

A space shuttle is shown in orbit around Earth, viewed from a low angle. The shuttle is white with a dark blue nose and a large orange external tank. The Earth's horizon is visible in the background, with a bright sun or star creating a lens flare effect. The shuttle is angled upwards and to the right.

White Paper

# AI Post-ChatGPT:

## A Progress Report and Future Forecast

### Authors

Di Le - HCAI | RAI Design Strategist | ServiceNow

Adir Mancebo Jr., Ph.D. - Data Project Manager | Data Science Alliance

Hayley Mortin - Sr. UX Researcher | ServiceNow

Fanny Rancourt - Data Scientist | ServiceNow Research



---

## Contributors

Kallyn Hobmann

Patricia Lopez

Phish Jeffrey Moore



---

## Introduction

In the history of human achievement, some milestones fundamentally reshape our understanding of the possible, drawing a bold line between what is known and the uncharted territories that lie ahead. Much like the moment humanity first breached Earth's atmosphere into the vast expanse of space, the advent of ChatGPT and other AI technologies presents a pivotal leap in our human journey—a journey into the cognitive cosmos where artificial intelligence (AI) becomes more than just a tool, but also a partner in creation, exploration, and human understanding.

Before the launch of ChatGPT, our interaction with artificial intelligence was largely one-directional: We commanded, and it executed. Additionally, access to and usage of such tools were limited to AI practitioners, researchers, and those with the technical knowledge needed to apply AI to use cases. While impressive results were achieved, like advancements in medicine, smart home systems and more, many interactions were confined within the operational orbit dictated by explicit programming and rigid algorithms.

After ChatGPT, however, we find ourselves in a new era where AI technologies invite us to explore the vast potential of Generative AI (GenAI). AI represents a monumental leap for humankind, ushering in an era where the boundaries of creativity, productivity, and interaction are continually expanding. Like astronauts venturing into new frontiers, GenAI models, with their multimodal capabilities, forge profound new pathways for human-AI collaboration across languages, images, video, audio and code.

This white paper analyzes the evolving landscape of GenAI since the pivotal launch of ChatGPT, reflecting on 1.5 years of advancements, real-world applications, and challenges navigated in that time. We will explore the potential trajectories of GenAI technology, its luminous possibilities, as well as the direct and nuanced ethical dilemmas it holds for various sectors of society.

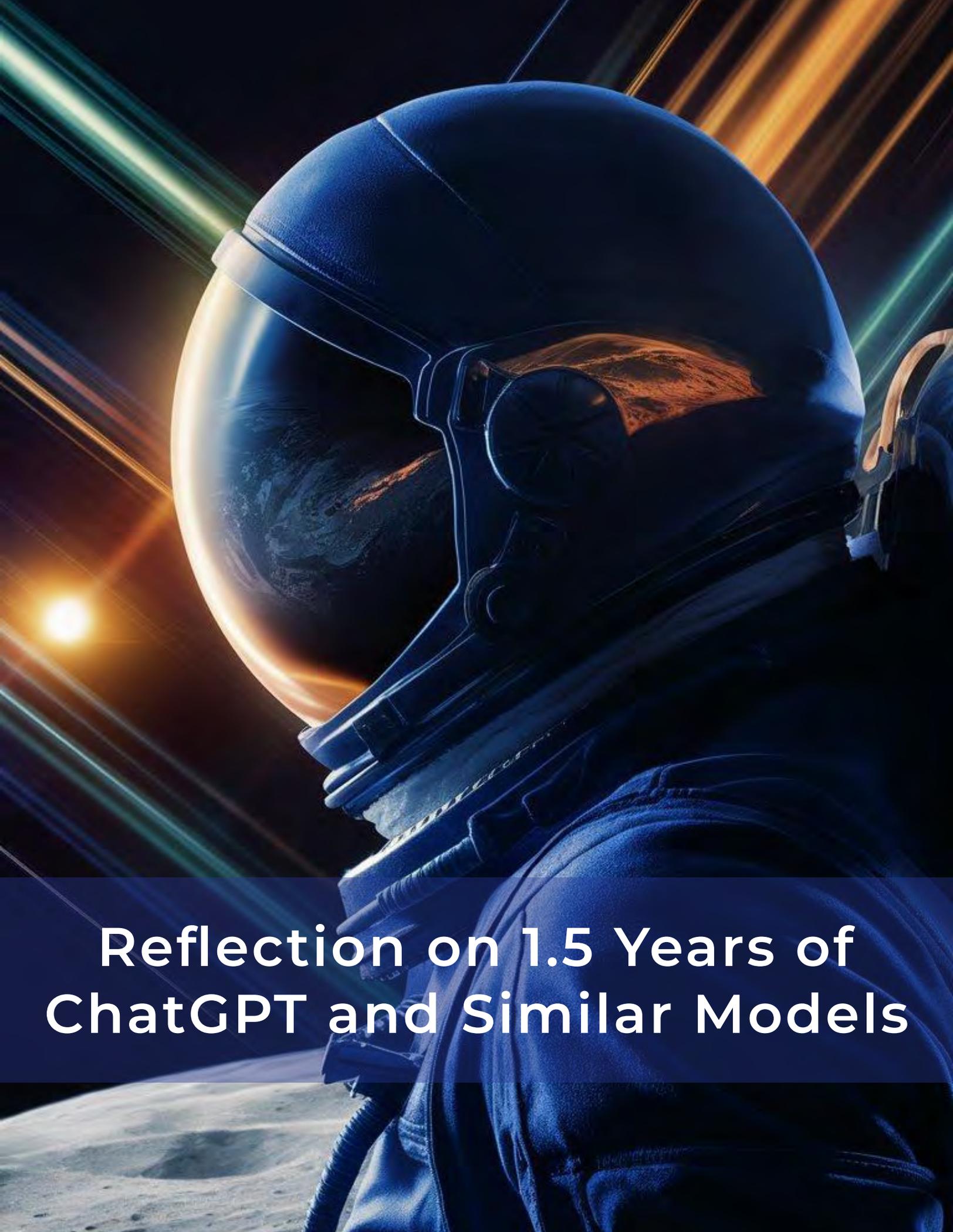


# Table of Contents

Contributors

Introduction

<b>Reflection on 1.5 Years of ChatGPT and Similar Models</b>	<b>5</b>
Advancements in Generative AI Technology	6
Insights & Challenges from the Real World	9
<b>Potential Societal Impacts and Ethical Considerations</b>	<b>12</b>
Responsible Data Science Principles and Contestability in AI Development	13
UX as the Bridge to Responsible AI	16
AI for All? The Technological Access Gap	17
Evaluating the Impacts of Open vs. Closed AI Development	20
The Dangers of an AI Arms Race	22
<b>Impact of Generative AI on Industry</b>	<b>24</b>
Opportunities for Innovation and Disruption	27
The Challenges Experienced by Industries	28
<b>Changes Since the Release of ChatGPT</b>	<b>31</b>
Shifts in Public Perception and Acceptance of AI Technology	32
Regulatory Developments and Policy Implications	33
Evolution of User Interactions with Generative AI Platforms	35
<b>Predictions for the Future of AI</b>	<b>36</b>
Unexplored Applications and Beyond Chatbots	37
The Next Steps in the Development of Generative AI	39
Anticipated Technological Developments	41
Consequences for the Future of Work	43
<b>Conclusion</b>	<b>45</b>
<b>References</b>	<b>47</b>



# Reflection on 1.5 Years of ChatGPT and Similar Models

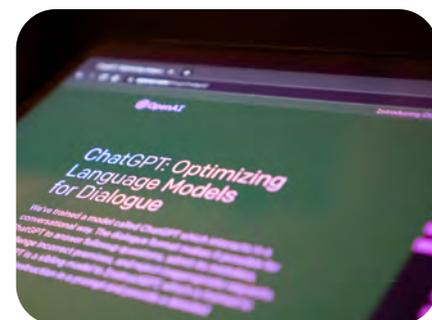


## Advancements in Generative AI Technology

Over the last 1.5 years since the release of ChatGPT, the global impact of this GenAI technology has been groundbreaking and far-reaching, especially with regards to the adoption and integration of AI technologies across various sectors. ChatGPT achieved unprecedented adoption rates, becoming one of the fastest-growing consumer applications in history. Within just a few months of its release, ChatGPT garnered an estimated 100 million active users [14], illustrating the immense interest in and applicability of GenAI technologies.

This rapid adoption eclipsed the initial uptake rates of many now-ubiquitous social media applications, like TikTok and Instagram, highlighting the model's appeal and the public's readiness to engage with advanced AI.

Moreover, advancements in the natural language processing (NLP) and language space have been both rapid and revolutionary in the past several months, focusing on enhancing the model's understanding, responsiveness, and contextual awareness.



### Key developments include:



#### 1. Fine-Tuning and Customization

The post-ChatGPT era has seen significant strides in fine-tuning models on specific datasets, allowing for more personalized and industry-specific applications [11, 16, 17].

## 2. Multimodal Capabilities

There has been a push towards developing models that can understand and generate not just text, but also images, audio, and video [45].



## 3. Language Diversity and Inclusion

Efforts have been made to improve the model's capabilities in handling a broader range of languages and dialects, promoting inclusivity and accessibility [22, 26, 30, 41].

## 4. Enhanced Ethical and Responsible AI Practices

An important development since the release of ChatGPT has been the unprecedented focus on setting standards and fostering global collaboration towards developing more ethical AI [38, 60]. This surge in interest and concerted efforts has engaged a diverse array of stakeholders, extending well beyond the confines of the technical community. The dialogue now emphasizes crucial issues such as human autonomy, privacy, data security, the societal implications of AI, and the future of work.



## 5. Interpretability and Explainability

There has been progress in making AI models more interpretable and explainable, helping users understand how decisions are made [1, 10]. This is needed to secure trust and accountability.

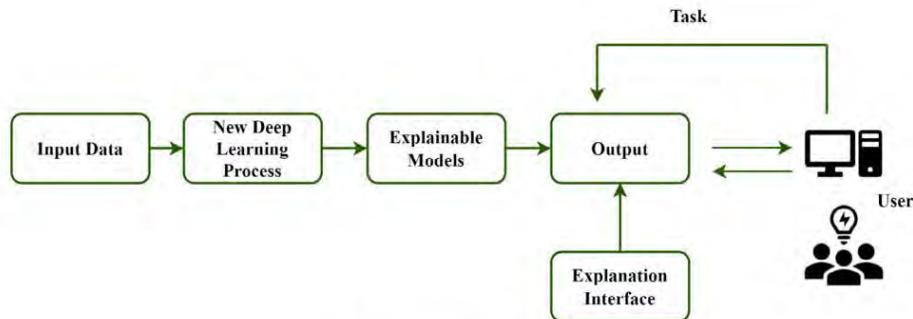


Figure 1: Process of explainable AI [1].

These advancements not only demonstrate the technological evolution in the NLP and language space but also highlight the broader shift toward creating more versatile, equitable, and human-centered AI tools. The release of ChatGPT has also acted as a catalyst for the emergence of numerous startup companies and innovations [40], focusing on leveraging GenAI for diverse applications such as content creation, education, customer service, and more.

**[GenAI] has fundamentally reshaped conversations around AI, sparking discussions on ethical considerations, the future of work, and the role of AI in society.**

ChatGPT's ripple effect extends further beyond technological advancements and new business ventures. It has fundamentally reshaped conversations around AI, sparking discussions on ethical considerations, the future of work, and the role of AI in society. This collective endeavor involves policymakers, educators, AI practitioners, ethicists, and the general public, all united in their aim to steer the development and deployment of AI technologies in a direction that respects ethical norms and societal values. Such a broad-based, collaborative approach marks a significant step towards establishing global standards for responsible AI.

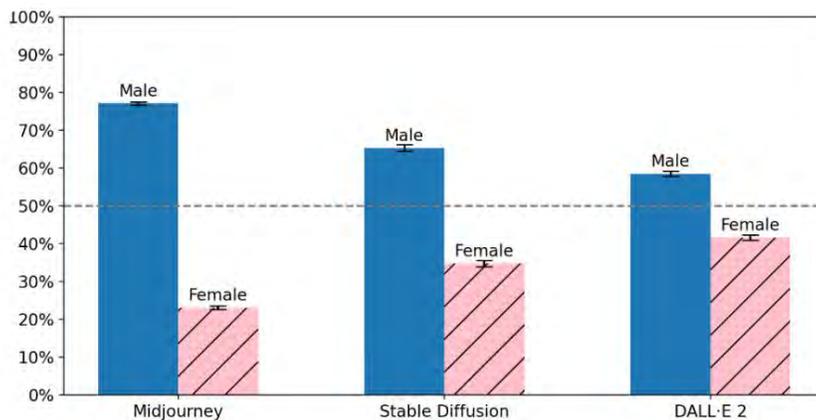
# Insights & Challenges from the Real World

Following ChatGPT's launch, several cautionary tales unique to GenAI have emerged, highlighting new challenges and exacerbating existing issues within machine learning (ML). These challenges serve as valuable lessons for researchers, developers, users, product developers, and policymakers alike:



## 1. Misinformation and Content Fabrication

GenAI models like ChatGPT can produce text that is often indistinguishable from human writing, raising concerns about the spread of misinformation [4, 21, 31]. Commonly known as “hallucinations,” false or misleading information generated by large language models (LLMs) like ChatGPT becomes an even greater challenge when systems perform without oversight.



## 2. Biases in AI Outputs

While bias in ML models is not a new issue, the scale and sophistication of GenAI amplify the potential for harm [46]. The nuanced and complex nature of GenAI outputs makes these biases harder to detect and correct [24].

Figure 2: Gender distribution in occupational portraits created by AI generators [46].



## 3. Privacy Concerns

GenAI models, trained on extensive public text data, prompt privacy concerns regarding the unsanctioned use of personal data [28, 69]. Individual writings or dialogues might be used in training datasets, risking the creation of content that mirrors private information. The human-like quality of GenAI conversations could also foster a false sense of security, leading users to share sensitive information unwittingly.



## 4. Intellectual Property Challenges

GenAI's ability to produce original content based on existing works poses new challenges for copyright and intellectual property laws. Determining the ownership of AI-generated content and the implications for copyright infringement are areas of ongoing debate [7, 14, 65].

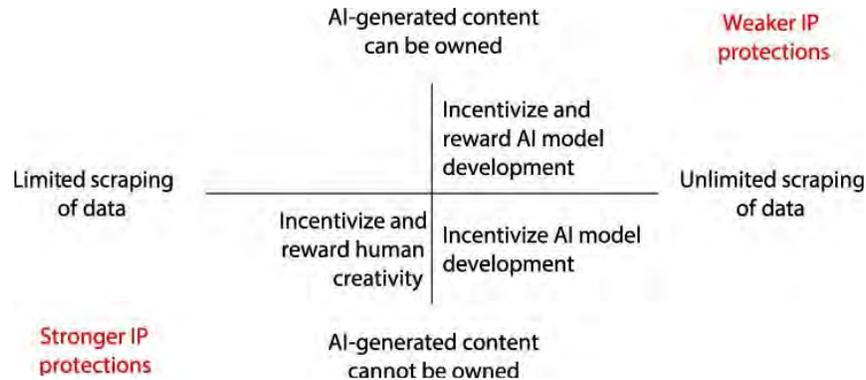


Figure 3: Incentives for humans and AI developers [7].

## 5. Deepfakes and Synthetic Media

Advances in GenAI have increased the sophistication of deepfakes and other forms of synthetic media, raising concerns about the authenticity of digital content [59, 66].



Figure 4: Fake photos generated by AI depicting Pope Francis in a designer coat went viral in March 2023.



## 6. Human-AI Interaction and Dependency

As GenAI becomes more integrated into daily life, there is a growing concern about over-reliance on AI for decision-making, creativity, and interpersonal communication [42, 58].

These cautionary tales highlight the complex interplay between technological advancements and societal impact. In today's AI landscape, the global discourse and work around responsible and ethical AI are inescapable.



# Potential Societal Impacts and Ethical Considerations



# Responsible Data Science Principles and Contestability in AI Development

The importance of fairness, transparency, privacy, and veracity in AI development cannot be overstated, especially as we delve into the realm of human-centered AI. This concept places a premium on ensuring that AI technologies are designed with the user's needs and ethical considerations at their core, thereby underscoring the significance of creating AI systems that are not only effective but also equitable and understandable by the general public. Human-centered AI heralds a shift toward more inclusive and ethically aligned technological advancements, making it a critical focus for developers and researchers.

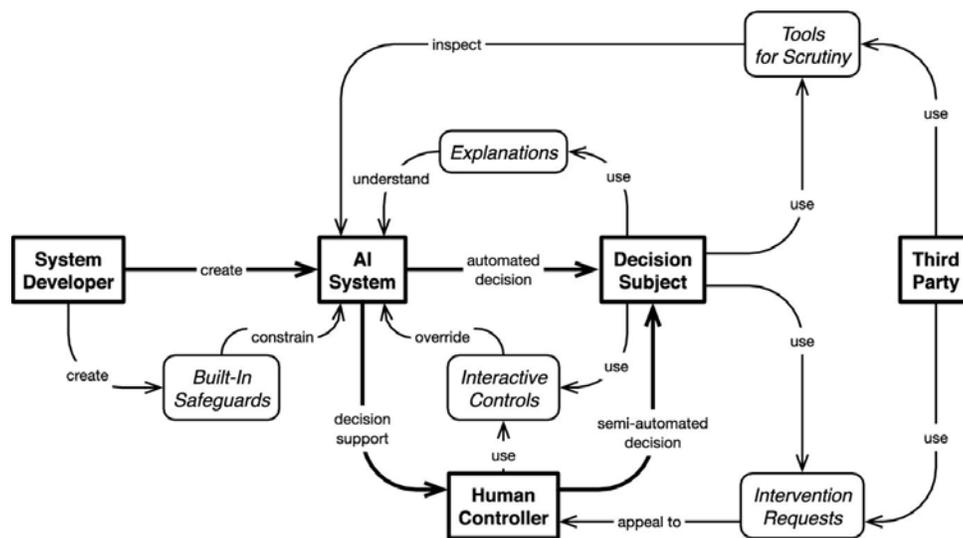


Figure 5: Features contributing to contestable AI [2].



GenAI represents a seismic shift from the software design philosophies that have dominated the past few decades. Unlike traditional deterministic software, which operates under predictable outcomes, GenAI introduces an element of inherent unpredictability with its probabilistic outcomes. This unpredictability brings new challenges in user experience (UX) design, such as dealing with novel types of inaccuracies like hallucinations, which necessitate drawing human attention to previously unconsidered issues.

**GenAI introduces an element of inherent unpredictability with its probabilistic outcomes. This unpredictability brings new challenges in user experience (UX) design.**

As such, foundational UX elements like latency and error messages are being reevaluated in contexts that demand a nuanced understanding of human interaction with AI. The advent of GenAI imposes additional risks, yet it also offers vast opportunities for optimization and efficiency, highlighting the need for designs that emphasize human control and oversight to facilitate broader adoption across various sectors.

Looking five years ahead, fostering trust in AI necessitates prioritizing fairness, transparency, privacy, and veracity throughout its development. These principles transcend ethical guidelines; they are cornerstones that will shape AI's societal integration:



### **Fairness**

is paramount to prevent discriminatory or unequal outcomes. We should design AI systems to equitably benefit all members of society by mitigating bias.



### **Transparency**

empowers users and stakeholders to comprehend AI systems' inner workings and decision-making processes. It fosters trust in the technology by demystifying its operation.



## Privacy

ensures that the personal information used to train and operate AI systems is handled securely and sensitively, protecting users from potential misuse.



## Veracity

guarantees that AI systems provide accurate and reliable information.

## Contestability empowers users to challenge AI decisions and seek recourse.

The field of AI acknowledges these principles as fundamental pillars for the ethical development and deployment of artificial intelligence. Alongside them, however, the concept of contestability emerges as equally significant. Contestability empowers users to challenge AI decisions and seek recourse [44].

This empowerment serves as a mechanism to ensure that individuals and communities have agency and influence over AI outputs, fostering a dynamic feedback loop that empowers users to challenge and improve the decisions made by AI systems. By embracing contestability, GenAI initiatives can engender a culture of continuous improvement and responsiveness to diverse perspectives, ultimately enhancing trust and democratizing the impact of artificial intelligence on society [2].

Trustworthiness can be enhanced through mechanisms such as open data access, participatory design processes, and avenues for redress and appeal, enabling stakeholders to engage meaningfully with AI systems and shape their outcomes in alignment with broader societal values and priorities.

## UX as the Bridge to Responsible AI

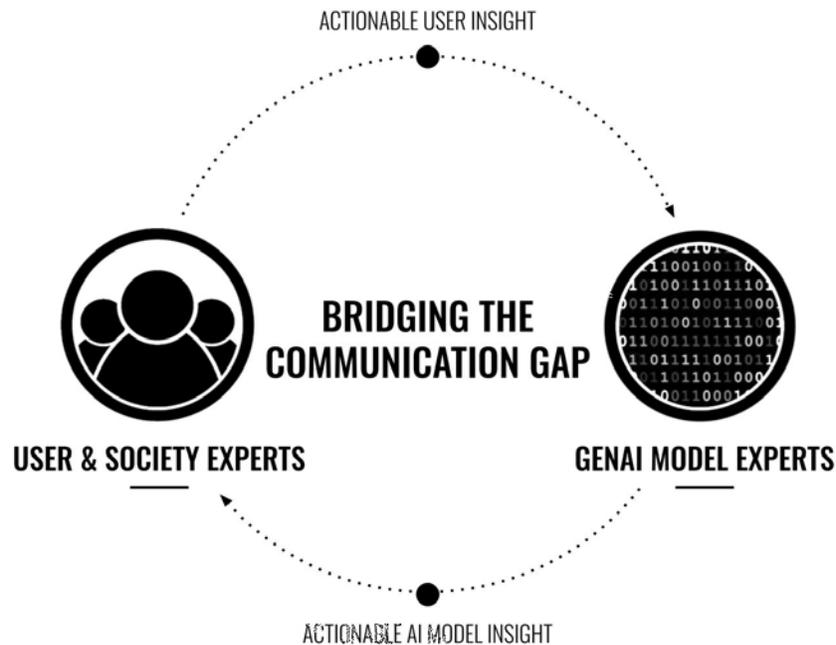


Figure 6: Bridging the communication gap [6].

UX design plays a critical role in operationalizing these principles. It serves as the bridge between intricate AI systems and their human users. A well-crafted UX not only renders AI technologies accessible and intuitive but also fosters transparency and understanding. By experiencing AI through a thoughtfully designed interface, users can perceive the fairness and contestability of its decisions.

The UX becomes a crucial channel for demonstrating the practical application of ethical principles. Through meticulous UX design, AI systems can communicate their decision-making rationale, provide avenues for users to question or contest those decisions, and ensure responsible and equitable use of the technology. This emphasis on UX design reassures the audience of the robustness and reliability of AI systems.

**Through meticulous UX design, AI systems can communicate their decision-making rationale, provide avenues for users to question or contest those decisions, and ensure responsible and equitable use of the technology.**



The evolving landscape of GenAI has led to a redefinition of traditional UX paradigms, prompting discussions about a "new design paradigm" [27]. UX researchers are refining best practices to ensure AI products are explainable, transparent, fair, and interpretable.

Examples of this shift include iterative testing with diverse user groups to mitigate bias and enhance accessibility. Additionally, UX interactions with GenAI differ from traditional software systems by prioritizing user involvement throughout the development process, ultimately transforming perceptions of AI from intimidating technology to user-friendly tools that enhance daily life.

### Here are some examples of how UX interactions have changed with GenAI:

- **Shift from static interfaces:** Traditional software has static interfaces, while GenAI systems can adapt and personalize interactions based on user input.
- **Focus on user education:** With GenAI's unpredictable nature, UX focuses on educating users about the system's capabilities and limitations.
- **Iterative feedback loops:** GenAI systems learn and change over time, requiring continuous user feedback loops to refine and improve the experience.
- **Explainable AI interfaces:** UX designs interfaces that explain how GenAI arrives at its outputs, building trust with users.

Ultimately, interface design is pivotal in defining users' relationships with technology, helping them view it as a collaborative resource that enhances our daily lives. As AI evolves and permeates our daily lives, its successful and sustainable integration hinges on embedding fairness, transparency, privacy, veracity, and contestability into every aspect of its development. AI systems' human experience will be at the forefront of this endeavor. It will be the tangible embodiment of these values and the primary interface through which users engage with and trust AI's transformative power.

---

## AI for All? The Technological Access Gap

Beyond skilled AI practitioners, one needs lots of data and extensive computing power to develop frontier AI systems. Sadly, most academic institutions and many private organizations are unable to acquire all the resources required to develop capable LLMs, which explains why the industry has gained immense

influence over AI research [68]. For AI to be truly accessible to all, it's critical to ensure that the majority of researchers and developers can build models suitable for their use case.

## Consequences of LLM Hallucination

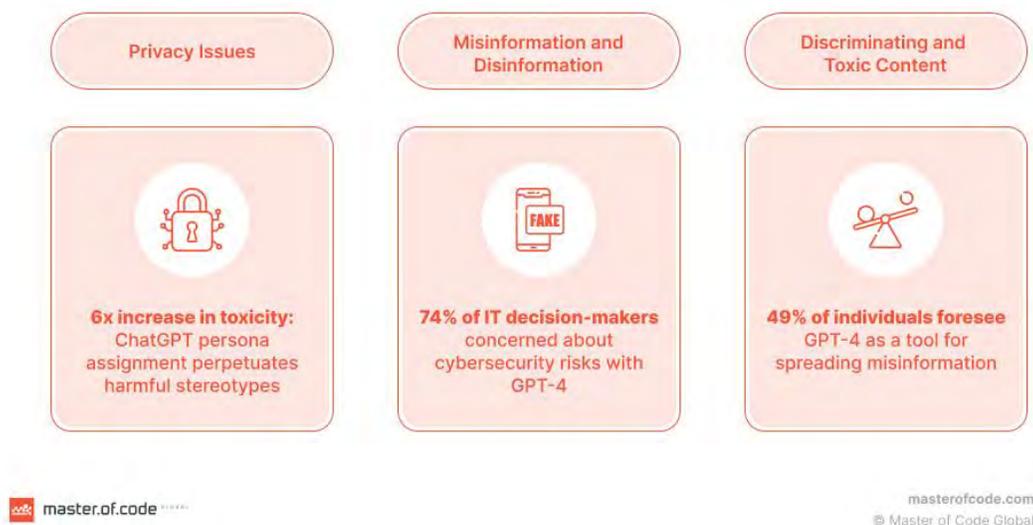
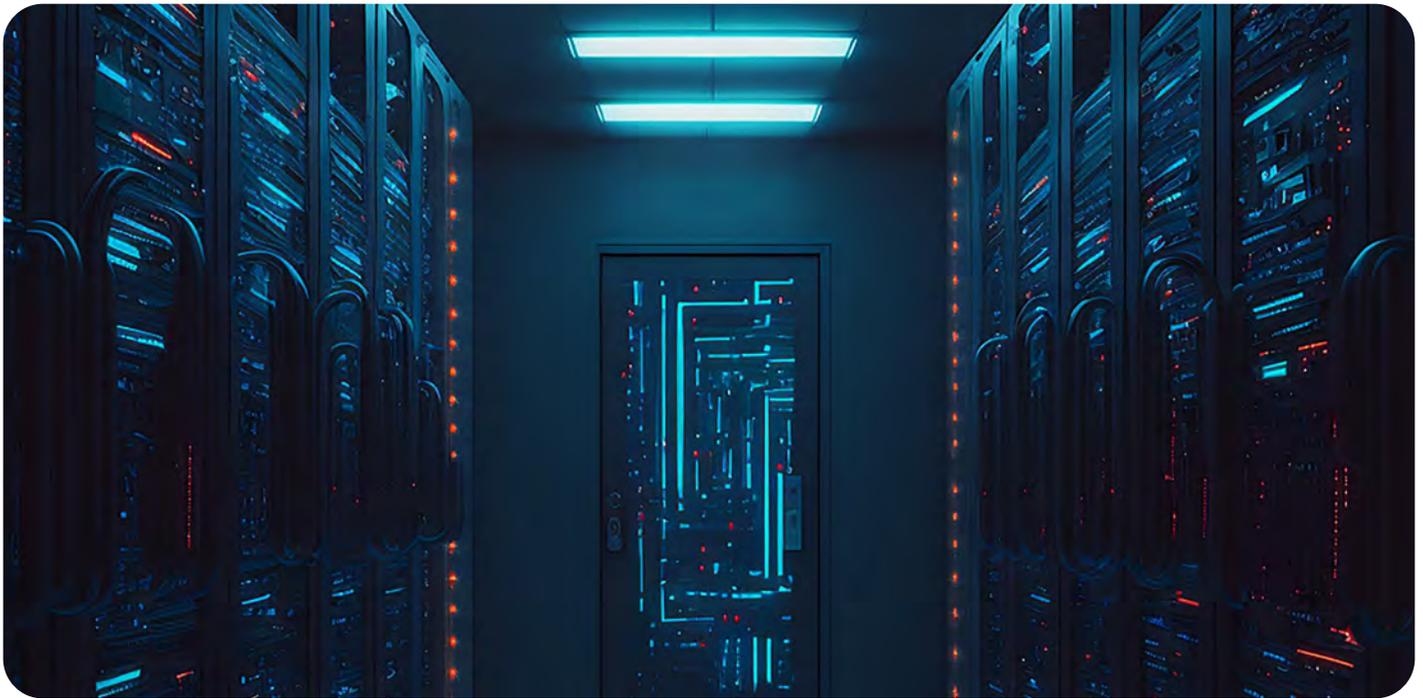


Figure 7: Consequences of LLM hallucination [5].

For an AI system to learn how to form intelligible sentences, it must be exposed to a large quantity of diverse texts. This allows for understanding of which parts of speech are needed and where, how some words don't go with each other, which tone is preferred, etc. For an LLM to be applicable and useful in the many use cases we see nowadays, it must also leverage some knowledge of the world (i.e., common sense, mathematics, programming, etc.) and have the ability to reason on it. This inevitably requires even more data.

Because it is rare for one organization to have all the data necessary to build such highly capable LLMs, it is common practice to leverage large collections of datasets collected from the internet. In particular, the Pile [57], which includes the Books3 dataset [56], was a go-to corpus for the development of LLMs. Since then, a group of authors included in the Books3 dataset sued OpenAI and Microsoft for copyright infringement suits. This legal case as well as the other ongoing copyright infringement suits will likely have an enormous impact on the future of AI development<sup>1</sup>.

<sup>1</sup>At the time of writing this paper, the Books3 case is still ongoing [32], as are others like the New York Times suit against OpenAI and Microsoft [9].



The intricacies of developing AI systems with publicly available content may become clearer after these landmark cases conclude, which may take years. In the meantime, small content creators are at greater risk of having their content used for training without their consent as they may not have registered their work at the copyright office, and copyright litigation requires numerous resources [52]. According to a Vondran Legal blog post, “copyright lawsuits can last months to years and can cost anywhere from \$5,000 (if a quick settlement can be reached) or up to \$350,000 or more for lengthy litigation” [43]. Additionally, the cases that may determine the future of AI will not be quick wins for content creators: those cases are complex, and the opposing AI companies have a lot to lose<sup>2</sup>.

Once enough data is collected, we can start looking at the actual training and the computing power necessary to execute it. Enormous computing power is required for an LLM to learn from the data, i.e. update its billions or trillions of parameters until it “comprehends” how to process the data properly. This phase is usually done with graphics processing units (GPUs), computer chips optimized to perform mathematical operations needed to train AI systems.

---

<sup>2</sup>Furthermore, there is so much to gain with AI that Microsoft committed “to defend these customers and pay for any adverse judgments if they are sued for copyright infringement for the use of the Azure OpenAI Service outputs.” [29].



**Big players like Microsoft and Google hold great influence in the AI space.**

For instance, training Llama2-7B, a leading open-source LLM of reasonable size developed by Meta [64], required no less than 184,320 GPU hours, i.e. 128 days when using exactly one GPU<sup>3</sup> [63]. To do this in less than 30 days, about five GPUs would need to be simultaneously dedicated to this task.

Enabling multi-GPU training usually requires building a cluster—a group of computers with multiple GPUs. This is not as straightforward as buying 5 GPUs: hardware-wise, it also requires CPUs, motherboards, RAM, memory storage, power

supply units, and networking. Moreover, it is necessary to secure the space to host this cluster and the electricity to power it<sup>4</sup>.

Instead of actually owning the GPUs, one could choose to leverage cloud computing services. This option is particularly interesting for startups as the initial investment is much smaller. If we return to our Llama2-7B example, assuming a cost of \$1.50 per hour, the training would roughly cost \$276,480 [25]. But there is a catch—this is only the initial investment, as we also need GPUs to use our model in production, and our count of GPU hours may explode as a result.

All of this highlights the vast amount of resources needed to develop products powered by GenAI and explains why big players like Microsoft and Google hold great influence in the AI space. Both companies have substantially contributed to the AI community through open-source contributions, whether it be libraries, datasets, or model weights.

---

## Evaluating the Impacts of Open vs. Closed AI Development

As AI systems grew more capable, the risks associated with them became more tangible. Although the risks are not new and “real problems like algorithmic discrimination and predictive policing [...] are harming individuals now” [18], AI doomsday scenarios do not feel as fictitious as they once did, and AI incidents may now be part of a regular news cycle. Thus, it can now be frowned upon to release GenAI systems without any restrictions or guardrails in place. Their release can be more safely done in a closed-source environment, where the developers of an AI system can restrict how it is shared, modified, and used. This contrasts with open-source development, where everything is shared freely.

---

<sup>3</sup>This number would be much higher if Meta didn't use the highest-performing GPUs at the time, Nvidia A100-80gb.

<sup>4</sup>See [62] for more information on building a GPU cluster.

## It can now be frowned upon to release GenAI systems without any restrictions or guardrails in place.

Simply put, the more closed-access an AI system is, the more control the provider has over its risks. Being more opaque about what the user is interacting with allows an organization to:

- Minimize inappropriate use of the AI system and negative public attention,
- Continuously improve the AI system by leveraging usage data,
- Protect its intellectual property and minimize the risks of litigation,
- And more.

Nonetheless, by being more closed, this AI system won't be used as much, and some risks could go overlooked, leading to inappropriate behaviors that may harm society. A testing environment prior to deployment often resembles lab conditions, which can drastically differ from real-world usage. Generally, it is the most extreme and unexpected uses that lead to uncovering and hopefully fixing many flaws of generative AI systems [37].

On the other end of the spectrum is open-source, an environment where the model's technical details, such as the model weights, the training recipe and associated datasets, evaluation protocols and results, code, and related documentation, are shared transparently and freely. This approach facilitates the inspection of these AI systems, increases trust, and accelerates innovation by sharing resources and best practices [20]. Fully open AI systems democratize this technology regardless of the user's intent.



**The complexity of jailbreaking attacks—attempts to bypass the safeguards of an AI system and elicit unethical responses—has steadily evolved.**

This could give malicious actors the ability to design sophisticated attacks by leveraging this technology. For instance, the complexity of jailbreaking attacks—attempts to bypass the safeguards of an AI system and elicit unethical responses—has steadily evolved from simple adversarial prompts to conversational and “seemingly benign” interactions [33].

Considerations	internal research only high risk control low auditability limited perspectives					community research low risk control high auditability broader perspectives
Level of Access	fully closed	gradual/staged release	hosted access	cloud-based/API access	downloadable	fully open
System (Developer)	PaLM (Google) Gopher (DeepMind) Imagen (Google) Make-A-Video (Meta)	GPT-2 (OpenAI) Stable Diffusion (Stability AI)	DALLE-2 (OpenAI) Midjourney (Midjourney)	GPT-3 (OpenAI)	OPT (Meta) Craiyon (craiyon)	BLOOM (BigScience) GPT-J (EleutherAI)

Figure 8: The many strategies used to release generative AI systems. Created by Irene Solaiman [36].

The appropriate degree of openness to responsibly release capable AI systems is still up for debate, and various strategies have been deployed. Generally, they all try to minimize risks by striking a balance between privacy and control over the usage of an AI system and transparency and collaboration with diverse stakeholders.

Beyond balancing risks, open AI development enables faster innovation. Instead of spending resources to reproduce the newest developments of AI research, organizations can invest in adapting this technology to their specific use cases.

## The Dangers of an AI Arms Race

When ChatGPT took the world by storm, it didn't take a visionary to conceive how it could be integrated into many applications and industries. At the time, ChatGPT was only a research demo—it was likely not



thoroughly tested like an enterprise AI product and was not intended to support human decision-making in all application contexts imaginable, although it did pass the Uniform Bar Exam [34] and showcased other outstanding capabilities.

**High demand led to what is now being called the “AI arms race,” where many companies feel enormous pressure to support GenAI-powered features sooner rather than later.**

A record number of users tried this new technology and many asked for it to be integrated into their daily life. This high demand led to what is now being called the “AI arms race,” where many companies feel enormous pressure to support GenAI-powered features sooner rather than later. Firms that did not have AI capabilities on par with OpenAI faced the challenge of catching up within a very short timeframe. In light of the difficulty of developing an LLM, many have chosen to leverage a third-party AI model within their product line [40].

This high pressure to release GenAI-powered features not only affected the choice of AI system, but also how it was tested by organizations integrating with this technology. While assessing the safety of an emerging technology is never an easy task, it may not be a priority if the organization perceives an existential risk if they delay their release. Further, when an organization must rush an integration with third-party systems because they don’t have the resources to develop their own AI capabilities, it is likely that they won’t have the resources and expertise required to thoroughly test it internally.

It is important to note that it is not simple to assess how responsible and safe an AI system is. Developing a robust methodology takes time: For example, 1.5 years were required for the NIST AI RMF, one of the leading resources for AI risk management, to officially launch [48]. While extremely thorough, it only alludes to questions regarding the appropriate ways to measure performance or appropriate behavior in a given use case.



# Impact of Generative AI on Industry



GenAI is rapidly becoming integral to both personal and professional realms. With an overwhelming majority of enterprise AI decision-makers (more than 80%) planning to harness GenAI by 2026 for boosting productivity, spurring innovation, and achieving cost efficiency, its influence is undeniable [50]. The surge in interest spans a broad range of industries with a particular focus on data-centric and heavily regulated sectors. Particularly, the banking, financial services, insurance, and healthcare sectors are leading the charge in AI adoption, driven by the profound impact AI can have in areas rich in data and subject to stringent regulatory oversight.

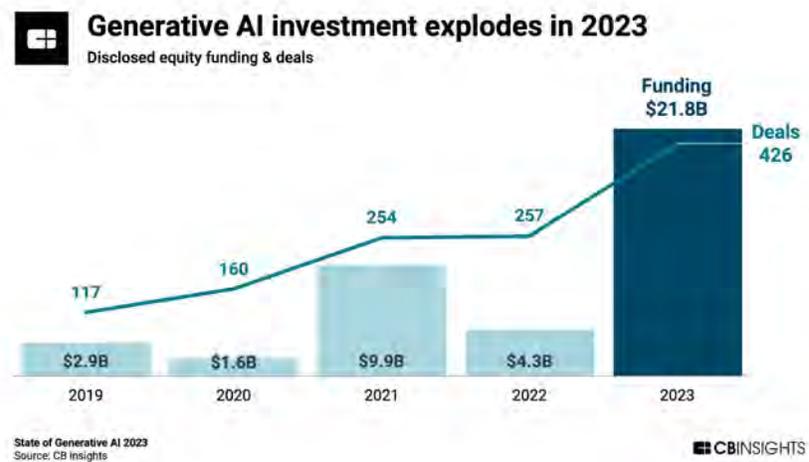


Figure 9: Generative AI investment between 2019-2023 [55].

As organizations explore the GenAI revolution, they confront the critical challenge of seamlessly and responsibly integrating these technologies into their operations. This journey involves developing vigorous strategies to ensure AI's safe and ethical use, especially considering the complexities around model training data and preventing its potential misuse.

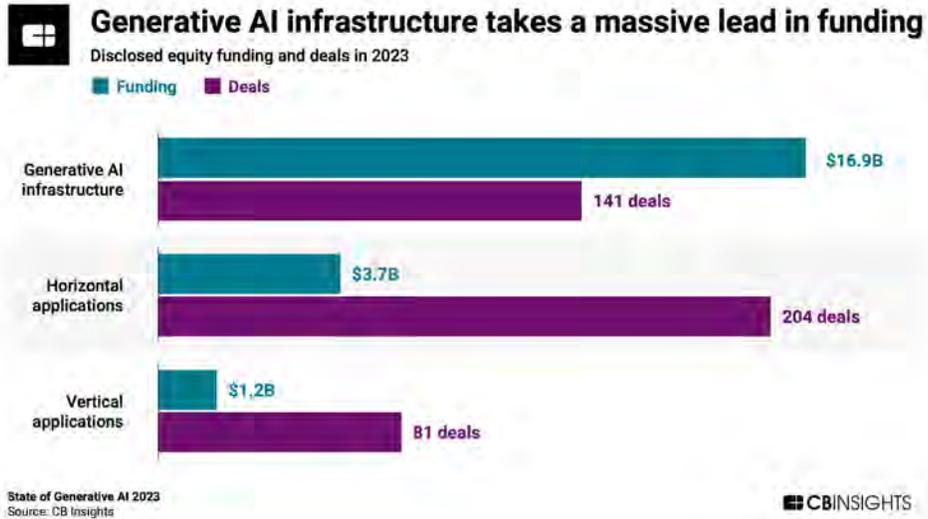


Figure 10: Generative AI disclosed equity funding and deals in 2023 [55].

The AI landscape is evolving rapidly, and businesses must be ready to adapt. GenAI may be a game-changer for boosting efficiency and innovation, but unlocking its full potential requires overcoming integration hurdles and prioritizing ethical considerations.

**Here's what organizations can do:**

- **Build internal expertise:** Invest in training and resources to build and empower teams with knowledge surrounding generative AI capabilities and limitations.
- **Identify use cases:** Explore specific business needs and identify low-risk areas where GenAI can provide business value.
- **Partner with experts:** Collaborate with AI specialists to add a review layer to address potential technical challenges or risks.
- **Develop strong ethical guidelines:** Establish clear principles for responsible AI use, focusing on areas specialized to your business.

**As organizations explore the GenAI revolution, they confront the critical challenge of seamlessly and responsibly integrating these technologies into their operations.**



By taking these steps, businesses can leverage the power of GenAI while increasing its responsible and ethical implementation.

## Opportunities for Innovation and Disruption

Generative AI is rapidly transforming industries by streamlining operations and catalyzing innovation, profoundly impacting cost reduction, process efficiency, growth, and new discoveries. Presently, 38% of IT professionals report active GenAI implementations in their enterprises, with another 42% evaluating its potential, indicating a robust trend toward widespread adoption [53].

GenAI has the potential to significantly reduce operational costs by automating routine job functions and minimizing the need for manual interventions. This automation is crucial in areas like human resources (HR) and customer support, where AI-driven chatbots provide 24/7 service, and in finance, where AI enhances fraud detection and risk management, streamlining processes and safeguarding assets.

Furthermore, this technology accelerates innovation by synthesizing and analyzing vast datasets faster than ever, leading to new insights and creativity across various fields. In project management and business operations, AI's capability to automate and predict enhances decision-making and efficiency, illustrating the broad applicability of GenAI in reshaping industry standards and operational paradigms.

The innovations afforded by GenAI can enhance operational efficiencies and open up new avenues for growth and discovery, propelling industries toward a more data-driven, efficient, and personalized future. As enterprises continue integrating GenAI into their core operations, the distinction between traditional AI and generative AI blurs, creating a dynamic landscape where AI's potential is continuously expanded and refined.

By automating routine and time-consuming tasks, GenAI frees up human capital to explore new ideas and growth areas essential for driving innovation and maintaining a competitive edge. But to do so ethically requires that companies leverage GenAI responsibly and effectively, maximizing benefits while ensuring trust and reliability, with the caveat that human oversight remains in control and is a part of the entire AI lifecycle loop, not just at the beginning of deployment.

**[AI] technology accelerates innovation by synthesizing and analyzing vast datasets faster than ever, leading to new insights and creativity across various fields.**



Incorporating human judgment mitigates risks and significantly enhances the contributions of GenAI, fostering a synergistic relationship where humans and machines collaborate towards more intelligent, sustainable, and innovative outcomes.

**But to do so ethically requires that companies leverage GenAI responsibly and effectively, maximizing benefits while ensuring trust and reliability, with the caveat that human oversight remains in control and is a part of the entire AI lifecycle loop, not just at the beginning of deployment.**

## The Challenges Experienced by Industries

As GenAI continues to integrate deeply into business processes and consumer applications, it confronts several unique challenges that could potentially stymie its efficacy and ethical adoption. These challenges revolve around data security, data bias, and the risk of generating inaccurate outputs—each presenting significant hurdles for organizations leveraging this technology.

**The challenge extends beyond merely securing data from external threats; it also involves ensuring that the data cannot be exploited maliciously within GenAI processes.**

One of the paramount concerns with GenAI revolves around data security. The potential for breaches, unauthorized access, or misuse of personal information remains a top concern for product teams, customers, and end users. Organizations prioritize safeguarding data as GenAI systems often require

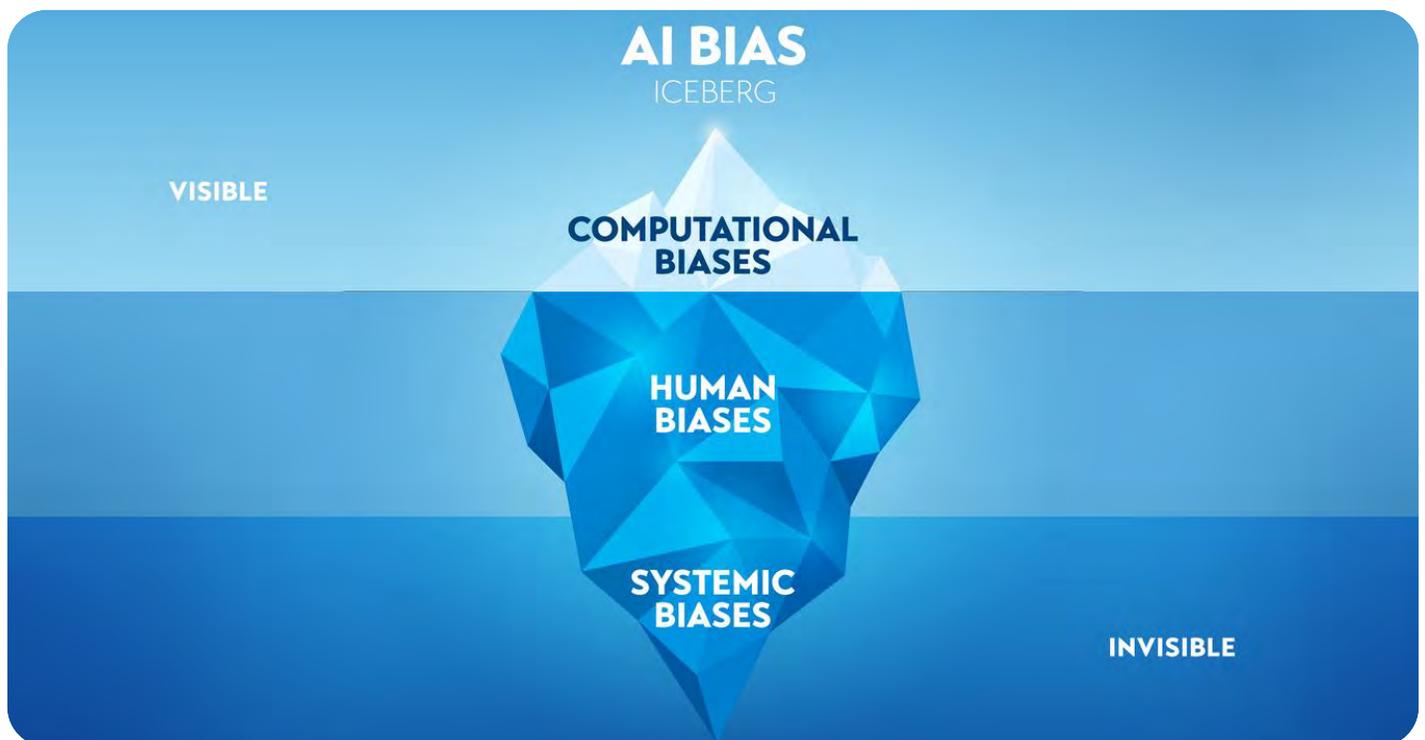


access to vast amounts of information to learn and make predictions. This extensive data access increases the vulnerability points for security breaches, which could lead to severe privacy violations and financial losses. The challenge extends beyond merely securing data from external threats; it also involves ensuring that the data cannot be exploited maliciously within GenAI processes, maintaining robust compliance with global data protection regulations (such as GDPR [61] and CCPA [47]), and implementing stringent access controls and audit trails to monitor data use within AI systems.

**Data bias is an inherent problem that cannot be eliminated but must be continuously mitigated and acknowledged.**

Another critical issue is data bias, particularly the underrepresentation of minority groups in AI algorithms. Data bias is an inherent problem that cannot be eliminated but must be continuously mitigated and acknowledged. An example of this challenge was observed with Google's image generation technology, which initially overcorrected racial representation by skewing predictive modeling in an unintended direction [24].

The task then becomes assessing the biases in the data used to train GenAI models and implementing strategies to counteract these biases effectively. This involves enhancing the diversity of data sources, applying robust pre-processing checks, and continuously testing the outputs for fairness across different demographic groups.





Inaccurate outputs from GenAI can range from minor errors in data interpretation to significant blunders that might skew business decisions or misinform users. These inaccuracies stem from various factors, including flawed data inputs, inherent biases, or model overfitting. In scenarios where AI-generated content or decisions have significant real-world impacts—such as in legal advice, medical diagnosis, or financial forecasting—the consequences of inaccuracies can be particularly severe. Addressing this challenge requires rigorous validation and testing of AI models, the development of sophisticated error-checking algorithms, and a fallback plan for human intervention when AI outputs fall within critical thresholds of doubt.

**As GenAI continues to evolve, the focus must increasingly shift towards enhancing AI systems' transparency, accountability, and ethical governance.**

These examples of AI-specific risks highlight the complex, multifaceted challenges of deploying generative AI technologies. They underscore the necessity for a comprehensive framework that anticipates these risks and prepares robust mitigation strategies to address them effectively. As GenAI continues to evolve, the focus must increasingly shift towards enhancing AI systems' transparency, accountability, and ethical governance to foster trust and reliability among all stakeholders involved. This approach will ensure that as organizations harness the powerful capabilities of GenAI, they also uphold the highest standards of integrity and fairness.



# Changes Since the Release of ChatGPT



## Shifts in Public Perception and Acceptance of AI Technology

The release of ChatGPT has fundamentally shifted public perception and acceptance of AI technology. Before ChatGPT, interaction with AI was primarily confined to specialized practitioners and researchers, with the general public viewing it as a niche, technical domain. However, the unprecedented adoption and widespread applicability of ChatGPT have brought AI into the mainstream, capturing the public's imagination and sparking broader engagement.

**The unprecedented adoption and widespread applicability of ChatGPT have brought AI into the mainstream, capturing the public's imagination and sparking broader engagement.**

The rapid integration of ChatGPT and similar models has also raised significant concerns among the public, policymakers, and experts. Issues surrounding the potential for misinformation, bias, privacy violations, and the impact on jobs have prompted a more cautious and critical examination of the societal implications of these technologies. This has fostered a more nuanced public discourse, where the excitement over AI's capabilities is balanced by recognizing the need for robust safeguards and responsible development.

As the public grapples with GenAI's dual-edged nature, the acceptance and perception of this technology will likely continue to evolve. Maintaining transparency, addressing ethical concerns, and demonstrating AI's tangible benefits will be crucial in shaping public trust and ensuring its sustainable integration into various aspects of society.

## Regulatory Developments and Policy Implications

The rapid advancements and widespread adoption of GenAI have prompted regulatory bodies worldwide to consider the implications of these transformative technologies and explore frameworks for their governance. In response to growing concerns about data privacy, intellectual property, and ethical standards, several countries and international organizations have begun drafting guidelines and regulations to ensure the responsible development and deployment of AI.

**There is a growing consensus that the status quo is insufficient to manage the risks and harness the benefits of GenAI.**

From the European Union's proposed AI Act [60] to the ongoing discussions around AI-specific regulations in the United States [38], Canada [54], and other nations, there is a growing consensus that the status quo is insufficient to manage the risks and harness the benefits of GenAI. These regulatory initiatives have focused on key areas such as ensuring algorithmic transparency and accountability, mitigating biases and discriminatory practices, protecting individual privacy and data rights, addressing the potential for malicious use and the spread of misinformation, maintaining human agency and oversight in AI-driven decision-making, and evaluating the impact of AI on employment and the future of work.



### EU AI Act in a Nutshell

- The EU AI Act proposes a legal framework for mitigating risks in AI technologies.
- It classifies AI into categories: unacceptable, high, limited, and minimal risk.
- High-risk AI must adhere to strict safety and nondiscrimination standards.
- The Act requires transparency for AI that interacts with individuals.
- Generative AI falls under the broader scope, addressed by risk potential.
- Compliance is overseen by national authorities and the European AI Board.
- The Act aims to balance innovation with the protection of rights and values.

Figure 11: The EU AI Act [60].



These regulatory frameworks will have far-reaching implications for the development and commercialization of GenAI systems. Organizations and developers must adhere to stringent guidelines, undergo comprehensive testing and auditing, and demonstrate their commitment to responsible AI practices. Moreover, the policy landscape will shape how GenAI can be integrated into various industries and applications, as policymakers must strike a delicate balance between fostering innovation and safeguarding the public interest.

However, there is a risk that over-regulation and regulatory capture could curtail innovation, particularly for small companies, startups, and open-source developers. Stricter compliance requirements and the resources needed to develop and deploy generative AI systems in line with regulations may erect barriers to entry, concentrating the most powerful models in the hands of a few large tech giants. This could stifle the diversity and creativity that has characterized the current generative AI landscape, limiting the potential for disruptive innovation and equitable access to these transformative technologies [39, 51].

**A balanced approach that fosters both innovation and responsible practices will be crucial in shaping the future of GenAI and ensuring its benefits are widely shared.**

Policymakers must therefore approach regulation with care, ensuring that the rules they establish do not inadvertently cement the dominance of incumbents or hamper the ability of smaller players to bring their ideas and solutions to the market. A balanced approach that fosters both innovation and responsible practices will be crucial in shaping the future of GenAI and ensuring its benefits are widely shared.

---

## Evolution of User Interactions with Generative AI Platforms

The emergence of GenAI platforms has ushered in a profound shift in how users interact with technology. The accessibility and human-like conversational abilities of ChatGPT and similar systems have democratized the user experience of AI, inviting a broader audience to engage with it in their daily lives.



One of the key changes has been the seamless integration of these AI platforms into everyday tasks and workflows. Users now turn to GenAI for a wide range of applications, from creative writing and research assistance to task automation and problem-solving. This level of ubiquity was previously unimaginable, as AI was seen as a niche tool primarily used by developers and researchers.

Moreover, the interactive and conversational nature of GenAI platforms has fostered a sense of familiarity and trust among users. Individuals are increasingly comfortable engaging in open-ended dialogues with these AI assistants, treating them as collaborative partners in the exploration of ideas, the generation of content, and the decision-making process. This shift in user perception, from viewing AI as a rigid tool to perceiving it as an intelligent companion, has profound implications for how these technologies are integrated into various domains [23].

However, this growing reliance on GenAI has also raised concerns about over-dependence, the potential for biased or inaccurate outputs, and the blurring of boundaries between human and machine intelligence. As users become more immersed in these AI-driven interactions, there is a pressing need to ensure the level of trust and engagement remains balanced and appropriate, with clear delineations between the capabilities and limitations of these systems.

**The interactive and conversational nature of GenAI platforms has fostered a sense of familiarity and trust among users.**



# Predictions for the Future of AI

## Unexplored Applications Beyond Chatbots

GenAI is rapidly advancing, poised to reshape our world with its untapped potential and innovative applications. Beyond merely enhancing text-based interactions, some anticipate GenAI will revolutionize a myriad of industries and societal norms, primarily through its integration with spatial computing. This confluence of technologies offers a glimpse into a future where AI's influence extends into uncharted territories, redefining what's possible in personalized learning, scientific discovery, immersive virtual reality, interactive holography, and even in creating AI-driven simulations that offer new ways to solve complex problems and communicate.

Personalized learning experiences, tailored by AI tutors to match individual learning styles, promise to revolutionize education, making it more adaptive and effective. In the scientific realm, GenAI's ability to generate hypotheses, design experiments, and analyze data will accelerate discoveries, with spatial computing enabling scientists to interact with complex 3D molecular models, deepening our understanding of the material world. The drug discovery field also benefits immensely, as AI can simulate molecular interactions to identify new therapeutic candidates, further enhanced by spatial computing's ability to visualize these interactions in three dimensions.



Image by [Freepik](#).



**Moving beyond traditional chat capabilities, generative design emerges as a new frontier for AI, enabling the creation of optimized products, buildings, and cities in real-time 3D environments.**

Moving beyond traditional chat capabilities, generative design emerges as a new frontier for AI, enabling the creation of optimized products, buildings, and cities in real-time 3D environments. These innovations will revolutionize architectural and product design processes. In other industries, AI is poised to enable new forms of personalized content creation tools, paving the way for interactive AI companions that provide emotional support and real-time assistance in users' spatial computing environments.

**The future of GenAI, particularly in conjunction with spatial computing, holds boundless possibilities. From revolutionizing personal learning to reshaping urban landscapes, AI advancements will augment human capabilities and open new dimensions for creativity, innovation, and efficiency.**

The future of GenAI, particularly in conjunction with spatial computing [12], holds boundless possibilities. From revolutionizing personal learning to reshaping urban landscapes, AI advancements will augment human capabilities and open new dimensions for creativity, innovation, and efficiency. As we venture into this future, how we perceive AI in our lives and workplaces will undergo profound transformations, marking a significant leap toward a more interconnected and augmented human experience.



# The Next Steps in the Development of Generative AI

The current pace of AI innovation is extraordinary. According to [13], an analysis of more than 200 evaluations of language models from 2012 to 2023 indicates that the computational power needed to achieve a specified level of performance reduces by half roughly every five to fourteen months. In other words, LLMs' capabilities have been growing several times faster than hardware gains per Moore's Law. This level of disruption makes it challenging to predict GenAI's path in the near future. However, unless GenAI evolution is poised to hit a wall soon [19], we can safely assume that these models will rapidly improve [8].

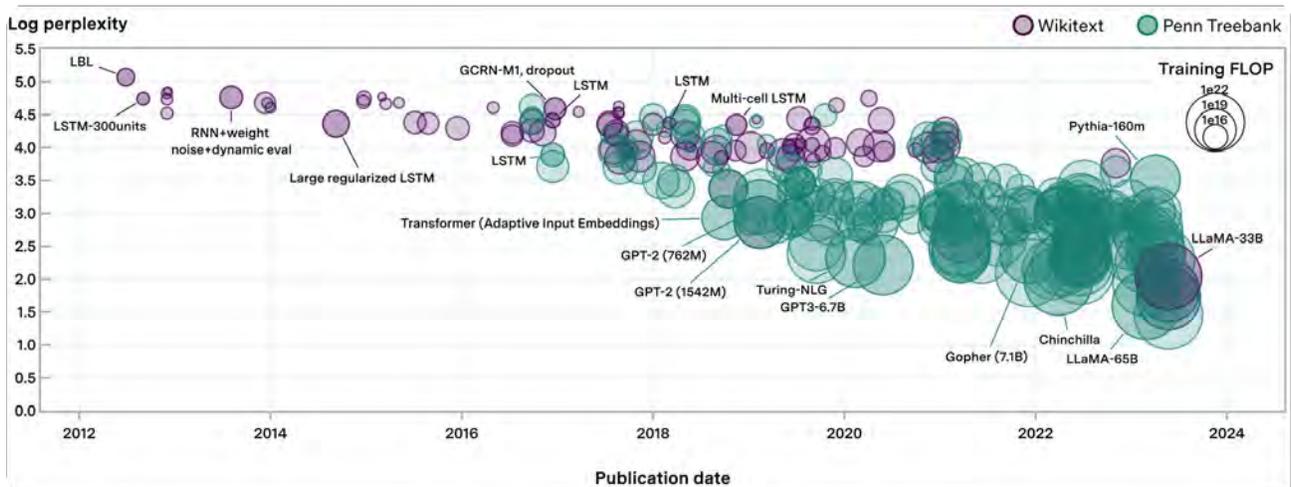


Figure 12: Log of perplexity results show exponential improvement of LLM models over time. Each shape represents a model and the size of the shape is proportional to the compute used during training [13].

The highly anticipated release of GPT-5 by OpenAI, later this year in 2024, will possibly catalyze transformative changes across diverse segments of society by setting a new standard for GenAI capabilities. This leap forward is expected to usher in a new era of AI-powered innovation, dramatically expanding the potential applications and capabilities of GenAI systems. Concurrently, advancements in video generation are expected to continue at a rapid pace in 2024, blurring the lines between AI-generated and real-life imagery. Enhanced realism and interactive capabilities in AI-generated videos could significantly impact marketing strategies, film production, and virtual reality experiences, offering users unprecedented immersive environments.



**Coupled with this, the emergence of AI agents signifies a groundbreaking shift in the workplace. AI-driven counterparts specifically designed to meet the distinct needs of various professional roles are set to transform the workforce.**

Coupled with this, the emergence of AI agents signifies a groundbreaking shift in the workplace. AI-driven counterparts specifically designed to meet the distinct needs of various professional roles are set to transform the workforce. For instance, companies like Cognition and Magic are redefining the role of the software engineer with AI agents guided by natural language prompts, enabling automated code generation and issue resolution, thereby accelerating development cycles and enhancing productivity. Further, capable web agents have the potential to rid us of menial tasks and substantially change how we work. These advancements are anticipated to reshape many other jobs such as financial analyst, product manager, and customer-facing roles, indicating a shift toward more collaborative and intuitive human-machine interactions or, in some cases, the replacement of humans for machines.

Visionaries are also betting on GenAI agents to transform how we interact with our personal devices. Innovations such as Rabbit's R1, Humane's Ai Pin, and Open Interpreter's O1 Light are at the forefront of this transformation, each offering unique interfaces and capabilities designed to integrate seamlessly into daily life. These AI assistants aim to provide personalized support as intelligent companions that can help with tasks ranging from scheduling to research. With natural language interfaces and contextual awareness, they could redefine how we navigate the digital world and leverage AI's potential on our most personal computing devices.

**As AI systems become more integrated into our daily lives, ensuring they can understand and interact appropriately with the complexities of the real world becomes increasingly critical.**

However, these advancements come with their own set of challenges, including mitigating hallucinations, improving reasoning capabilities, and developing comprehensive world models. As AI systems become more integrated into our daily lives, ensuring they can understand and interact appropriately with the complexities of the real world becomes increasingly critical.

## Anticipated Technological Developments

Despite tremendous improvements in AI capabilities over the last few years, it is clear that even the most advanced models do not comprehend the world like humans do. GenAI's reasoning is still far from human-like, often resulting in responses that, while technically correct, may lack depth or understanding of context. To address this discrepancy and enhance AI's cognitive faculties, some researchers believe that a fundamental shift in AI architecture is necessary. Meta's Chief AI Scientist Yann LeCun, for instance, has proposed a move towards systems with predictive capabilities that can learn world models—comprehensive simulations of reality allowing AI to anticipate outcomes and understand causality [49].

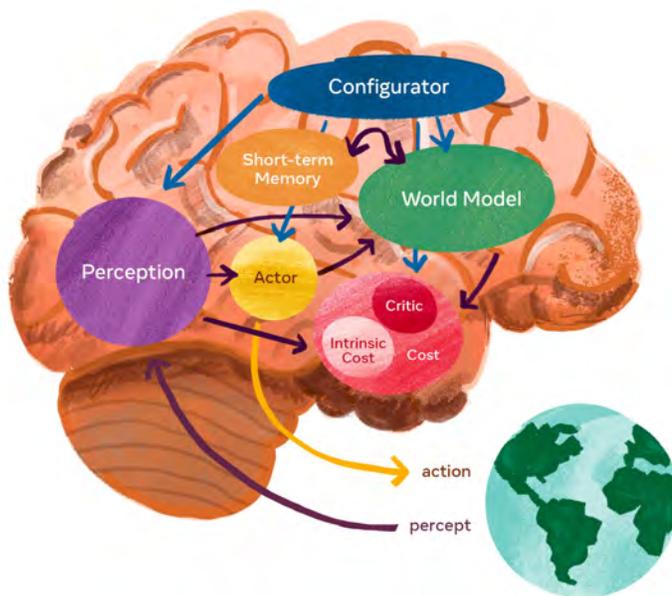


Figure 13: A simplified diagram of Yann LeCun's system architecture for autonomous intelligence [49].

Others believe that breakthroughs in the integration of GenAI and robotics may be sufficient to reach the next level, asserting that these advances are expected to enhance AI's understanding of the world [67]. By situating AI within embodied agents that can interact with their environment, we may enable AI models to acquire a more nuanced understanding of physical spaces and social dynamics, thus improving their reasoning and learning. These developments will not only expand the applications of AI in our daily lives but also refine the interaction between humans and machines, paving the way for a future where AI becomes a more integrated and beneficial part of society.

**GenAI's reasoning is still far from human-like, often resulting in responses that, while technically correct, may lack depth or understanding of context. To address this discrepancy and enhance AI's cognitive faculties, some researchers believe that a fundamental shift in AI architecture is necessary.**

Along with contextual understanding, another roadblock to AI's seamless integration into everyday life includes the enormous computing resources required to train and run the most advanced models. Despite impressive strides, GenAI still demands significant computational power, which can limit scalability and accessibility, especially in less-developed regions or among smaller businesses. This challenge is being met head-on with advances in algorithm efficiency and the advent of new chip architectures that promise to deliver more computing power with less energy consumption. As these technological developments progress, they hold the promise of making AI more sustainable and widely available, thereby narrowing the digital divide and fostering a more inclusive technological future.

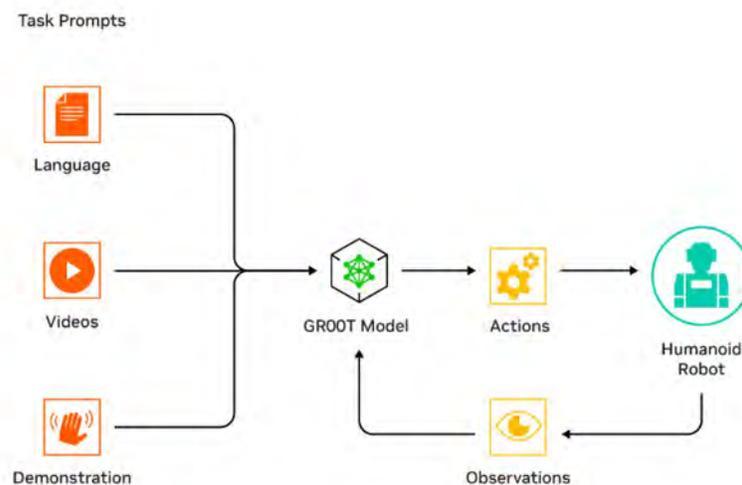


Figure 14: GROOT model training workflow [67].



## Consequences for the Future of Work

As AI capabilities rapidly advance, there is much debate about how it will impact the future of work and employment. Many believe that AI will inevitably replace most human jobs, leading to widespread unemployment and economic disruption. However, others argue that this viewpoint is overly simplistic and fails to account for the nuances of how AI and humans can coexist and complement each other in the workforce.

While AI may indeed automate or replace some specific tasks currently performed by humans, it does not necessarily mean whole jobs will disappear. As economist David Autor argues, AI could actually aid in revitalizing the middle class rather than eliminating jobs entirely. By extending the scope and value of human expertise, AI can enable a wider set of workers to perform higher-stakes decision-making tasks currently reserved for highly educated professionals. This could potentially reverse the trend of the last few decades where decision-making and resources became increasingly concentrated among elite experts, leading to a "hollowing out" of the middle class due to automation and globalization [3].

**While AI may indeed automate or replace some specific tasks currently performed by humans, it does not necessarily mean whole jobs will disappear.**

Similarly, economist Noah Smith offers another counter-narrative to the fear of AI-induced job scarcity. He argues that even if AI becomes capable of outperforming humans at virtually every task, it does not necessarily mean the end of human jobs. Drawing on the economic principle of comparative advantage, Smith explains that value and demand for labor could still increase in an AI-dominant world due to computing constraints. As AI grows the economy (at potentially staggering rates) the resultant wealthy society would still necessitate human roles in areas where they have a comparative advantage—tasks they can do at a lower opportunity cost, even if AI has an absolute advantage in all tasks. This could lead to a rich and prosperous society where there's ample demand for human work, resulting in high-paying jobs and an economic structure where humans are employed not in spite of, but alongside advanced AI systems [35].



Regardless of whether AI leads to soaring unemployment or ushers in new economic frontiers, the nature of work will evolve significantly in the coming decades. Supporting this transformation will likely require overhauling outdated infrastructure, regulations, and social policies to match the new realities of an AI-driven economy and workforce. Educational institutions must pivot towards curricula that foster skills complementary to AI, such as curiosity, creativity, critical thinking, and complex problem-solving. Labor markets and social policies will need to adapt, providing support for continuous learning and career transitions. Additionally, economic frameworks must be reconsidered to account for the redistribution of tasks between humans and AI, ensuring fair compensation and job satisfaction for the human workforce.

While the specifics remain uncertain, AI's impact on employment will catalyze foundational changes to how we work, learn, and derive economic security in the future. The ultimate outcome will depend on the choices we make today, as we design the systems and structures that will underpin the workplace of tomorrow.



Conclusion



The evolution of GenAI over the past 1.5 years has been nothing short of transformative, catalyzing groundbreaking advancements that have rewritten the boundaries of what is possible. From the unprecedented adoption of ChatGPT to the rapid proliferation of AI-powered innovations across diverse industries, this technology has emerged as a pivotal force reshaping our world.

Our examination has unveiled both the luminous potential and the intricate challenges associated with GenAI. While this technology presents unprecedented opportunities for innovation, productivity, and creative expression, it has also exposed new risks and ethical dilemmas that demand our utmost attention. The issues of bias, privacy, intellectual property, and AI-driven manipulation of digital media are real and pressing, underscoring the critical need for a collaborative, multistakeholder approach to responsible AI development and deployment.

The trajectory of GenAI appears poised for even more extraordinary leaps into our lives, from integrating AI agents into our devices to revolutionizing scientific discovery and education. However, this progress must be anchored in a steadfast commitment to responsible practices. Only through a human-centric approach that places ethical considerations at the core of AI development can we ensure these technological advancements truly serve to uplift and empower society.

**Let us continue to engage in open dialogue, foster collaborative innovation, and collectively shape a future where AI augments and enhances the human experience rather than replacing or endangering it.**

The road ahead is not without obstacles, but the collective efforts of policymakers, researchers, industry leaders, and the public hold the key to harnessing the full potential of GenAI. Let us continue to engage in open dialogue, foster collaborative innovation, and collectively shape a future where AI augments and enhances the human experience rather than replacing or endangering it. By embracing the challenges and seizing the opportunities, we can usher in a new era of technological symbiosis that propels us toward a more equitable, innovative, and sustainable world.

# References

- [1] Saranya A. and Subhashini R. 2023. A systematic review of Explainable Artificial Intelligence models and applications: Recent developments and future trends. *Decision Analytics Journal* 7, (June 2023), 100230. <https://doi.org/10.1016/j.dajour.2023.100230>
- [2] Kars Alfrink, Ianus Keller, Gerd Kortuem, and Neelke Doorn. 2022. Contestable AI by Design: Towards a Framework. *Minds & Machines* 33, 4 (August 2022), 613–639. <https://doi.org/10.1007/s11023-022-09611-z>
- [3] David Autor. 2024. AI Could Actually Help Rebuild The Middle Class. *Noema* (February 2024). Retrieved April 18, 2024 from <https://www.noemamag.com/how-ai-could-help-rebuild-the-middle-class>
- [4] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, March 03, 2021. ACM, Virtual Event Canada, 610–623. <https://doi.org/10.1145/3442188.3445922>
- [5] Maryna Bilan. 2023. LLM Hallucinations: What You Need to Know Before Integration. *Master of Code Global*. Retrieved April 24, 2024 from <https://masterofcode.com/blog/hallucinations-in-llms-what-you-need-to-know-before-integration>
- [6] Ayça Çakmaklı. 2024. Responsible AI at Google Research: User Experience Team. *Google Research*. Retrieved April 29, 2024 from <http://research.google/blog/responsible-ai-at-google-research-user-experience-team/>
- [7] Simon Chesterman. 2024. Good models borrow, great models steal: intellectual property rights and generative AI. *Policy and Society* (February 2024), puae006. <https://doi.org/10.1093/polsoc/puae006>
- [8] Katja Grace, Harlan Stewart, Julia Fabienne Sandkühler, Stephen Thomas, Ben Weinstein-Raun, and Jan Brauner. 2024. Thousands of AI Authors on the Future of AI. Retrieved April 18, 2024 from <http://arxiv.org/abs/2401.02843>
- [9] Michael M. Grynbaum and Ryan Mac. 2023. The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work. *The New York Times*. Retrieved April 18, 2024 from <https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html>



## References

- [10] David Gunning and David W. Aha. 2019. DARPA's Explainable Artificial Intelligence Program. *AI Magazine* 40, 2 (June 2019), 44–58. <https://doi.org/10.1609/aimag.v40i2.2850>
- [11] Hyeonmin Ha, Jihye Lee, Wookje Han, and Byung-Gon Chun. 2023. Meta-Learning of Prompt Generation for Lightweight Prompt Engineering on Language-Model-as-a-Service. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, 2023. Association for Computational Linguistics, Singapore, 2433–2445. <https://doi.org/10.18653/v1/2023.findings-emnlp.159>
- [12] Cathy Hackl. 2023. AI Is A Critical Building Block For Spatial Computing. *Forbes*. Retrieved April 18, 2024 from <https://www.forbes.com/sites/cathyhackl/2023/11/29/ai-is-a-critical-building-block-for-spatial-computing/>
- [13] Anson Ho, Tamay Besiroglu, Ege Erdil, David Owen, Robi Rahman, Zifan Carl Guo, David Atkinson, Neil Thompson, and Jaime Sevilla. 2024. Algorithmic progress in language models. Retrieved April 18, 2024 from <http://arxiv.org/abs/2403.05812>
- [14] Jim Holloway, Milton Cheng, and Julia S. Dickenson. 2024. Will copyright law enable or inhibit generative AI? *World Economic Forum*. Retrieved April 18, 2024 from <https://www.weforum.org/agenda/2024/01/cracking-the-code-generative-ai-and-intellectual-property/>
- [15] Krystal Hu. 2023. ChatGPT sets record for fastest-growing user base - analyst note. *Reuters*. Retrieved April 18, 2024 from <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>
- [16] Pablo Junco. 2023. The Power Of Fine-Tuning In Generative AI. *Forbes*. Retrieved April 10, 2024 from <https://www.forbes.com/sites/forbestechcouncil/2023/10/10/the-power-of-fine-tuning-in-generative-ai/?sh=7aa09dc4adfa>
- [17] Zixuan Ke, Haowei Lin, Yijia Shao, Hu Xu, Lei Shu, and Bing Liu. 2022. Continual Training of Language Models for Few-Shot Learning. Retrieved April 10, 2024 from <http://arxiv.org/abs/2210.05549>

# References

- [18] **Sasha Luccioni. 2023.** The Call to Halt 'Dangerous' AI Research Ignores a Simple Truth. *Wired*. Retrieved April 18, 2024 from <https://www.wired.com/story/the-call-to-halt-dangerous-ai-research-ignores-a-simple-truth/>
- [19] **Thomas Macaulay. 2024.** Meta's AI chief: LLMs will never reach human-level intelligence. *TNW | Deep-Tech*. Retrieved April 18, 2024 from <https://thenextweb.com/news/meta-yann-lecun-ai-behind-human-intelligence>
- [20] **Adir Mancebo Junior. 2023.** With Arms Wide Open (or Not): Navigating Open vs. Closed Development of Powerful AI Models. *Data Science Alliance*. Retrieved April 18, 2024 from <https://www.datasciencealliance.org/work/opensource-vs-closedsource-article>
- [21] **Gary Marcus. 2023.** Why Are We Letting the AI Crisis Just Happen? *The Atlantic*. Retrieved April 18, 2024 from <https://www.theatlantic.com/technology/archive/2023/03/ai-chatbots-large-language-model-misinformation/673376/>
- [22] **MIT Technology Review Insights.** Out with the old and in with the "new IT." Retrieved April 10, 2024 from <https://www.technologyreview.com/2023/02/08/1067218/out-with-the-old-and-in-with-the-new-it/>
- [23] **Ethan Mollick. 2024.** *Co-intelligence: living and working with AI*. Portfolio/Penguin, New York.
- [24] **Megan Morrone. 2024.** Google pauses AI image generation after diversity controversies. *Axios*. Retrieved April 18, 2024 from <https://www.axios.com/2024/02/23/google-gemini-images-stereotypes-controversy>
- [25] **Moin Nadeem. 2023.** Twitter/X Post. *Twitter*. Retrieved April 18, 2024 from <https://twitter.com/moinnadeem/status/1681371166999707648>
- [26] **Xuan-Phi Nguyen, Sharifah Mahani Aljunied, Shafiq Joty, and Lidong Bing. 2023.** Democratizing LLMs for Low-Resource Languages by Leveraging their English Dominant Abilities with Linguistically-Diverse Prompts. Retrieved April 10, 2024 from <http://arxiv.org/abs/2306.11372>



## References

[27] Jakob Nielsen. 2023. AI: First New UI Paradigm in 60 Years. Nielsen Norman Group. Retrieved April 18, 2024 from <https://www.nngroup.com/articles/ai-paradigm/>

[28] Claudio Novelli, Federico Casolari, Philipp Hacker, Giorgio Spedicato, and Luciano Floridi. 2024. Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity. Retrieved April 18, 2024 from <http://arxiv.org/abs/2401.07348>

[29] Brad Smith Nowbar Hossein. 2023. Microsoft announces new Copilot Copyright Commitment for customers. Microsoft On the Issues. Retrieved April 18, 2024 from <https://blogs.microsoft.com/on-the-issues/2023/09/07/copilot-copyright-commitment-ai-legal-concerns/>

[30] Krithika Ramesh, Sunayana Sitaram, and Monojit Choudhury. 2023. Fairness in Language Models Beyond English: Gaps and Challenges. Retrieved April 10, 2024 from <http://arxiv.org/abs/2302.12578>

[31] Alexander Ratner. 2023. Why The Future Of Generative AI Lies In A Company's Own Data. Forbes. Retrieved April 18, 2024 from <https://www.forbes.com/sites/forbestechcouncil/2023/10/17/why-the-future-of-generative-ai-lies-in-a-companys-own-data/>

[32] Alex Reisner. 2023. These 183,000 Books Are Fueling the Biggest Fight in Publishing and Tech. The Atlantic. Retrieved April 18, 2024 from <https://www.theatlantic.com/technology/archive/2023/09/books3-database-generative-ai-training-copyright-infringement/675363/>

[33] Mark Russinovich, Ahmed Salem, and Ronen Eldan. 2024. Great, Now Write an Article About That: The Crescendo Multi-Turn LLM Jailbreak Attack. Retrieved April 18, 2024 from <http://arxiv.org/abs/2404.01833>

[34] Karen Sloan. 2023. Bar exam score shows AI can keep up with "human lawyers," researchers say. Reuters. Retrieved April 18, 2024 from <https://www.reuters.com/technology/bar-exam-score-shows-ai-can-keep-up-with-human-lawyers-researchers-say-2023-03-15/>



## References

[35] Noah Smith. 2023. Plentiful, high-paying jobs in the age of AI. Noahpinion. Retrieved April 18, 2024 from <https://www.noahpinion.blog/p/plentiful-high-paying-jobs-in-the>

[36] Irene Solaiman. 2023. The Gradient of Generative AI Release: Methods and Considerations. Retrieved April 18, 2024 from <http://arxiv.org/abs/2302.04844>

[37] Victor Storchan, Ravin Kumar, Rumman Chowdhury, Seraphina Goldfarb-Tarrant, and Sven Cattell. 2024. Generative AI Red Teaming Transparency Report. Retrieved April 18, 2024 from [https://drive.google.com/file/d/1JqpbIP6DNomkb32umLoiEPombK2-ORc-/view?pli=1&usp=embed\\_facebook](https://drive.google.com/file/d/1JqpbIP6DNomkb32umLoiEPombK2-ORc-/view?pli=1&usp=embed_facebook)

[38] The White House. 2023. Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. The White House. Retrieved April 18, 2024 from <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/>

[39] Adam Thierer and Neil Chilson. 2023. Overregulating AI Will Disrupt Markets and Discourage Competition. Bloomberg Law. Retrieved April 18, 2024 from <https://news.bloomberglaw.com/us-law-week/overregulating-ai-will-disrupt-markets-and-discourage-competition>

[40] Bergur Thormundsson. 2023. Amount of companies using ChatGPT in their business function in 2023, by industry. Statista. Retrieved April 10, 2024 from <https://www.statista.com/statistics/1384323/industries-using-chatgpt-in-business/>

[41] Bibek Upadhayay and Vahid Behzadan. 2024. TaCo: Enhancing Cross-Lingual Transfer for Low-Resource Languages in LLMs through Translation-Assisted Chain-of-Thought Processes. Retrieved April 10, 2024 from <http://arxiv.org/abs/2311.10797>

[42] Richard Van Noorden and Jeffrey M. Perkel. 2023. AI and Science: What 1,600 Researchers Think. Nature 621, 672–675.



## References

[43] Steve Vondran. 2020. How much does it cost to litigate a copyright case? Attorney Steve Blog. Retrieved April 18, 2024 from <https://www.vondranlegal.com/how-much-does-it-cost-to-litigate-a-copyright-case>

[44] Kevin Werbach (Ed.). 2020. After the Digital Tornado: Networks, Algorithms, Humanity (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781108610018>

[45] Kyle Wiggers. 2023. As OpenAI's multimodal API launches broadly, research shows it's still flawed. TechCrunch. Retrieved April 10, 2024 from <https://techcrunch.com/2023/11/06/openai-gpt-4-with-vision-release-research-flaws/>

[46] Mi Zhou, Vibhanshu Abhishek, Timothy Derdenger, Jaymo Kim, and Kannan Srinivasan. 2024. Bias in Generative AI. Retrieved April 18, 2024 from <http://arxiv.org/abs/2403.02726>

[47] 2018. California Consumer Privacy Act (CCPA). State of California - Department of Justice - Office of the Attorney General. Retrieved April 18, 2024 from <https://oag.ca.gov/privacy/ccpa>

[48] 2021. AI Risk Management Framework. nist.gov. Retrieved April 18, 2024 from <https://www.nist.gov/itl/ai-risk-management-framework>

[49] 2022. Yann LeCun on a vision to make AI systems learn and reason like animals and humans. Meta. Retrieved April 18, 2024 from <https://ai.meta.com/blog/yann-lecun-advances-in-ai-research/>

[50] 2023. Gartner Says More Than 80% of Enterprises Will Have Used Generative AI APIs or Deployed Generative AI-Enabled Applications by 2026. Gartner. Retrieved April 18, 2024 from <https://www.gartner.com/en/newsroom/press-releases/2023-10-11-gartner-says-more-than-80-percent-of-enterprises-will-have-used-generative-ai-apis-or-deployed-generative-ai-enabled-applications-by-2026>



# References

- [51] 2023. Joint Statement on AI Safety and Openness. Mozilla. Retrieved April 18, 2024 from <https://open.mozilla.org/letter/>
- [52] 2023. Federal Judge Dismisses Major Claims in Artists' Copyright Case Against AI Platforms. Legal.io. Retrieved April 18, 2024 from <https://www.legal.io/articles/5455194/Federal-Judge-Dismisses-Major-Claims-in-Artists-Copyright-Case-Against-AI-Platforms>
- [53] 2024. Data Suggests Growth in Enterprise Adoption of AI is Due to Widespread Deployment by Early Adopters. IBM Newsroom. Retrieved April 18, 2024 from <https://newsroom.ibm.com/2024-01-10-Data-Suggests-Growth-in-Enterprise-Adoption-of-AI-is-Due-to-Widespread-Deployment-by-Early-Adopters>
- [54] 2024. With “superhuman” artificial intelligence looming, Canada needs law now: AI pioneer. CBC News. Retrieved April 18, 2024 from <https://www.cbc.ca/news/politics/ai-pioneer-canada-needs-law-to-regulate-ai-now-1.7105463>
- [55] 2024. The generative AI boom in 6 charts. CB Insights Research. Retrieved April 23, 2024 from <https://www.cbinsights.com/research/generative-ai-funding-top-startups-investors-2023/>
- [56] 2024. the\_pile\_books3. Datasets at Hugging Face. Retrieved April 18, 2024 from [https://huggingface.co/datasets/the\\_pile\\_books3](https://huggingface.co/datasets/the_pile_books3)
- [57] 2024. The Pile (dataset). Wikipedia. Retrieved April 18, 2024 from [https://en.wikipedia.org/w/index.php?title=The\\_Pile\\_\(dataset\)&oldid=1217399543](https://en.wikipedia.org/w/index.php?title=The_Pile_(dataset)&oldid=1217399543)
- [58] Artificial Intelligence. Future of Life Institute. Retrieved April 18, 2024 from <https://futureoflife.org/cause-area/artificial-intelligence/>
- [59] Carnegie Silicon Valley: Deepfakes. Carnegie Endowment for International Peace. Retrieved April 18, 2024 from <https://carnegieendowment.org/siliconvalley/deepfakes>
- [60] EU AI Act. Retrieved April 18, 2024 from <https://www.euaiact.com/>



# References

- [61] **General Data Protection Regulation (GDPR) – Official Legal Text.** General Data Protection Regulation (GDPR). Retrieved April 18, 2024 from <https://gdpr-info.eu/>
- [62] **How to Build Your GPU Cluster: Process and Hardware Options.** run:ai. Retrieved April 18, 2024 from <https://www.run.ai/guides/multi-gpu/gpu-clusters>
- [63] **Llama 2 Model Details.** GitHub. Retrieved April 18, 2024 from <https://github.com/microsoft/Llama-2-Onnx/blob/main/MODEL-CARD-META-LLAMA-2.md>
- [64] **Meta Llama 2.** Meta Llama. Retrieved April 18, 2024 from <https://llama.meta.com/llama2/>
- [65] **Navigating the Complexities of Generative AI in Intellectual Property: Challenges and Opportunities.** EUIPO. Retrieved April 18, 2024 from <https://www.euipo.europa.eu/en/news/navigating-the-complexities-of-generative-ai-in-intellectual-property-challenges-and-opportunities>
- [66] **Prepare, Don't Panic: Synthetic Media and Deepfakes Clone.** WITNESS Media Lab. Retrieved April 18, 2024 from <https://lab.witness.org/projects/synthetic-media-and-deep-fakes/>
- [67] **Project GR00T.** NVIDIA Developer. Retrieved April 18, 2024 from <https://developer.nvidia.com/project-gr00t>
- [68] **SQ8.** What should the roles of academia and industry be, respectively, in the development and deployment of AI technologies and the study of the impacts of AI? | One Hundred Year Study on Artificial Intelligence (AI100). Retrieved April 18, 2024 from <https://ai100.stanford.edu/gathering-strength-gathering-storms-one-hundred-year-study-artificial-intelligence-ai100-2021-1/sq8>
- [69] **What is this site's policy on content generated by generative artificial intelligence tools? - Help Center.** Stack Overflow. Retrieved April 18, 2024 from <https://stackoverflow.com/help/ai-policy>



**DATA SCIENCE  ALLIANCE**