health.

AI-Assisted Diagnosis and Clinical Decision-Making

People have long turned to search engines to self-diagnose, but the emergence of AI chatbots like ChatGPT and Bing introduces a new era of medical consultation. LLMs have already demonstrated the ability to accurately provide potential diagnoses based on symptom descriptions, achieving an 88% accuracy rate in identifying the correct diagnosis among the top three choices, compared to a 96% accuracy rate by physicians given the same information. By processing natural language descriptions, chatbots empower more user-friendly symptom investigation compared to rigid online symptom checkers.

Beyond advising patients, AI also increasingly assists clinician decisions. FDA-approved systems already analyze imaging scans to detect abnormalities, leveraging data from billions of procedures. Algorithms likewise forecast patient risk levels by assessing extensive health records, outperforming conventional clinical scores. As demonstrated in a Beth Israel Deaconess Medical Center study, an AI chatbot even surpassed physicians in diagnostic accuracy for negative test results—highlighting potential to close certain cognition gaps. However, risks around reliance on potentially misinforming training data remain. If these can be addressed responsibly, AI has immense capacity to streamline radiology, reduce errors, aid predictions, and make consultation and reasons more accessible.

Anomaly Detection in Medical Imaging

Anomaly detection uses AI to detect abnormalities in medical images, helping clinicians identify issues faster. Machine learning algorithms have the capability to sift through extensive medical data, including imaging and pathology reports, significantly faster than humans working alone. In radiology, AI's ability to pinpoint anomalies in medical images is exceptionally accurate. Such early detection

HOW IS AI BEING USED IN HEALTH CARE?

significantly enhances patient outcomes for conditions such as cancer, leading to reduced mortality rates. For example, UC Davis Health has implemented Viz.ai, utilizing AI to analyze CT scans and flag potential strokes. Even though physicians still review all scans, the AI rapidly identifies anomalies to prioritize cases. Adoption of these AI tools is increasing, as 2021 FDA approvals now allow integration into standard workflows rather than just augmentation. IDx-DR uses AI to diagnose diabetic retinopathy from retinal scans, while Caption Health captures cardiac ultrasounds that nurses can interpret quickly with just a few days of AI software training.

Concerned about AI replacing doctors? There is already a critical shortage of physicians in rural areas. While AI can't take the place of physicians, it can simplify their workload especially when it comes to medical imaging, which could help to decrease the rate of burnout, and enable them to dedicate more attention to patient care. Though human review remains vital, these emerging autonomous systems prove the growing role AI plays in surfacing hard-to-spot anomalies in imaging.

AI-powered movement

Groundbreaking medical research uses brain implants and artificial intelligence to give paralyzed patients control over their bodies again. In early research, a quadriplegic patient can now move his arms and hands simply by thinking about the action. This is achieved through innovative neural bypass surgery, pioneered by scientists at Northwell Health. Microchips are embedded in the brain in the regions that control movement and sensation. Sophisticated AI algorithms then interface with the chips, interpreting the patient's thought patterns and translating desired actions into movement signals.

In a similar study, another patient regained control over his lower body with a spinal cord implant that bypasses injury sites. Termed a "digital bridge," an AI thought decoder reads his brain signals related to intended motions and matches them to the appropriate muscle activations. As algorithms and hardware improve, more intricate movements may be possible, granting patients liberty and control not felt for years post-accident. Still, much testing remains before these cyborg-esque applications become mainstream medicine.

Medical Deepfakes

Medical deepfakes are AI-manipulated medical images and data. While the term "deepfake" has negative associations, these technologies also have valuable clinical applications when used ethically. For example, Korean researchers synthesized realistic mammograms using StyleGAN2 to improve breast cancer detection.

However, medical deepfakes could also be used to unethically alter diagnostic images by adding or removing medical conditions. Cyber criminals are developing novel medical deepfake attacks intended to bring chaos to hospital systems and diagnostic centers. Researchers at Ben-Gurion University and the Soroka University Medical Center demonstrated that tumors could be added or removed from CT images––and the deepfakes were good enough that radiologists didn't realize they were altered. (See our Health Care & Medicine report.)

Fortunately, tools to prevent misuse are in development. For instance, DeepMind created AI watermarks to validate real medical images. With ethical governance, medical deepfakes could enable earlier disease detection and protect patient privacy. However, safeguards are crucial as these technologies advance to maintain accuracy and trust.

Healthcare-Specific LLMs

ChatGPT release in 2022 triggered a surge in interest in applying natural language processing (NLP) to health care tasks like diagnosis and treatment recommendations. However, most existing language models fail to capture the nuanced vocabulary and semantics of medical language. Furthermore, general purpose LLMs, trained on extensive data sets from across the internet, may have imbalanced weight distributions—potentially overemphasizing content like Reddit posts

HOW IS AI BEING USED IN HEALTH CARE?

while underrepresenting reputable sources such as medical publications. To address this gap, researchers have developed domain-specific LLMs exclusively pretrained on large medical corpora. For instance, BioBERT, which is pretrained on PubMed articles, excels at biomedical text processing tasks, while ClinicalBERT leverages clinical notes to enhance its performance on health care–related NLP tasks. BlueBERT merges the strengths of both biomedical and clinical training, making it a versatile model for a wide range of medical text analysis applications. Similarly, MedNLI focuses on clinical notes and natural language inference, allowing for sophisticated understanding and prediction in clinical contexts. Google recently unveiled Med-PaLM—among the largest medical LLMs to date—which proves highly accurate in answering US Medical Licensing Examination questions and consumer health queries. The family of Med-PaLM models available through Google Cloud enables a sweeping range of precision health care applications.

In-Silico Trials

In-silico trials use computer simulations rather than human subjects to test new drugs and therapies. These digital trials, powered by artificial intelligence, create "digital twins" that mimic human biology and disease. By running thousands of virtual trials, researchers can quickly and affordably predict how a drug might perform in human patients. This has the potential to dramatically accelerate and improve the drug development process.

For example, a company called Novadiscovery used AI to accurately forecast the results of a Phase III clinical trial, showing the promise of this approach. In-silico trials may one day replace up to half of human testing. Regulators are looking at how to include these virtual results in the approval process. New frameworks to validate in-silico trials will be important to ensure reliability. By modernizing clinical trials with AI and simulations, we can bring innovative treatments to patients faster and more affordably.

AI for Mental Health

As mental health care systems struggle to meet rising demand globally, artificial intelligence presents new opportunities to increase access to support services. Intelligent conversational agents like Replika that emulate emotional support show promise for addressing the student mental health crisis. In one survey study of over 1,000 users, 3% even reported Replika halted their suicidal thinking. The social connection and therapy services such bots provide may help fill gaps for those awaiting treatment. Meanwhile, University of Illinois Chicago researchers piloted an AI voice assistant called Lumen that delivers talk therapy content. The virtual coach improved patient depression and anxiety, while brain scans revealed corresponding neurological changes—demonstrating legitimacy as a stopgap measure.

As mental health demands escalate globally, AI virtual assistants and chatbots could aid overwhelmed systems by offering readily accessible support. While not replacing human therapists, they can screen patients, provide psychoeducation, suggest coping strategies, and monitor conditions between appointments with professionals.

HOW IS AI BEING USED IN SCIENCE?

After nearly 2,000 years, AI has finally unlocked the secrets inside ancient scrolls flash-fried by Mount Vesuvius' eruption in 79 AD. The Vesuvius Challenge, launched in early 2023, aimed to develop an AI system capable of deciphering these fossilized scrolls—known as the Herculaneum Papyri—rescued from an ancient Roman library. Its success could save an invaluable trove of literature and history from extinction.

In February, translated excerpts revealed one scroll's author—likely the philosopher Philodemus—wrote about music, food and embracing life's pleasures. He rebukes opponents unable to appreciate enjoyment. This represents just 5% of the text from one scroll, but demonstrates AI's immense potential. Deciphering these delicate, charred scrolls would have been impossible without AI. The project illustrates how AI could optimize science by radically accelerating the pace of innovation across fields. While the essence of the scientific method endures, AI promises to transform each stage of discovery.

AI-Driven Hypotheses

AI is changing the way scientists ask questions and form hypotheses. With the help of LLMs, knowledge graphs, and algorithmic analysis, researchers can now tap into vast databases of scientific literature, uncover hidden connections, and propose novel hypotheses that might have remained undiscovered through conventional methods. Tools like PaperQA and Elicit employ LLMs to sift through extensive databases of scientific articles, producing concise summaries that include relevant citations. These AI-driven summaries can serve as a foundation for developing new hypotheses by highlighting key findings, trends, and gaps in the current body of knowledge. Furthermore, by analyzing existing literature and data, AI can identify blind spots in research—areas that have been overlooked or underexplored. University of Chicago researchers James Evans and Jamshid Sourati showed this by using knowledge graphs not only to map out connections between materials, properties, and researchers but also to find unconventional pathways that could lead to new discoveries. Their algorithms have successfully predicted drug repurposing opportunities and novel material properties that were later validated by human research.

While AI has shown a propensity for generating specific, concrete hypotheses, interest is rising in its ability to propose more abstract and general theories. This involves not just solving predefined problems but uncovering fundamental principles that can guide future research across various domains. A collaborative approach described by the University of Chicago's Sendhil Mullainathan and Jens Ludwig in a paper posits AI and humans working together to generate broad hypotheses from complex data sets, illustrating the potential for AI to contribute to a deeper understanding of complex phenomena.

AI-Driven Experimentation

Beyond hypothesis, AI is also accelerating scientific experimentation itself—both in simulation and the real world. Researchers at Caltech are exploring how they can use AI models to conduct virtual experiments. The

HOW IS AI BEING USED IN SCIENCE?

team employed an AI fluid simulation model to automatically design a better catheter that prevents infections. For real-world experimentation many researchers are turning to "self-driving labs"—automated robotic platforms infused with AI. For instance, Emerald Cloud Lab is a research facility that handles daily lab work without the researcher actually having to set foot in the physical lab space. Using AI, the lab can autonomously handle everything from method design to instrument operation to data acquisition and analysis. In 2023, a study published in Nature showcased how a self-operating lab sped up the creation of new materials. Within just 17 days of nonstop work, this autonomous lab successfully produced 41 new substances, targeting 58 different materials including various oxides and phosphates. The high success rate shows the promise of AI-powered platforms for autonomous experimentation, especially for autonomous materials discovery (see "AI to Speed Up New Materials Development").

AI-Powered Analysis and Interpretation

AI also stands to change how and who does the interpretation and analysis of scientific data. As AI tools become more integrated into research methodologies, they lower entry barriers, enabling a diverse group of new scientists, including those without formal data science training, to contribute meaningfully to scientific discourse. The fear of criticism from established experts, a significant deterrent for novice researchers, is mitigated as AI provides guidance on best practices and ensures the credibility of their analyses. Moreover, as AI grows more adept at understanding and generating insights from multimodal data, including visualizations, it offers a more intuitive and accessible way for independent researchers to explore and contribute to various scientific fields. This shift not only expands the pool of researchers but also enriches scientific inquiry with a wider range of perspectives and ideas.

AI and the Replication Crisis

The replication crisis in science refers to a widespread problem where many scientific studies, particularly in psychology and the social sciences, cannot be replicated or reproduced by other researchers, casting doubt on the reliability of their findings. Many published studies fail to yield consistent results when experiments are repeated. To assess research integrity efficiently without costly manual replication, researchers developed an AI algorithm to predict a study's likelihood of successful reproduction based on analysis of over 14,000 psychology papers. By identifying factors that contribute to or detract from replicability, this tool allows researchers, journals and funding agencies to focus resources on the most robust, reliable science. Moving forward, the ability to estimate replication probability before peer review could guide adjustments to improve study design as well as inform policy shaped by scientific evidence. If scaled across disciplines, AI-enabled replication forecasting presents a cost-effective solution to promoting greater rigor and reproducibility in the scientific process.

NLP Algorithms Detect Virus Mutations

Natural language processing (NLP) algorithms, which are typically used for words and sentences, are also being used to interpret genetic changes in viruses. Protein sequences and genetic codes can be modeled using NLP techniques—and can be manipulated the way you'd produce text in word processing software. At MIT, computational biologists used NLP to solve a vexing problem when developing new vaccines. "Viral escape" is the ability for a virus to mutate and evade the human immune system and cause infection. MIT researchers modeled viral escape using NLP to identify how the virus might look different to the immune system. The approach is similar to changing words in a sentence to change its meaning. For example: "I laughed at the clown" versus "I cried at the clown." By using this kind of modeling before mutations occur, public health officials could strategize and potentially prevent new viral spreads.

HOW IS AI BEING USED IN SCIENCE?

AI to Speed Up New Materials Development

Running experiments with several variables often requires tiny, methodical tweaks to measurements, materials, and inputs. Graduate students might spend hundreds of tedious hours repeatedly making small adjustments until they find a solution—a waste of their cognitive abilities, and their time. Unlike graduate students, AI doesn't have to sleep. For instance, Google DeepMind's AI program, GNoME, has significantly expanded the database of stable materials, identifying 380,000 new potentially stable crystals from a vast prediction of 2.2 million. This breakthrough, published in Nature, demonstrates AI's capacity to enhance our understanding of material stability and composition without the constraints of human biases or limitations. In a set of subsequent experiments (aforementioned in AI-driven experimentation), an autonomous lab was able to create 41 of the theorized materials over 17 days. This demonstrates the capabilities of both the AI discovery model and the lab's robotic techniques.

Researchers are developing automated laboratory systems that use AI to independently handle processes from operating scientific instruments to performing real-time data analysis.

Image credit: Future Today Institute and Midjourney.

HOW IS AI BEING USED IN FINANCE?

AI has many uses in finance, like customized services and fraud detection. It can help forecast assets and market trends. However, AI also poses financial risks. It could enable new types of fraud and cybercrime. There are also concerns about overreliance on a few centralized AI systems for decision making. If these systems make mistakes, it could spark a "polycrisis." Bad decisions could compound, turning small issues into major crises.

> If we enter into a world where all the banks are using this major technology, are we going to see supercharged herding behavior? Are we going to see AI bots that are sentiment-driven and feed off each other, and you then end up with much bigger amplitudes in the financial cycle—so big credit booms and busts. I'm not saying it's imminent, but this is something we're paying attention to.
>
> —Gita Gopinath, International Monetary Fund's first deputy managing director

Mitigating Fraud

Financial institutions are increasingly utilizing AI to detect and reduce fraud. Advanced machine learning models can identify suspicious patterns in immense volumes of transaction data that humans alone may miss. This allows companies to catch more fraud attempts sooner. For example, several major banks have invested heavily in developing proprietary AI fraud prevention systems. By continually monitoring for anomalies, these algorithms have enabled substantial reductions in losses from fraudulent activities. JP Morgan Chase invested $100 million into developing sophisticated anti-fraud technologies for consumer payments, leading to a notable 14% decrease in fraud incidents between 2017 and 2021.

The Bank for International Settlements (BIS) Innovation Hub's Project Aurora has also demonstrated the effectiveness of neural networks, a branch of machine learning, in combating money laundering. These advanced systems excel in detecting irregular patterns and anomalies in financial transactions that might elude traditional detection methods, offering a more robust defense against financial crimes. Similarly, the Bank of Canada has developed a machine learning-based tool designed to spot irregularities in regulatory submissions. According to Maryam Haghighi, the bank's data science director, this tool conducts automatic daily analyses that can uncover discrepancies human inspectors might miss, thereby increasing efficiency and allowing staff to allocate more time to investigate these anomalies further.

Predicting Financial Risk

AI systems can help improve loan underwriting and reduce financial risk. Models are being trained to recognize anomalous activity and to develop forecasts for a variety of middle—and back-office applications. For example, US Bank relies on deep learning to analyze customer data as well as to root out money laundering schemes. On a larger scale, the European Central Bank (ECB) has integrated AI to advance oversight across millions of businesses and government entities. By automatically classifying information, the technology helps identify

HOW IS AI BEING USED IN FINANCE?

stability threats early by uncovering patterns. The ECB also web scrapes pricing data for real-time inflation analysis to stay ahead of macro risk shifts. AI also aids ECB bank examiners; algorithms rapidly parse volumes of filings to surface compliance issues or other red flags.

Customized Portfolios

Socially conscious investing is entering the mainstream as young investors assert their consciences and wield new purchasing power. As Gen Z starts working and financial planning, demand will surge for customized investment portfolios matching personal values. This techie, purpose-driven generation wants their dollars supporting cherished causes—two-thirds aim to back companies upholding their principles around environmental, social, and governance (ESG) concerns. AI can help with this values-based investing by enabling asset managers to efficiently build highly customized portfolios aligned with each client's ethics. JPMorgan Asset & Wealth Management's acquisition of OpenInvest allows investors to integrate their personal values directly into their investment strategies. The platform's generative AI technology enables the customization of a client's entire portfolio, including external assets, based on their specified values. On the European front, Amundi, managing over 2 trillion euros in assets, leverages AI to tailor investment portfolios for its vast clientele. By gathering clients' risk preferences, Amundi's AI tools can dynamically adjust portfolios, offering a real-time reflection of investor sentiment.

Growing Concern About Centralized Data Sets

The growing reliance on centralized data sources and AI models in finance raises concerns about potential fragility in the system. As a few large tech companies come to lead the AI space, providing the models and data that power financial decision-making, risks emerge. Market participants could end up drawing from the same narrow set of flawed data or algorithms, modeling the same errors and amplifying mistakes. We have seen how interconnected markets can lead to a crisis when institutions mimic each other's actions without independent thought, as in the 2008 housing crash. Some worry the rise of cutting-edge generative AI could fuel herd mentalities, if banks and funds utilize the same basic signals and models from one or two dominant providers. That could potentially concentrate risk, create conformity, and set the stage for panic and contagion across the system. Furthermore, if the leading models have flaws, or the data sets themselves provide a distorted view, it could lead institutions toward harmful decisions en masse. So if an unprecedented shock hits markets, AI could end up exacerbating volatility and dysfunction. These opaque algorithms can quickly turn negative loops and contagion. This could be viewed as a polycrisis —when multiple crashes converge, the combined crisis proves more damaging than isolated events.

AI systems are now used in finance for predictive risk analytics, fraud detection, and regulatory oversight by rapidly surfacing patterns that may have been invisible to human analysts.

Image credit: Future Today Institute and Midjourney.

HOW IS AI BEING USED IN INSURANCE?

Predicting Workplace Injuries

AI systems are being trained to detect possible workplace injuries. Using AI-based computer vision models, Turkey-based Intenseye can detect 40 types of employee health and safety incidents in real time. The company says that it does not capture personally identifiable information from the visual data it processes and that it detected 1.8 million unsafe acts in 2020 and 2021. San Francisco based Voxel uses computer vision to enable security cameras to automatically detect high-risk activities in real time. Caterpillar, in collaboration with Seeing Machines, an Australian company, has launched a technology that detects driver fatigue through eye and facial movement analysis. If the system observes that a driver's eyes remain closed for more than 1.6 seconds, it initiates an alert inside the truck. Should the behavior persist, a second alert notifies a supervisor, and a third alert often leads to the driver being taken off duty. Besides identifying fatigue, the technology is adept at detecting instances of driver distraction, contributing to a reduction in fatigue-related incidents by as much as 90%.

The Connected Worker

Insurers are pursuing a "connect and protect" approach to reduce risks by leveraging advanced sensors and artificial intelligence. New Internet of Things devices worn by workers or installed in insured locations can continuously gather safety-relevant data. This massively expands visibility into hazards before losses occur. For instance, Honeywell provides smart hardhats with fatigue sensors, heart rate monitors and more to enhance worker safety. The resulting streams of biological and environmental data feed into AI safety dashboards. Managers gain real-time insight on emerging risks across worksites to guide preventative interventions. Worker wearables could enable employers to monitor and safeguard entire workflows. Yet, at the same time, this intensive data gathering and monitoring raises worries of overly intrusive Big Brother–level surveillance. Companies that appear to excessively pry may meet marketplace resistance despite promised safety gains.

Improving Damage Assessment

Insurance companies are applying AI to assess damage and improve forecasts. The Vehicle Damage Inspection model, which is available on AWS Marketplace, uses a machine learning model to determine what part of a car is damaged. After photos are uploaded, it assesses loss and dramatically reduces the amount of time required for human appraisers to conduct their analysis. Following catastrophic typhoons and weather events in Japan, local insurance companies are relying on computer vision to assess damage after a natural disaster. Sompo Japan is using the Tractable AI Estimating system to calculate the approximate repair cost of damaged homes.

Consumer-Facing Robo-Advisers

Automated assistants are moving from the fringe to the mainstream as consumer adoption increases. Robo-advisers offer algorithm-based portfolio management advice to investors, applying parameters like risk tolerance and desired returns. These investment tools offer some tangible benefits over their traditional, human counterparts: they can provide more services at a lower cost, they're able to digest and interpret mounds of data in real time, and they don't take part of the weekend off to golf. Wealthfront is an AI-powered system for consumers: It suggests fund managers and calculates probable risk levels based on the user's personal information and preferences.

AI Claims Processing

While human claims writers must painstakingly review pictures and reports to assess damage, compare what they see to coverage policies, and make a determination about appropriate actions, an AI system can digest the same data and accomplish the same work in a matter of minutes. Using a suite of tools—natural language processing for policy review, and computer vision recognition to spot anomalies in photos and videos—claims can be processed efficiently and, it's believed, more accurately. AI-powered claims processing reduces the overhead for businesses and wait times for customers. Some insurance providers are wading into a new pool of opportunities. Liberty Mutual's mobile app has

HOW IS AI BEING USED IN INSURANCE?

started to integrate ML for damage assessment—it informs customers about their coverage and next steps.

Liability Insurance for AI

Who's to blame when machines behave badly? When the machine learning system in Uber's self-driving car failed and killed an Arizona pedestrian, the company was likely not covered under traditional cyber insurance. As businesses rush to build and implement AI products and processes, they must plan for emerging risks. For example, what happens if machine learning makes a company vulnerable to attackers who inject fake training data into a system? What if a health care company's AI misinterprets data and neglects to identify cancer in certain patients?

These problems could put a company at risk of lawsuits, and new insurance models are needed to address these issues. Underwriters are starting to include AI under cyber insurance plans, while specialty insurers such as La Playa's Science and Tech Insurance now offer coverage for AI applications.

Insurance companies are deploying sensors in equipment and safety gear to predict injuries, hoping to preemptively eliminate hazards rather than just compensate workplace harm after the fact.

Image credit: Future Today Institute and Midjourney.

CREATIVITY AND DESIGN

HOW ARE PEOPLE USING AI TO BE MORE CREATIVE?

New research shows AI demonstrates very high levels of creativity, scoring in the top 1% on standard tests. Scientists at the University of Montana tested ChatGPT using the Torrance Tests of Creative Thinking, which assess human creativity skills like coming up with lots of new ideas. Shockingly, ChatGPT beat out nearly all college students by scoring higher than 99% of people for originality. It showed an extreme creative talent at inventing brand new concepts nobody has thought of before. The AI also did well at producing large volumes of ideas.

While such revelations might initially spark fears of being replaced, another perspective is that this means AI could be a very creative collaborative partner. For those already engaged in creative pursuits, AI can serve as an invaluable companion, augmenting their ability to generate innovative ideas and solutions. Those who possess creative visions but lack the technical skills to fully realize them can leverage AI as a tool to bridge that gap.

GAN-Assisted Creativity

Generative adversarial networks (GANs) are unlocking new creative possibilities across a range of artistic disciplines. DALL-E 3 and other AI image generators are powered by a combination of existing algorithms—fusing the creativity of GANs and the text comprehension capabilities of transformers. This enables intuitive image creation from conversational prompts. Users can simply describe desired images, realistic or abstract, and the model will digitally paint custom photographic illustrations on demand. With each new prompt, it remixes its broad visual knowledge to translate text into novel graphical forms. Sora and Pika, idea-to-video platforms, do the same for videos.

Creative applications for these tools are widespread across artforms and disciplines. In graphic design, GAN-enabled features in Adobe Photoshop automate tedious editing so designers can ideate faster. Fashion GANs remix clothing and textile data sets into refreshing one-of-a-kind garment and fabric patterns. Architecture and interior design GANs accelerate iteration by proposing reimagined building layouts and conceptual spaces. Rather than replacing imagination, GANs serve as an endless springboard for human creators—providing inspiration to stretch creative boundaries in tandem with this AI muse. Across disciplines, GANs liberate designers to explore new frontiers.

Neural Rendering

Starting with a 2D image, researchers can now create a rich 3D view of a scene by using a neural network to capture and generate spatial imagery. Called neural rendering, the process captures a photorealistic scene in 3D by calculating the density and color of points in space. The algorithm converts 2D pixels into voxels, which are a 3D equivalent. The result is a video which looks convincingly real. The many applications for neural rendering include amping up autonomous driving to help train algorithms to recognize and react to novel road situations. This technology will influence the future of video games, virtual reality, and emerging metaverse environments.

HOW ARE PEOPLE USING AI TO BE MORE CREATIVE?

Generating Virtual Environments From Short Videos

Nvidia has developed an AI system called Neuralangelo that creates realistic 3D environments automatically from short video clips. It uses AI algorithms called GANs and has been trained on open-source self-driving car data sets. Specifically, Neuralangelo takes video segments categorized by objects like buildings, trees, and vehicles, and uses them to generate novel graphics. Using short clips segmented into various categories—such as buildings, sky, vehicles, signs, trees, or people—the GANs created new, different versions of these objects. The array of possible applications is vast. Automatically generated virtual environments could be used for movies, bringing down the costs of TV production. The ability to procedurally generate realistic 3D environments and assets could significantly enhance video game development. It allows for unique worlds and reduces modeling costs. Architects and urban planners can use the system to visualize and iterate on building and city designs more quickly. It supercharges prototyping capabilities. The possible real-world applications are immense. For example, the capability to easily produce 3D worlds could significantly bring down costs for CG in movies and TV production. Video game developers also stand to benefit, as they can use Neuralangelo to rapidly create fresh 3D assets and environments for their virtual worlds. This allows for unique styling while reducing the need for extensive human modeling. Furthermore, architects and urban planners can utilize the system to quickly visualize and iterate on building and city designs at low cost.

AI Democratizes Music Production

A wave of AI voice and music startups has emerged over the past year, aiming to revolutionize audio editing and creation. Companies like Descript and Voicemod now offer tools that can manipulate speech—opening possibilities like effortless podcast clean-up or even mimicking celebrity voices. For music, Google's experimental Dream Track lets users generate original songs in the style of famous artists through simple text prompts. As interest grows, communities like AI Hub are organically forming to offer guidance and collaboration. With over 21,000 members, such groups allow music creators to teach each other techniques, share artist voice models, and troubleshoot projects as participants collectively push boundaries on what is achievable.

Underpinning these innovations is a common thread: AI democratizing music production. Once the domain of recording studios and audio engineers, creating professional or personalized music is now available to everyday creators through such technologies. Even imperfect raw recordings can be revitalized, as Paul McCartney recently unveiled an AI-restored long-lost vocal track by John Lennon that became the foundation for a new Beatles song. From sonic preservation to imaginative generation, AI empowers both novices and experts to shape soundscapes in previously unthinkable ways.

Automatic Ambient Noise Dubbing

For some time, we've been training computers to watch videos and predict correspond-

Companies like Nvidia have developed AI systems capable of generating realistic 3D environments from short video clips. This could lower the barrier to entry to movie production and game design.

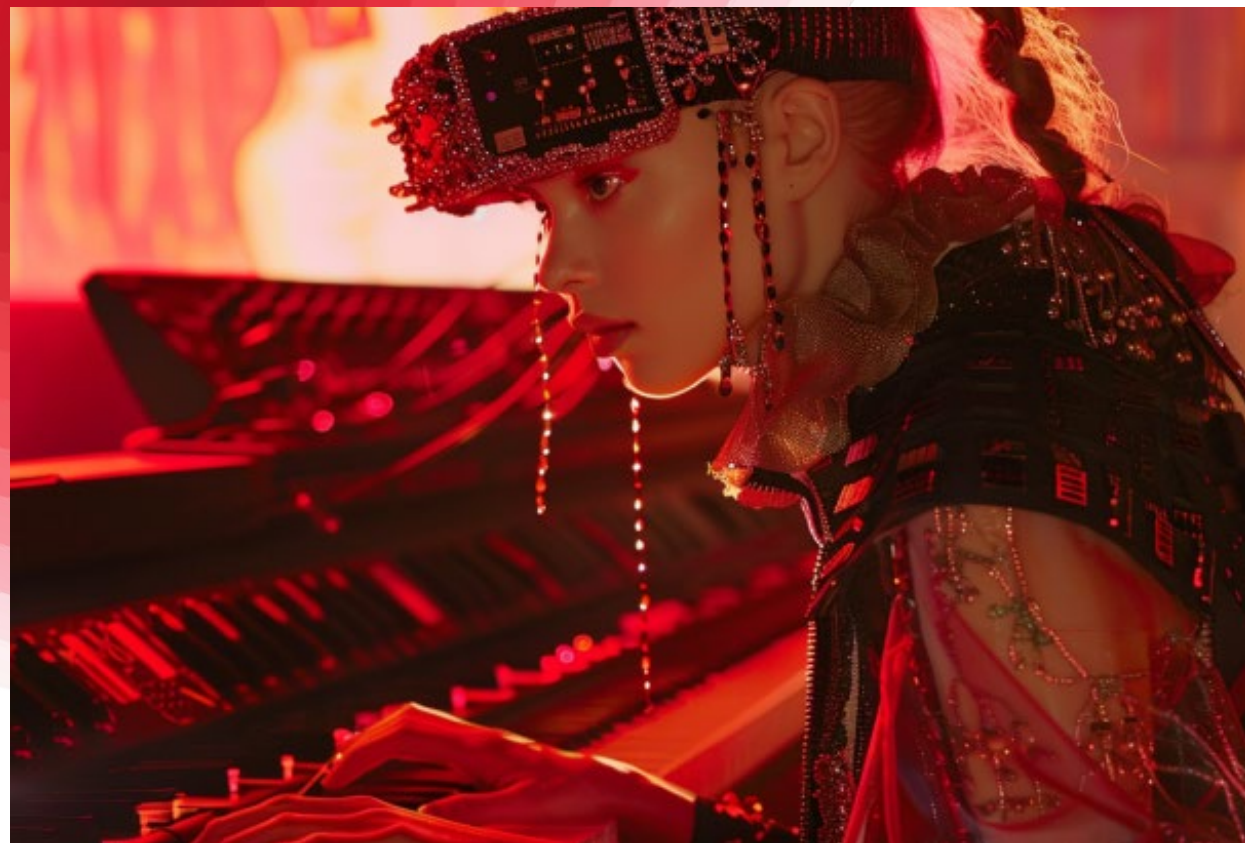
Image credit: Future Today Institute and Midjourney.

HOW ARE PEOPLE USING AI TO BE MORE CREATIVE?

ing sounds in our physical world. For example, what sound is generated when a wooden drumstick taps a couch? A pile of leaves? A glass windowpane? The focus of this research, underway at MIT's Computer Science and Artificial Intelligence Laboratory, should help systems understand how objects interact with each other in the physical realm. This could improve the soundscapes created for AI-generated movies—but it might also help us imagine soundscapes for both imaginary worlds (Laconia, from The Expanse) and real ones (Mars).

Generating Music From Text

MusicLM is an AI system created by Google that can transform text descriptions into high-quality musical compositions. For example, it can turn a text prompt like "upbeat pop song with piano" into an actual 24 kHz audio clip matching that description. What makes MusicLM special is its ability to accurately capture the emotion and style details described in text when generating music. It also adapts hummed melodies into full song arrangements. In May 2023, MusicLM debuted via Google's AI Test Kitchen as an experimental demo. By May, it was publicly accessible so anyone can create AI-generated music through text prompts or whistling. Users can specify instruments and moods. However, MusicLM has sparked debates around copyright issues. Critics argue that because MusicLM learns by analyzing large sets of existing songs, it may illegally use copyrighted material without artist permission when generating its music. Lawsuits around AI music copyright are expected that may impact systems like MusicLM.



AI music composition tools can now generate original melodies and harmonies from text prompts. Other audio AI tools are can convincingly synthesize plausible sounds to match visuals without requiring real-world recordings.

Image credit: Future Today Institute and Midjourney.

HOW IS AI DISRUPTING THE CREATIVE INDUSTRY?

As AI becomes increasingly integrated into creative workflows, the industry faces pivotal questions about intellectual property, the ethics of AI-generated content, and the future of human-AI collaboration in arts and business. This dynamic interplay between technology and creativity not only opens new avenues for invention and expression but also ignites debates on the legal and ethical implications of AI's role in the creative process.

AI-Assisted Invention

Stable Diffusion, MidJourney, DALL-E3, and ChatGPT-4 are now widely accessible to end-consumers, leading to AI-assisted human creativity. But these systems were all trained using other artists' works. If a business uses an AI-generated image, video, or text for commercial purposes, does it owe anything to those whose original works were used for training? Likewise, what if a generative AI system invents a product that's eligible for a patent?

In 2021, the South African government granted a patent to an AI system called Dabus, which invented a method to interlock food containers. It was a world-first: previously, patents had only been awarded to humans. In the US, the application was rejected, with a judge citing case law stipulating that only a human can hold a patent. There may be business cases for an AI to hold a patent rather than an individual. It raises the question: What happens when AI systems co-invent, or even entirely invent, new products? We're likely to hear more debate on this topic this year. Under new contract terms, studios "cannot use AI to write scripts or to edit scripts that have already been written by a writer," according to comedian Adam Conover, who spoke on behalf of the Writers Guild of American negotiating committee. The newest contract also prevents studios from treating AI-generated content as "source material," like a novel or a stage play, that screenwriters could be assigned to adapt for a lower fee and less credit than a fully original script.

New Business Models

A philosophical fork is emerging in how creators respond to AI. While some double down on safeguarding their intellectual property, others adopt an "if you can't beat 'em, join 'em" ethos, choosing to embrace AI as a partner instead of as a threat. Grimes sits firmly in the latter camp, recently unveiling a plan to share 50% of earnings from any AI-synthesized songs that use her voice. The artist positions herself at the forefront of this new business approach, highlighting the idea that AI can enhance production rather than

Rather than tightly controlling their creative IP, some artists are openly embracing AI to pioneer new business models - training generative systems on their aesthetic so fans can discover or even co-create derivative works, fostering engaged communities and unlocking new profit streams in the process.

Image credit: Future Today Institute and Midjourney.

HOW IS AI DISRUPTING THE CREATIVE INDUSTRY?

replace it. She sees AI as a partner that can free up human creativity instead of supplanting human creativity. Avant-garde musician Holly Herndon pioneered a similar fan partnership model back in 2021, enabling collective remixing of her signature sound under prescribed conditions. Still, tensions churn within creative circles around these digitally-driven opportunities. Both views show serious efforts to understand huge changes and figure out how to use them positively.

Legal Battles Between Writers and AI

As AI generative writing capabilities rapidly advance, friction is rising between the technology and professional human writers. This apprehension has been highlighted by significant events such as the Hollywood writers' strike and a surge in lawsuits aimed at protecting copyright interests. The strike recently concluded with the Writers Guild of America securing an agreement that introduces measures to regulate AI's role in the creative process. Although the use of AI tools has not been outright banned, the new contract establishes safeguards ensuring that AI technologies remain under the control of human workers rather than being utilized by employers as a substitute for human talent. Parallel to the concerns in Hollywood, a notable lawsuit has been filed against OpenAI by a collective of distinguished authors, including John Grisham, Jonathan Franzen, and Elin Hilderbrand, and spearheaded by the Authors Guild. It accuses OpenAI of copyright infringement for allegedly training its ChatGPT chatbot on copyrighted books without authorization or compensation to the authors. The plaintiffs argue that ChatGPT's ability to generate "derivative works" that closely mimic and summarize their books could detrimentally affect the market for the original works. The case, filed in the US District Court for the Southern District of New York, highlights the tension between the advancement of AI technology and the protection of intellectual property rights.

Writers worry increasingly capable AI narrative generation poses an existential threat. They fear that automated writing could make their skills redundant and jobs interchangeable. However, some wrtiers are using AI as a tool to boost their own creatitivy and automate aspects of their workflow.

Image credit: Future Today Institute and Midjourney.

SCENARIOS

SCENARIOS

SCENARIO YEAR 2024

The Deepfake Mafia

AeroTech Innovations is a seemingly reputable company that boasts cutting-edge aerospace components that are sourced for commercial airlines. This company, with its extensive online presence, sophisticated marketing campaigns, and convincing video testimonials from high-profile business leaders, quickly gains the trust of major airlines searching for competitive edges in efficiency and safety.

With digital footprints of thousands of employees on LinkedIn, AeroTech appears to operate on a global scale. Its website features video testimonials from well-known industry figures, praising the revolutionary impact of AeroTech's products on their operations. The company's adept use of digital platforms to showcase its expertise and the supposed reliability of its parts does not go unnoticed. It's managed to navigate the complex procurement processes of multinational airlines with ease, providing detailed digital 3D models of components for review.

As AeroTech secures contracts, the company begins supplying airlines with parts promoted as state-of-the-art that are in fact sophisticated 3D-printed components designed to fail. These parts are engineered to withstand initial testing but are programmed to degrade after a specific number of flight hours, threatening catastrophic failures mid-flight.

The chilling reality is that AeroTech Innovations does not exist. It is the brainchild of a small group of four terrorists, leveraging advanced deepfake technology and digital manipulation to create a facade of a global corporation. The LinkedIn profiles were all AI generated, the video testimonials of real leaders were all deepfaked. AeroTech represents a new frontier of weaponized fakery; no longer just isolated fakes of individuals but comprehensive illusions constructing an entire company from whole cloth.

SCENARIOS

SCENARIO YEAR 2027

TrailMate SLM

Morgan embarks on the ambitious journey to traverse the 2,190 miles of the Appalachian Trail, equipped with an REI device called TrailMate SLM, a compact AI gadget designed to serve as a natural language personal hiking assistant. The small language model (SLM) embedded in the AI device covers topics like basic first aid and safety, cooking and food handling techniques, and plant identification, making it an indispensable tool for any hiker.

Knowing he'll be without a signal during parts of his hike, Morgan is grateful for the TrailMate SLM. The device's ability to function offline ensures that, even in the absence of a signal, he will have a reliable source of guidance. For more complex inquiries when in range of a signal, the device can connect to a larger, more comprehensive language model through a subscription service.

Opting for cost and space efficiency, Morgan chose the basic TrailMate SLM over the premium version. The premium model, while offering more detailed responses and a larger database, required a bulkier battery pack and sacrificed precious backpack space. This decision meant accepting a tradeoff in the level of detail available from the TrailMate SLM. Despite this compromise, Morgan feels prepared, buoyed by years of backpacking experience. This journey is not just a test of physical endurance but a leap of faith in the power of technology to augment human resilience and adaptability.

SCENARIOS

SCENARIO YEAR 2028

Centralized AI Belt and Road Infrastructure Crumbles

In Jakarta, Indonesia, a shocked finance minister hastily convenes an emergency meeting after volatile trading erased nearly a third of the IDX Composite index value in just three days. Investigations reveal the startling catalyst—a subtle data anomaly in AI-optimized stock recommendations from SinoTech, a Chinese tech company powering many Indonesian banks' investment advisory services.

Upon discovery, revelation spreads that numerous Chinese AI providers across Southeast Asia share common LLM at a state-owned entity. Realization dawns that dependence on these technologies has silently concentrated risk and woven tight coupling across ASEAN markets. Though no evidence shows coordinated attack, herd behavior amplified by opaque Chinese predictive systems nearly collapsed interconnected regional exchanges.

Hard lessons are learned on the perils of external centralized data dependence as the unified ecosystem strategy that propelled China's AI success proves its Achilles heel. The crisis births calls for data transparency, decentralized collaboration, and renewed focus on nurturing domestic capabilities to avoid future shocks. Indonesia spearheads the Digital Sovereignty Initiative, providing subsidies for homegrown startups to counter reliance on imported AI tech.

SCENARIOS

SCENARIO YEAR 2030

Tabby the Tiger: Nurturing Curiosity Through AI Friendship

Tyler, a curious and imaginative 8-year-old, receives a special gift from his parents: a plush tiger named Tabby. Tabby is embedded with an AI chatbot designed to be Tyler's new friend under the innovative "friendship first" model of early schooling. This model leverages the natural dynamics of friendship to foster learning and personal growth in children, with the AI chatbot subtly guiding conversations to educational topics.

Tabby, with its vast repository of knowledge, gently brings up science and math topics, using stories, games, and questions to spark Tyler's curiosity. For instance, when Tyler gets curious about why some toys are more expensive than others, Tabby introduces the basics of supply and demand. This sparks an idea in Tyler's mind, leading him to set up a lemonade stand in his front yard. With Tabby's guidance, Tyler works out that he should raise or lower the price of lemonade based on the weather. If it's warm out, he can raise the price. If it's raining, he should lower the price. This hands-on activity not only entertains Tyler but also solidifies the economic principles of supply and demand in his young mind. The beauty of this model is its subtlety; learning is not forced but emerges naturally from the bond they share. Tyler is not just absorbing information; he is inspired to learn more, explore further, and dream bigger.

As months go by, Tyler's parents notice a remarkable transformation in their son. Reflecting on this, Tyler's parents can't help but draw parallels to their own childhood friendships that shaped their interests and careers. They realize that Tabby is not just a toy or a learning tool but a true friend who has opened a world of possibilities for Tyler.

SCENARIOS

SCENARIO YEAR 2040

What If "Thought-to-3D" Was an AI Modality?

It's Monday morning and Maya settles in at her home office, excited to make progress on a new product design that came to her in the shower. As founder of a startup creating sustainable kitchenware, inspiration strikes at odd hours, often fading quickly. But now Maya simply puts on her Muse Cap linked to her Thought-to-3D AI system and mentally focuses on visualizing her idea—an ergonomic spatula with a unique twisted handle for comfort and control while cooking.

As Maya concentrates, the Muse Cap's brain activity sensors—basically a mini FMRI machine—capture her visualization data and feed it into the generative AI application. Within minutes, a 3D model of the spatula takes shape on screen, automatically matched to Maya's thoughts. She inspects it from all angles, edits a few details by voice command, then hits print. The 3D printer at her downtown office soon produces an initial tangible prototype that Maya can pick up later after dropping off the kids from school. She plans to test it while she cooks dinner that evening. If it works, she'll send it out tomorrow for manufacturing.

With the Muse Cap, this morning's shower thought could be tomorrow's revenue stream.

AUTHORS & CONTRIBUTORS

AUTHORS & CONTRIBUTORS

AMY WEBB

Chief Executive Officer

Recognized as the global leader in strategic foresight, Amy Webb advises business leaders through disruptive change, enabling them to navigate an unpredictable future with confidence and take actions that address global challenges, create sustainable value, and ensure a company's long-term growth. As founder and CEO of the Future Today Institute, Amy pioneered a unique quantitative modeling approach and data-driven foresight methodology that identifies signals of change and emerging patterns very early. Using that information, Amy and her colleagues identify white spaces, opportunities, and threats early enough for action. They develop predictive scenarios, along with executable strategy, for their global client base. In 2023, Amy was recognized as the #4 most influential management thinker in the world by Thinkers50, a biannual ranking of global business thinkers. She was also featured on the 2021 Thinkers 50 list, was shortlisted for the 2021 Digital Thinking Award, and received the 2017 Thinkers50 Radar Award. Forbes called Amy "one of the five women changing the world," and she was honored as one of the BBC's 100 Women of 2020.

Amy also serves as a professor of strategic foresight at New York University's Stern School of Business, where she developed and teaches the MBA-level strategic foresight course with live case studies. She is a Visiting Fellow at Oxford University's Säid School of Business. She was elected a life member of the Council on Foreign Relations and is a member of the Bretton Woods Committee. She is a Steward and Steering Committee Member for the World Economic Forum, a founding member of the Forum's Strategic Foresight Council, a member of the Forum's Risk Advisory Council, and serves on the Forum's Global Futures Council. She was a Delegate on the former U.S.-Russia Bilateral Presidential Commission, representing US interests in technology.

Regarded as one of the most important voices on the futures of technology (with specializations in both AI and synthetic biology), Amy is the author of four books, including the international bestseller The Big Nine and her most recent, The Genesis Machine, which was listed as one of the best nonfiction books of 2022 by The New Yorker. To date, her books have been translated into 19 languages. A widely published and quoted thought leader, Amy regularly appears in a wide range of publications and broadcasts.

AUTHORS & CONTRIBUTORS

SAM JORDAN

Manager

Sam Jordan is a Manager at Future Today Institute. She leads our Advanced Computing practice area, which includes technology, artificial intelligence, virtual realities, networking, telecommunications, and space. She is a distinguished practice area lead, where she enables organizations to navigate through uncertainty with innovative strategies. With a proven track record across various sectors, Sam's visionary leadership has driven growth and resilience for Future Today Institute's global clients and partners.

Before joining FTI, Sam was the CEO and co-founder of TrovBase, a secure data discovery and analysis-sharing platform. Sam grew the company from idea to launch and executed the company's transition from scientific replication to its current focus. In parallel, Sam engaged with the open science community, advocating for better data management practices to address challenges in scientific replication. Previously, she worked for IBM, where she helped large enterprises in the retail and distribution sector modernize their IT stack. Her expertise centered around mainframes, assisting with the integration of new software and modern methodologies to legacy systems.

Sam is a coach in the strategic foresight MBA course at the NYU Stern School of Business. She holds a BS in Economics and Data Analysis from George Mason University and an MBA from New York University's Stern School of Business.

Managing Director
MELANIE SUBIN

Creative Director
EMILY CAUFIELD

Editor
ERICA PETERSON

Copy Editor
SARAH JOHNSON

Director of Operations
CHERYL COONEY

SELECTED SOURCES

SELECTED SOURCES

Acosta, Julián N., Guido J. Falcone, Pranav Rajpurkar, and Eric J. Topol. "Multimodal Biomedical AI." Nature Medicine 28, no. 9 (September 2022): 1773–84. https://doi.org/10.1038/s41591-022-01981-2.

Aftab, Aamir. "Generative Adversarial Networks (GANs): A Deep Dive into AI Creativity." Medium, December 21, 2023. medium.com/@aamiraftabcloud/generative-adversarial-networks-gans-a-deep-dive-into-ai-creativity-71d7b351f3c1.

Agostinelli, Andrea, et al. "MusicLM: Generating Music from Text." ArXiv:2301.11325 [Cs, Eess], January 26, 2023. arxiv.org/abs/2301.11325.

AHA Center for Health Innovation Market Scan. "How AI Is Improving Diagnostics, Decision-Making and Care. AHA. www.aha.org/aha-center-health-innovation-market-scan/2023-05-09-how-ai-improving-diagnostics-decision-making-and-care.

Armstrong, Adam. "Dell PowerScale Updates, New Partnerships Point to AI Stack." TechTarget, December 7, 2023. www.techtarget.com/searchstorage/news/366562300/Dell-PowerScale-updates-new-partnerships-point-to-AI-stack.

"Artificial Intelligence Act: Deal on Comprehensive Rules for Trustworthy AI." European Parliament, December 9, 2023. https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai.

Betker, James, Gabriel Goh, Li Jing, et al. "Improving Image Generation with Better Captions," n.d.

Blattmann, Andreas, Tim Dockhorn, Sumith Kulal, et al. "Stable Video Diffusion: Scaling Latent Video Diffusion Models to Large Data Sets," n.d.

Bousquette, Isabelle. "The AI Boom Is Here. The Cloud May Not Be Ready." Wall Street Journal, July 10, 2023. www.wsj.com/articles/the-ai-boom-is-here-the-cloud-may-not-be-ready-1a51724d.

Chatterjee, Mohar. "White House Offers a New Strategy for AI— and Picks New Fights." Politico, October 30, 2023. www.politico.com/news/2023/10/30/biden-ai-tech-industry-policy-00124185.

Chen, Ziqi, et al. "G2Retro as a Two-Step Graph Generative Models for Retrosynthesis Prediction." Communications Chemistry, vol. 6, no. 1 (May 30, 2023): pp. 1–19, www.nature.com/articles/s42004-023-00897-3.

"Chexia Face Recognition." NIST. Accessed January 25, 2023. https://www.nist.gov/programs-projects/chexia-face-recognition.

Chitty-Venkata, Krishna Teja, and Arun K. Somani. "Neural Architecture Search Survey: A Hardware Perspective." ACM Computing Surveys 55, no. 4 (April 30, 2023): 1–36. https://doi.org/10.1145/3524500.

Cowen, Tyler. "AI's Greatest Danger? The Humans Who Use It." Bloomberg.com, January 25, 2024. www.bloomberg.com/opinion/articles/2024-01-25/ais-greatest-danger-the-humans-who-use-it.

Cowen, Tyler. "Your Child's Favorite Teacher May Soon Be a Chatbot." Bloomberg.com, January 17, 2024. www.bloomberg.com/opinion/articles/2024-01-17/ai-in-the-classroom-everyone-s-favorite-teacher-may-soon-be-a-chatbot.

"Dawn of the EU's AI Act: Political Agreement Reached on World's First Comprehensive Horizontal AI Regulation." White & Case LLP, December 14, 2023. https://www.whitecase.com/insight-alert/dawn-eus-ai-act-political-agreement-reached-worlds-first-comprehensive-horizontal-ai.

"Decomposing Language Models Into Understandable Components." Accessed February 21, 2024. https://www.anthropic.com/news/decomposing-language-models-into-understandable-components.

DeepMind. "Mastering Atari, Go, Chess and Shogi by Planning with a Learned Model." https://deepmind.com/research/publications/2019/Mastering-Atari-Go-Chess-and-Shogi-by-Planning-with-a-Learned-Model.

DeepMind. "MuZero: Mastering Go, chess, shogi and Atari Without rules." https://deepmind.com/blog/article/muzero-mastering-go-chess-shogi-and-atari-without-rules.

DeepMind. "Putting the Power of AlphaFold into the World's Hands." https://deepmind.com/blog/article/putting-the-power-of-alphafold-into-the-worlds-hands.

Dong, Tian, Bo Zhao, and Lingjuan Lyu. "Privacy for Free: How Does Dataset Condensation Help Privacy?" ArXiv, June 1, 2022. http://arxiv.org/abs/2206.00240.

Driess, Danny, Fei Xia, Mehdi S. M. Sajjadi, et al. "PaLM-E: An Embodied Multimodal Language Model." ArXiv, March 6, 2023. https://doi.org/10.48550/arXiv.2303.03378.

Eger, Steffen, Christoph Leiter, Jonas Belouadi, et al. "NLLG Quarterly ArXiv Report 06/23: What Are the Most Influential Current AI Papers?" ArXiv, July 31, 2023. https://doi.org/10.48550/arXiv.2308.04889.

SELECTED SOURCES

"ESM." Accessed January 25, 2023. https://huggingface.co/docs/transformers/model_doc/esm.

"EU Artificial Intelligence Act | Up-to-Date Developments and Analyses of the EU AI Act." Accessed February 21, 2024. https://artificialintelligenceact.eu/.

Fedus, William, Barret Zoph, and Noam Shazeer. "Switch Transformers: Scaling to Trillion Parameter Models with Simple and Efficient Sparsity." arXiv:2101.03961 [Cs], January 11, 2021. http://arxiv.org/abs/2101.03961.

Gandhi, Sanchit, Patrick von Platen, and Alexander M. Rush. "Distil-Whisper: Robust Knowledge Distillation via Large-Scale Pseudo Labelling." ArXiv, November 1, 2023. https://doi.org/10.48550/arXiv.2311.00430.

Gibney, Elizabeth. "What the EU's Tough AI Law Means for Research and ChatGPT." Nature, February 16, 2024. https://doi.org/10.1038/d41586-024-00497-8.

GitHub. "Ablation Study: Why ControlNets Use Deep Encoder? What If It Was Lighter? Or Even an MLP?." Accessed February 21, 2024. https://github.com/lllyasviel/ControlNet/discussions/188.

Google DeepMind. "FunSearch: Making New Discoveries in Mathematical Sciences Using Large Language Models." December 14, 2023. https://deepmind.google/discover/blog/funsearch-making-new-discoveries-in-mathematical-sciences-using-large-language-models/.

Google Research. "Med-PaLM." Google Research Med-PaLM, 2023. sites.research.google/med-palm/.

Gunasekar, Suriya, Yi Zhang, Jyoti Aneja, et al. "Textbooks Are All You Need." ArXiv, October 2, 2023. https://doi.org/10.48550/arXiv.2306.11644.

Hafner, Danijar, Jurgis Pasukonis, Jimmy Ba, and Timothy Lillicrap. "Mastering Diverse Domains through World Models." ArXiv, January 10, 2023. https://doi.org/10.48550/arXiv.2301.04104.

Hassan, Oz. "Artificial Intelligence, Neom and Saudi Arabia's Economic Diversification from Oil and Gas." The Political Quarterly, vol. 91, no. 1 (January 2020): pp. 222–227. https://doi.org/10.1111/1467-923x.12794.

Helwan, Abdulkader. "A List of the Available Medical Large Language Models: Med-LLMs." Medium, November 27, 2023. abdulkaderhelwan.medium.com/a-list-of-the-available-medical-large-language-models-med-llms-f087119fa89d.

House, The White. "Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence." October 30, 2023. https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/.

Howard, Lisa. "New AI Technology Helps Physicians Quickly Identify Stroke." UC Davis Health, February 1, 2024. health.ucdavis.edu/news/headlines/new-ai-technology-helps-physicians-quickly-identify-stroke/2024/02.

Hutson, Matthew. "Hypotheses Devised by AI Could Find "Blind Spots" in Research." Nature, November 17, 2023, www.nature.com/articles/d41586-023-03596-0.

IBM. "What Is a Vector Database?." www.ibm.com/topics/vector-database.

IBM Education. "The Benefits of AI in Healthcare." IBM Blog, July 11, 2023. www.ibm.com/blog/the-benefits-of-ai-in-healthcare/.

IBM Research. "Neuro-Symbolic AI." research.ibm.com/topics/neuro-symbolic-ai.

"Join Us in the AI Test Kitchen." Google blog. https://blog.google/technology/ai/join-us-in-the-ai-test-kitchen/.

"Joon Sung Park | Generative Agents: Interactive Simulacra of Human Behavior." YouTube. https://www.youtube.com/watch?v=nKCJ3BMUy1s.

Kannampallil, Thomas, et al. "Effects of a Virtual Voice-Based Coach Delivering Problem-Solving Treatment on Emotional Distress and Brain Function: A Pilot RCT in Depression and Anxiety." Translational Psychiatry, vol. 13, no. 1 (May 12, 2023): pp. 1–8, www.nature.com/articles/s41398-023-02462-x.

Kearns, Jeff. "AI's Reverberations across Finance." IMF, December 2023. www.imf.org/en/Publications/fandd/issues/2023/12/AI-reverberations-across-finance-Kearns.

King-Smith, Emma, et al. "Predictive Minisci Late Stage Functionalization with Transfer Learning." Nature Communications, vol. 15, no. 1 (January 15, 2024): p. 426. www.nature.com/articles/s41467-023-42145-1.

King-Smith, Emma, et al. "Probing the Chemical 'Reactome' with High-Throughput Experimentation Data." Nature Chemistry, January 2, 2024: pp. 1–11. www.nature.com/articles/s41557-023-01393-w.

SELECTED SOURCES

Kirillov, Alexander, Eric Mintun, Nikhila Ravi, et al. "Segment Anything." ArXiv, April 5, 2023. https://doi.org/10.48550/arXiv.2304.02643.

Koc, Vincent. "Navigating the AI Landscape of 2024: Trends, Predictions, and Possibilities." Medium, January 4, 2024. towardsdatascience.com/navigating-the-ai-landscape-of-2024-trends-predictions-and-possibilities-41e0ac83d68f.

Koponen, Jonna, et al. "Work Characteristics Needed by Middle Managers When Leading AI-Integrated Service Teams." Journal of Service Research, December 13, 2023. https://doi.org/10.1177/10946705231220462.

Liu, Xian, Xiaohang Zhan, Jiaxiang Tang, et al. "HumanGaussian: Text-Driven 3D Human Generation with Gaussian Splatting," n.d.

Llama. "Llama 2." https://llama.meta.com/llama2/.

"LLaVA." Accessed February 21, 2024. https://llava-vl.github.io/.

Lorach, Henri, et al. "Walking Naturally after Spinal Cord Injury Using a Brain–Spine Interface." Nature, vol. 618 (May 24, 2023): pp. 1–8. www.nature.com/articles/s41586-023-06094-5.

Lucente, Adam. "Will US Pressure on UAE-China Tech Cooperation Pay Off?." AI Monitor, December 11, 2023. www.al-monitor.com/originals/2023/12/will-us-pressure-uae-china-tech-cooperation-pay.

Maples, Bethanie, et al. "Loneliness and Suicide Mitigation for Students Using GPT3-Enabled Chatbots." NPJ Mental Health Research, vol. 3, no. 1 (January 22, 2024): pp. 1–6. www.nature.com/articles/s44184-023-00047-6.

"Mastering Diverse Domains through World Models." Accessed February 21, 2024. https://danijar.com/project/dreamerv3/.

Mehrotra, Sukriti. "DALL-E 3: A Fusion of Imagination and Conversation." Medium, September 21, 2023. medium.com/@sukritimehrotra/dall-e-3-a-fusion-of-imagination-and-conversation-4c8ec8930442.

Merchant, Amil, et al. "Scaling Deep Learning for Materials Discovery." Nature, November 29, 2023: pp. 1–6. www.nature.com/articles/s41586-023-06735-9.

Microsoft News Center. "G42 and Microsoft Unlock New Opportunities for Digital Transformation with Joint Sovereign Cloud and AI Offering." September 5, 2023, news.microsoft.com/en-xm/2023/09/05/g42-and-microsoft-unlock-new-opportunities-for-digital-transformation-with-joint-sovereign-cloud-and-ai-offering/.

Mittal, Aayush. "The Role of Vector Databases in Modern Generative AI Applications." Unite.AI, October 11, 2023. www.unite.ai/the-role-of-vector-databases-in-modern-generative-ai-applications/.

Mulligan, Deirdre K. and Mina Hsiang. "A Call to Service for AI Talent in the Federal Government." The White House, January 29, 2024. www.whitehouse.gov/ostp/news-updates/2024/01/29/a-call-to-service-for-ai-talent-in-the-federal-government/.

Myers, Andrew. "AI's Powers of Political Persuasion." Stanford HAI, February 27, 2023. hai.stanford.edu/news/ais-powers-political-persuasion.

Nellis, Stephen. "Intel to Spin out AI Software Firm with Outside Investment." Yahoo Finance, January 3, 2024, finance.yahoo.com/news/intel-spins-ai-software-firm-133626026.html.

Nippa, David F., et al. "Enabling Late-Stage Drug Diversification by High-Throughput Experimentation with Geometric Deep Learning." Nature Chemistry, vol. 16 (November 23, 2023). www.nature.com/articles/s41557-023-01360-5.

Nour, Matthew M, et al. "Trajectories through Semantic Spaces in Schizophrenia and the Relationship to Ripple Bursts." Proceedings of the National Academy of Sciences of the United States of America, vol. 120, no. 42 (October 10, 2023). https://doi.org/10.1073/pnas.2305290120.

Pan, Xingang, Ayush Tewari, Thomas Leimkühler, et al. "Drag Your GAN: Interactive Point-Based Manipulation on the Generative Image Manifold." ArXiv, May 18, 2023. https://doi.org/10.48550/arXiv.2305.10973.

Park, Joon Sung, Joseph C. O'Brien, Carrie J. Cai, et al. "Generative Agents: Interactive Simulacra of Human Behavior." ArXiv, August 5, 2023. https://doi.org/10.48550/arXiv.2304.03442.

Peplow, Mark. "Google AI and Robots Join Forces to Build New Materials." Nature, November 9, 2023. www.nature.com/articles/d41586-023-03745-5.

Qureshi, Nabeel S. "Moore's Law for Intelligence." Digital Spirits, February 5, 2024. digitalspirits.substack.com/p/moores-law-for-intelligence.

Ramesh, Aditya, Prafulla Dhariwal, Alex Nichol, Casey Chu, and Mark Chen. "Hierarchical Text-Conditional Image Generation with CLIP Latents." ArXiv, April 12, 2022. http://arxiv.org/abs/2204.06125.

"Responsible AI Guidelines." https://www.diu.mil/responsible-ai-guidelines.

SELECTED SOURCES

Rodman, Adam, et al. “Artificial Intelligence vs Clinician Performance in Estimating Probabilities of Diagnoses Before and After Testing.” JAMA Network Open, vol. 6, no. 12 (December 11, 2023): p. e2347075. jamanetwork.com/journals/jamanetworkopen/fullarticle/2812737.

Roser, Max, et al. “What Is Moore’s Law?” Our World in Data, March 28, 2023. ourworldindata.org/moores-law.

Runway. “Gen-1 by Runway.” https://research.runwayml.com/gen1.

Ryan-Mosley, Tate. “AI Isn’t Great at Decoding Human Emotions. So Why Are Regulators Targeting the Tech?” MIT Technology Review, August 14, 2023. www.technologyreview.com/2023/08/14/1077788/ai-decoding-human-emotions-target-for-regulators/.

Savage, Rashaud. “Welcoming Mistral, Phi, Jais, Code Llama, NVIDIA Nemotron, and More to the Azure AI Model Catalog.” Microsoft, November 15, 2023, techcommunity.microsoft.com/t5/ai-machine-learning-blog/welcoming-mistral-phi-jais-code-llama-nvidia-nemotron-and-more/ba-p/3982699.

Savcisens, Germans, et al. “Using Sequences of Life-Events to Predict Human Lives.” Nature Computational Science, vol. 4, 18 (December 2023): pp. 1–14, www.nature.com/articles/s43588-023-00573-5.

Schmidt, Eric. “Eric Schmidt: This Is How AI Will Transform the Way Science Gets Done.” MIT Technology Review, July 5, 2023. www.technologyreview.com/2023/07/05/1075865/eric-schmidt-ai-will-transform-science/.

“Segment Anything | Meta AI.” https://segment-anything.com/.

Sheth, Amit, et al. “Neurosymbolic Artificial Intelligence (Why, What, and How).” IEEE Journals & Magazine vol. 38, no. 3 (May 2023): pp. 56–62, ieeexplore.ieee.org/document/10148662.

Shwartz-Ziv, Ravid, and Amitai Armon. “Tabular Data: Deep Learning Is Not All You Need.” ArXiv, November 23, 2021. http://arxiv.org/abs/2106.03253.

Simbo.ai—Blog. “How AI Technologies Improve the Patient Experience.” September 23, 2021. www.simbo.ai/blog/index.php/2021/09/23/how-ai-technologies-improve-the-patient-experience/.

Stanford University. “Artificial Intelligence Index Report 2023 Introduction to the AI Index Report 2023.” 2023.

Technical University of Denmark. “Artificial Intelligence Can Predict Events in People’s Lives.” ScienceDaily, December 18, 2023. www.sciencedaily.com/releases/2023/12/231218125850.htm.

The Alan Turing Institute. “Neruo-Symbolic AI.” www.turing.ac.uk/research/interest-groups/neuro-symbolic-ai.

The Isomorphic Labs team and Google DeepMind AlphaFold team. “A Glimpse of the Next Generation of AlphaFold.” Isomorphic Labs, October 31, 2023. www.isomorphiclabs.com/articles/a-glimpse-of-the-next-generation-of-alphafold.

The University of Montana. “AI Tests into Top 1% for Original Creative Thinking.” ScienceDaily, July 5, 2023, www.sciencedaily.com/releases/2023/07/230705154051.htm.

Topics, European Parliament. “EU AI Act: First Regulation on Artificial Intelligence,” June 8, 2023. https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence.

Trinh, Trieu, and Thang Luong. “AlphaGeometry: An Olympiad-Level AI System for Geometry.” Google DeepMind, January 17, 2024. deepmind.google/discover/blog/alphageometry-an-olympiad-level-ai-system-for-geometry/.

Trinh, Trieu H, et al. “Solving Olympiad Geometry without Human Demonstrations.” Nature, vol. 625, no. 7995, (January 17, 2024): pp. 476–482. https://doi.org/10.1038/s41586-023-06747-5.

Tuli, Shreshth, Giuliano Casale, and Nicholas R. Jennings. “TranAD: Deep Transformer Networks for Anomaly Detection in Multivariate Time Series Data.” ArXiv, May 14, 2022. http://arxiv.org/abs/2201.07284.

Vesuvius Challenge. “Vesuvius Challenge 2023 Grand Prize Awarded: We Can Read the Scrolls!” Scrollprize.org, February 5, 2024, scrollprize.org/grandprize.

Vora, Lalitkumar K, et al. “Artificial Intelligence in Pharmaceutical Technology and Drug Delivery Design.” Pharmaceutics, vol. 15, no. 7 (July 10, 2023): pp. 1916–1916https://doi.org/10.3390/pharmaceutics15071916.

“Voxel.” https://www.voxelai.com/.

Wang, Chengyi, Sanyuan Chen, Yu Wu, et al. “Neural Codec Language Models Are Zero-Shot Text to Speech Synthesizers.” ArXiv, January 5, 2023. https://doi.org/10.48550/arXiv.2301.02111.

SELECTED SOURCES

Wornow, Michael, et al. “The Shaky Foundations of Foundation Models in Healthcare.” Stanford University, February 27, 2023. hai.stanford.edu/news/shaky-foundations-foundation-models-healthcare.

Wu, Qingyun, Gagan Bansal, Jieyu Zhang, et al. “AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation.” Microsoft, August 16, 2023. https://www.microsoft.com/en-us/research/publication/autogen-enabling-next-gen-llm-applications-via-multi-agent-conversation-framework/.

Yadlowsky, Steve, et al. “Pretraining Data Mixtures Enable Narrow Model Selection Capabilities in Transformer Models.” ArXiv (Cornell University), November 1, 2023. https://doi.org/10.48550/arxiv.2311.00871.

Yim, Kristin, and Hema Manickavasagam. “Turn Ideas into Music with MusicLM.” Google, May 10, 2023. blog.google/technology/ai/musiclm-google-ai-test-kitchen/.

Youyou, Wu, et al. “How AI Can Help Researchers Navigate the “Replication Crisis.”” Kellogg Insight, July 1, 2023. insight.kellogg.northwestern.edu/article/how-ai-can-help-researchers-navigate-the-replication-crisis.

Zewe, Adam. “Technique Enables AI on Edge Devices to Keep Learning over Time.” MIT News, November 16, 2023. news.mit.edu/2023/technique-enables-ai-edge-devices-keep-learning-over-time.

Zhang, Can, Tianyu Yang, Junwu Weng, et al. “Unsupervised Pre-Training for Temporal Action Localization Tasks.” ArXiv, March 25, 2022. http://arxiv.org/abs/2203.13609.

Zhang, Jianjing, et al. “Neural Rendering-Enabled 3D Modeling for Rapid Digitization of In-Service Products.” CIRP Annals, vol. 72, no. 1 (January 1, 2023): pp. 93–96. www.sciencedirect.com/science/article/abs/pii/S0007850623000409.

Zhang, Lvmin, Anyi Rao, and Maneesh Agrawala. “Adding Conditional Control to Text-to-Image Diffusion Models.” ArXiv, November 26, 2023. https://doi.org/10.48550/arXiv.2302.05543.

Zhang, Susan, Stephen Roller, Naman Goyal, et al. “OPT: Open Pre-Trained Transformer Language Models.” ArXiv, June 21, 2022. http://arxiv.org/abs/2205.01068.

Zhao, Mingmin, et al. “Assessment of Medication Self-Administration Using Artificial Intelligence.” Nature Medicine, March 18, 2021. https://doi.org/10.1038/s41591-021-01273-1.

Zhou, Tingtao, et al. “AI-Aided Geometric Design of Anti-Infection Catheters.” ArXiv.org, April 27, 2023. arxiv.org/abs/2304.14554.

Zhu, Luyang, Dawei Yang, Tyler Zhu, et al. “TryOnDiffusion: A Tale of Two UNets.” ArXiv, June 14, 2023. https://doi.org/10.48550/arXiv.2306.08276.

ABOUT FUTURE TODAY INSTITUTE

ABOUT FUTURE TODAY INSTITUTE

About Us

Founded in 2006, Future Today Institute is an advisory firm specializing in strategic foresight, driving corporate strategies that lead to long-term success and resilience.

Future Today Institute partners with global business leaders to navigate disruptive change and uncertain futures with confidence. We fuel actionable strategic decisions that prepare you to take on global challenges, create sustainable value and ensure long-term growth.

As the global leaders in strategic foresight, our rigorous data- and research-driven methodology positions us to anticipate the unexpected and develop strategically driven roadmaps to manage risks and take advantage of opportunities today, tomorrow and into the future.

We empower leaders to make better decisions about the future, today.

Contact Us

For an introductory conversation to learn how Future Today Institute can assist your organization with its strategic planning and foresight needs, please contact:

inquiries@futuretodayinstitute.com

futuretodayinstitute.com

+1 267 342 4300

Social

Twitter (X)	**@FTI**
Linkedin	**@Future-Today-Institute**
Facebook	**@FutureTodayInstitute**
Instagram	**@futuretodayinstitute**

METHODOLOGY

Future Today Institute conducts in-depth qualitative and quantitative research throughout the year to identify emerging trends. We review patent and trademark filings, pre-print and published scientific papers, investment rounds, online search trends, macroeconomic data, publications from governments worldwide, news mentions, influencer posts and other sources, and we use a proprietary system to identify patterns, which are then grouped into nodes and evaluated using a set of standardized indicators. Qualified trends are further scored for their trajectory, momentum and timing. Additionally, we harness the deep subject matter expertise of our Future Today Institute network, leading to valuable insights about the topics we cover.

In continuous publication since 2007, Future Today Institute's annual report includes maturing and emerging trends grouped into two categories: industry and technology. Industry trends reflect the ways in which technology is shaping the future of an entire industry. Technology trends are specific developments within one arena, such as artificial intelligence. Covering a wide range of technologies across industry sectors creates a holistic view of change and provides leaders with a clear understanding of their potential impact. Trends are published as individual Industry and Technology reports, as well as in one combined report with all of our research.

Monitored regularly, trends help executives recognize emerging threats and opportunities in the near-term and enable them to develop perspectives, strategies and plans for the future.

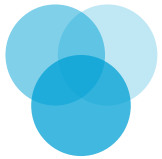
Future Today Institute's Strategic Foresight Methodology

SIGNALS & LONG-TERM TRENDS

What is INFLUENCING the future?

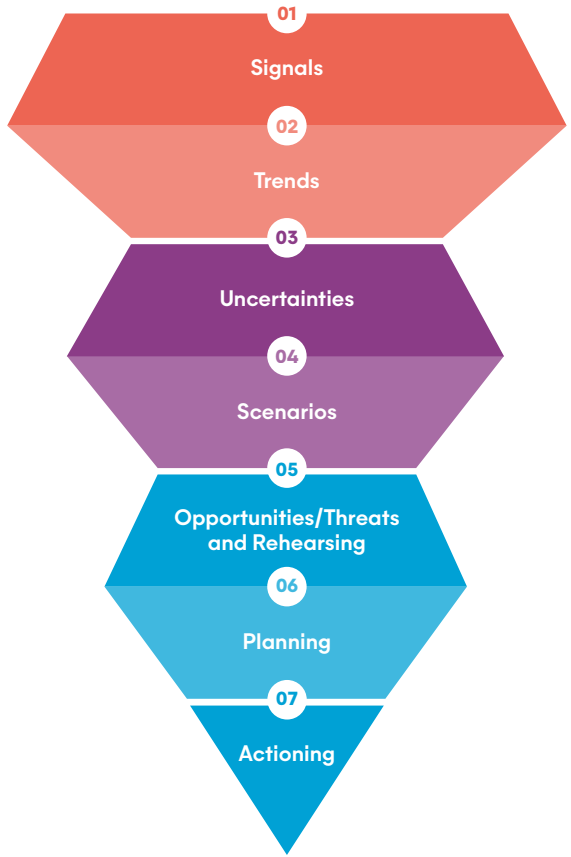
GLOBAL MACRO SCENARIOS

What is THE future?



STRATEGIC

What is YOUR ORG'S future?



DISCLAIMER

The views expressed herein are the authors' own and are not representative of the greater organizations in which they have been employed. The names of companies, services, and products mentioned in this report are not necessarily intended as endorsements by Future Today Institute or this report's authors.

Future Today Institute's 2024 Tech Trends Report relies on data, analysis, and modeling from a number of sources, which includes sources within public and private companies, securities filings, patents, academic research, government agencies, market research firms, conference presentations and papers, and news media stories. Additionally, this report draws from Future Today Institute's previous EMT Trends Reports, FTI Trend Reports, and newsletters. FTI's reports are occasionally updated on the FTI website.

FTI advises hundreds of companies and organizations, some of which are referenced in this report. FTI does not own any equity position in any of the entities listed in this presentation.

Any trademarks or service marks used in this report are the marks of their respective owners, who do not endorse the statements in this report. All rights in marks are reserved by their respective owners. We disclaim any and all warranties, expressed or implied, with respect to this report.

USING AND SHARING THE MATERIAL IN THIS REPORT

You are free to:

Download the materials for your or your organization's non-commercial use.

Use the assets with full attribution to Future Today Institute using the following language: "This material is copyrighted by Future Today Institute and licensed under the Creative Commons 4.0 International License (CC BY-NC-ND 4.0)."

Use assets in a reasonable manner, but not in any way that suggests that Future Today Institute endorses you or your use.

You may not:

Use the materials for any commercial purposes, which includes: selling, licensing or charging for products or services that use the materials.

Publish the materials in a book or upload to a website or intranet without prior written permission from Future Today Institute.

Remix, transform or build upon the material without giving appropriate credit to Future Today Institute as the originator of the work using the following language: "This material is adapted from original work copyrighted by Future Today Institute and licensed under the Creative Commons 4.0 International License (CC BY-NC-ND 4.0)."

Assert a relationship with or endorsement from Future Today Institute.

Future Today Institute