

Guest editor
Michael Eiden

Contributing authors

Ryan Abbott
Stephen J. Andriole
Michael Bateman
Foivos Christoulakis
Paul Clermont
Remy Gillet
Curt Hall

Nicholas Johnson
Philippe Monnot
Michael Papadopoulos
Elizabeth Rothman
Greg Smith
Eystein Thanisch

CUTTER
AN ARTHUR D. LITTLE
COMMUNITY

AMPLIFY

Vol. 36, No. 8, 2023



**Generative AI:
A Conversation with the Future**

CONTENT

4

OPENING STATEMENT

Michael Eiden,
Guest Editor



8

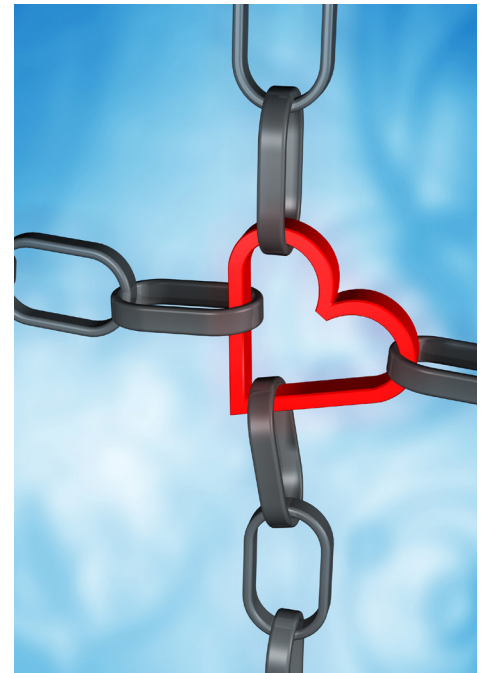
GENERATIVE AI IN THE ENTERPRISE: STATUS, PRACTICES & TRENDS

Curt Hall

18

GENERATIVE AI: LOVE, HATE, IGNORE, OR JUST REGULATE?

Stephen J. Andriole





26

LLM SECURITY CONCERNS SHINE A LIGHT ON EXISTING DATA VULNERABILITIES

Michael Papadopoulos, Nicholas Johnson, Michael Eiden, Philippe Monnot, Foivos Christoulakis, and Greg Smith

34

IP LAW IN THE ERA OF GENERATIVE AI

Ryan Abbott and Elizabeth Rothman



40

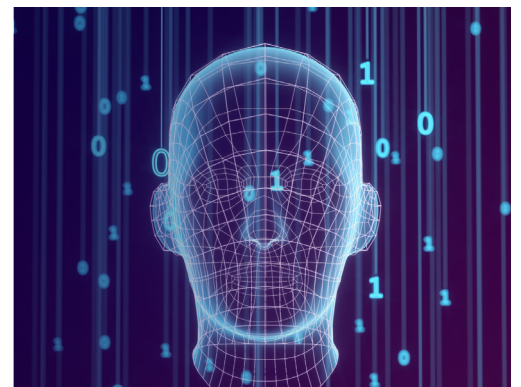
ENVIRONMENTAL IMPACT OF LARGE LANGUAGE MODELS

Greg Smith, Michael Bateman, Remy Gillet, and Eystein Thanisch

50

WHAT'S REALLY AHEAD FOR GENERATIVE AI?

Paul Clermont



GENERATIVE AI: A CONVERSATION WITH THE FUTURE

BY MICHAEL EIDEN, GUEST EDITOR

The technology industry is certainly not immune to hyperbole, but the speed of development in generative artificial intelligence (GAI) over the course of this year is unprecedented. In fact, it's mind-boggling. For quite a while, the term "AI" has been bandied about, applied to several market segments, and talked up in numerous articles — yet somehow never delivered the "wow factor" that was promised. That changed on 30 November 2022 with the release of OpenAI's ChatGPT.

Online AI bots that mimic human conversation aren't new, but ChatGPT seemed to operate on a different level, able to almost instantly answer complex questions and engage in philosophical debate. Most importantly, it did this in the style and syntax of a human, creating the experience of talking to a genuine artificial intelligence.

After catapulting GAI into the media spotlight, it became apparent that ChatGPT was still far from perfect, and attention shifted to copyright issues raised by a GAI model that learned by scraping the Internet for content. Within the technology industry, many believed that only major companies would have the computing power and resources necessary to train the large language models (LLMs) needed to properly run GAI apps.

That stance contrasts sharply with the open source community, which was inspired to new levels of innovation after the initial "leak" of Meta's LLaMA in early March. Things have moved quickly since the start of the year, and it's now possible to train much smaller language models on commodity hardware that can still solve very complex tasks. This is a genuine game changer in the potential application of GAI across all aspects of business and our lives online.

GAI will prove transformative, changing the way companies and organizations work forever, streamlining existing processes beyond recognition. Rather than working in a world in which interactions with data are rigidly rule-based and transactional, GAI will allow for questions and inquiries at a highly sophisticated level, producing in-depth, detailed answers in a format much more user-friendly than a standard data dump. Rather than performing a transaction, it will be like having a complicated conversation.

For example, internal knowledge management is lacking at most companies — they sit on a trove of information that is poorly indexed and saved in a variety of formats and languages. Imagine being able to ask your system to generate a report on your company's experience in a particular area, including customer use cases, summarized in German? Or how about using GAI to perform primary research on the latest academic or scientific papers or to ensure that your company's operations are always compliant with global regulatory frameworks, no matter how often they change?

GAI also has the potential to transform online customer relations. Today's online bots can lead to deeply frustrating experiences, causing unseen reputational damage. What if customers could talk to an online assistant that not only understood what they wanted (or could quickly find out) but also had access to real-time financial information, could instantly look up all known solutions to a problem, and could recommend a product based on detailed requirements? What if it could do all this in a conversational style tailored for the individual, rather than using predefined scripts?

The potential applications for GAI are almost limitless, saving companies enormous amounts of time and money compared to current processes, whether they relate to internal knowledge sharing and exploitation or external market analysis and customer service.

We are still at the very beginning of this revolutionary curve, and if we are to fully enjoy its advantages, there are important issues to be resolved in areas such as intellectual property (IP) protection, regulation, security, and environmental impact. This *Amplify* takes a look at these issues — and more.

IN THIS ISSUE

We begin with a fascinating dive into data on key GAI trends. Cutter Expert and frequent *Amplify* contributor Curt Hall examines findings from a Cutter survey of more than 100 organizations worldwide. So many respondents are already using GAI tools that Hall calls the rate of adoption “amazing.” Most companies are still using basic tools, but many report they’re open to using a range of tools, including domain-specific ones. Hall’s article looks at enterprise adoption of LLMs, strategy and oversight for GAI adoption and usage, and enterprise experience with GAI to date. In addition to graphs showing the survey results, Hall highlights direct quotes from respondents, including this one from a communications executive:

We are doing comparisons that show promise in amplifying the creativity and productivity of our teams in significant ways, which vary depending on the type of job. This is already showing that the use of these tools will lead to significant savings.

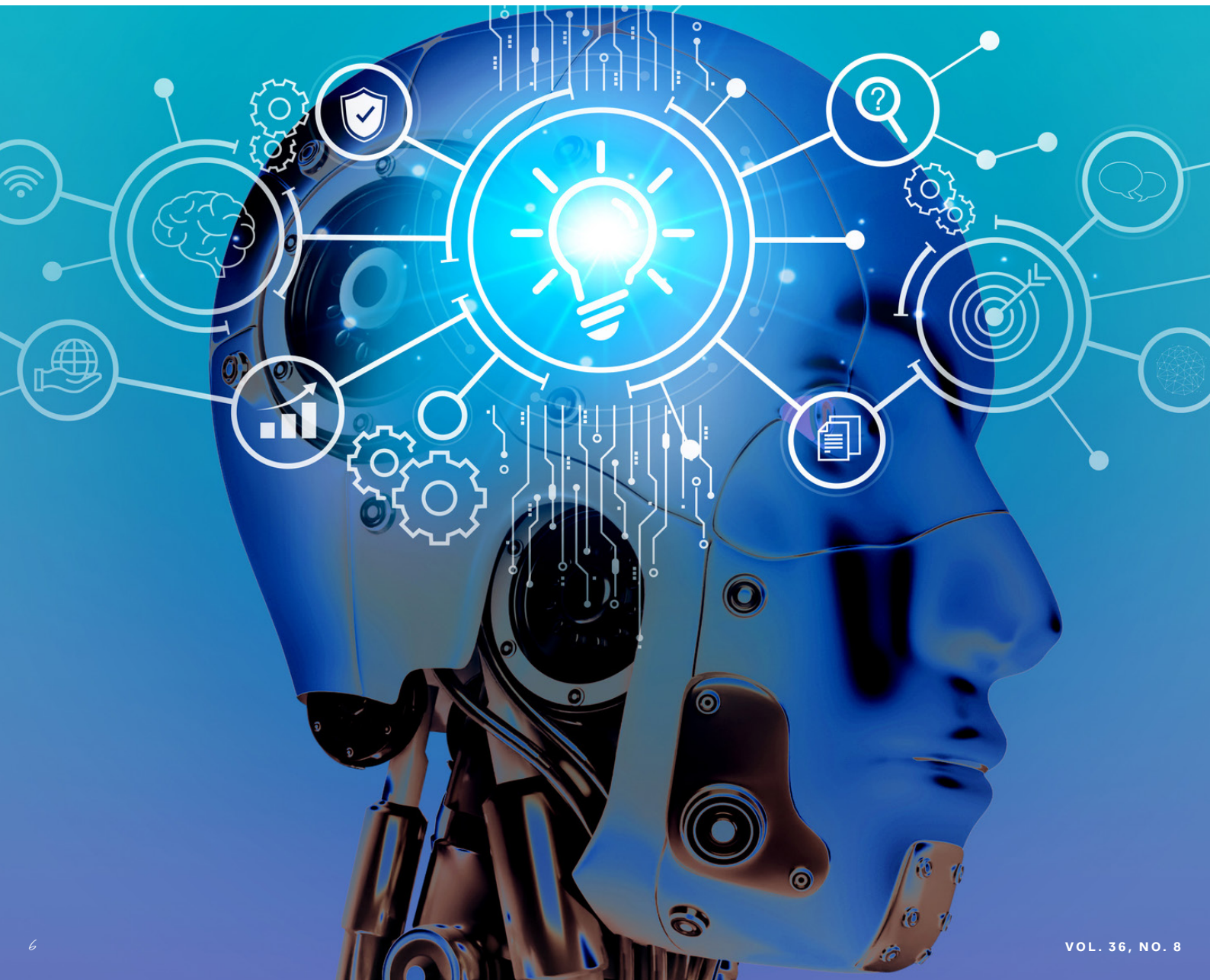
Next, Cutter Fellow Stephen J. Andriole presents a no-holds-barred discussion of the predictions and fearmongering swirling around GAI. Clearly, Andriole says, we should stop panicking and start thinking about how to optimize GAI. We should also acknowledge that some form of regulation is necessary. Andriole turns to ChatGPT and Bard (who else?) for advice on potential regulation, looks closely at what other countries and regions are doing in this area, and highlights the importance of addressing IP infringement issues. He concludes by saying that regulatory decisions should not be anchored in technology capabilities, pointing out that social, political, and economic concerns about the impact of regulation will exert as much, if not more, influence on the regulatory scenarios that emerge.

In our third piece, me and my Arthur D. Little (ADL) colleagues Michael Papadopoulos, Nicholas Johnson, Philippe Monnot, Foivos Christoulakis, and Greg Smith debunk the idea that security concerns about LLMs are entirely new. We examine each concern to show that these issues are merely new manifestations of existing security threats — and thus manageable. “LLMs highlight and stress test existing vulnerabilities in how organizations govern data, manage access, and configure systems,” we assert. The article concludes with a list of 10 specific ways to improve LLM-adoption security.

THE POTENTIAL APPLICATIONS FOR GAI ARE ALMOST LIMITLESS, SAVING COMPANIES ENORMOUS AMOUNTS OF TIME AND MONEY COMPARED TO CURRENT PROCESSES

**GAI WILL PROVE
TRANSFORMATIVE,
CHANGING
THE WAY
COMPANIES AND
ORGANIZATIONS
WORK FOREVER,
STREAMLINING
EXISTING
PROCESSES
BEYOND
RECOGNITION**

Our fourth article comes from Ryan Abbott and Elizabeth Rothman who believe we must address the legal, ethical, and economic implications of AI-generated output if we want to foster innovation, promote the responsible use of AI, and ensure an equitable distribution of the benefits arising from AI-generated works. The authors look at the complicated relationship between AI and IP, then discuss the Artificial Inventor Project, which filed two patent applications for AI-generated inventions back in 2018 in the UK and Europe. The project aims to promote dialogue about the social, economic, and legal impact of frontier technologies like AI and generate stakeholder guidance on the protectability of AI-generated output. Clearly, say Abbott and Rothman, AI systems challenge our existing IP frameworks and necessitate a thorough rethinking of what rules will result in the greatest social value.



Next, ADL's Greg Smith, Michael Bateman, Remy Gillet, and Eystein Thanisch scrutinize the environmental impact of LLMs. Specifically, they compare carbon dioxide equivalent (CO₂e) emissions from LLMs with using appliances such as electric ovens and kettles, streaming videos, flying from New York City to San Francisco, and mining Bitcoin. Next, the authors look at how fit-for-purpose LLMs and increased renewable energy usage could help LLM operators reduce their carbon footprint. Finally, this ADL team points out the relationship between smaller LLMs and responsible, democratized AI.

Finally, Cutter Expert Paul Clermont takes a down-to-earth look at what we can expect from AI in the near term. For one thing, he says, we're still in the garbage-in, garbage-out phase with

LLMs; for another, it's nowhere close to artificial general intelligence. There are, of course, ethical and social implications, including the fact the AI puts what we don't like about today's Internet (disinformation, loss of privacy, and more) on steroids. A host of new legal issues also needs attention, Clermont notes, which may lead to governments playing a role in the evolution of AI usage that they did not assume in the advent of the computer or the Internet.

We hope the articles in this issue of *Amplify* offer you insightful ways to examine and ponder the potential of GAI going forward and recognize the importance of viewing this complicated technology with an objective eye.

About the guest editor

MICHAEL EIDEN

Michael Eiden is a Cutter Expert, Partner and Global Head of AI & ML at Arthur D. Little (ADL), and a member of ADL's AMP open consulting network. Dr. Eiden is an expert in machine learning (ML) and artificial intelligence (AI) with more than 15 years' experience across different industrial sectors. He has designed, implemented, and productionized ML/AI solutions for applications in medical diagnostics, pharma, biodefense, and consumer electronics. Dr. Eiden brings along deep expertise in applying supervised, unsupervised, as well as reinforcement ML methodologies to a very diverse set of complex problem types. He has worked in various global technology hubs, such as Heidelberg (Germany), Cambridge (UK), and Silicon Valley (US), with clients ranging from small and medium-sized enterprises to globally active organizations. Dr. Eiden earned a doctorate in bioinformatics. He can be reached at experts@cutter.com.

GENERATIVE AI IN THE ENTERPRISE: STATUS, PRACTICES & TRENDS

Author

Curt Hall

In its relatively brief existence, generative artificial intelligence (GAI) has both amazed and alarmed, due to its ability to perform tasks previously considered too dependent on human knowledge or creative skills. Today, everyone from rank-and-file employees to managers and CxOs at almost every organization in the world is trying to figure out how to effectively and safely apply the technology.

At the moment, it can be difficult to determine the current status of GAI within the enterprise and what the future holds. To gain insight into these and other important questions, Cutter conducted a survey from April–May 2023 about how organizations are adopting GAI and what they see as the possible impacts on their businesses and industries. We also asked about key trends organizations are encountering, or foresee arising, when adopting the technology.

This article aims to assist organizations' efforts to leverage GAI by examining some of the findings. Specifically, it covers the following:

- Current status of GAI in the enterprise and future plans, including the GAI technologies and commercial GAI products organizations are using/planning to use
- Enterprise adoption of large language models (LLMs)
- Strategy, oversight, and employee support for GAI adoption and usage
- Enterprise experience with GAI to date

Our findings are based on the responses of 103 global organizations. Where applicable, we offer anonymous quotes from participants who were kind enough to share their thoughts and experiences with GAI. For more on survey methods, including demographics, see end of article.

GA I IN THE ENTERPRISE: CURRENT STATUS & FUTURE PLANS

Our first key question considers to what extent organizations are currently using or planning to use GAI. As shown in Figure 1, just under half of surveyed organizations are already using tools like ChatGPT, DALL-E, and Jasper AI. Another 14% indicate they plan to do so within the next six to 12 months. A further 19% report that their organizations are seriously considering its use.

This rate of adoption is, quite frankly, amazing. In my decades of conducting surveys measuring enterprise adoption of advanced and emerging information technologies, I've never seen anything like this — certainly not for AI! Although GAI has only been generally available (in the form of commercial products) for about seven months, 81% of respondents say their organizations are either using it, planning to implement it, or seriously investigating doing so. It's also notable that only 11% of surveyed organizations have no plans to use GAI in the foreseeable future.

Figure 2 provides additional insight into adoption by organizations using GAI. It appears that respondents are somewhat unsure as to how GAI is making its way into their organizations: for the majority, unapproved and approved use of GAI tools is about equal (49% versus 46%, respectively).

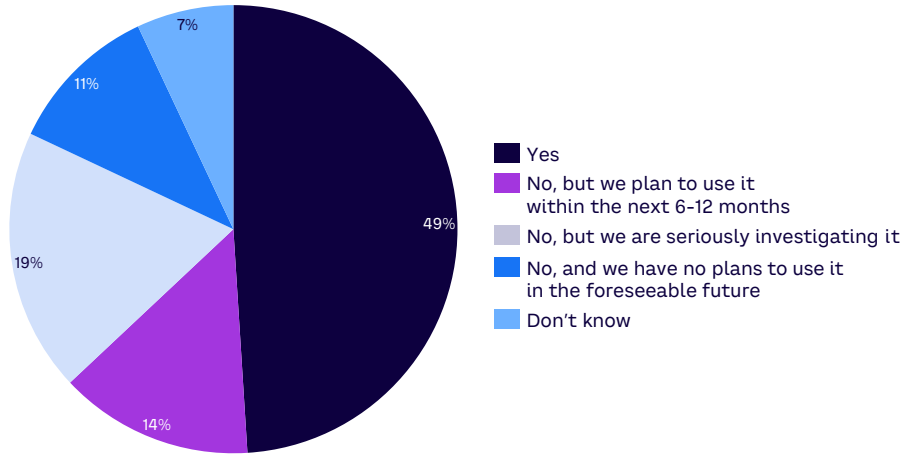


Figure 1. Is your organization using, or planning to use, GAI technology like ChatGPT, DALL-E, Jasper, or similar?



Figure 2. How is your organization currently using GAI?

This trend can be attributed to the relative newness of this popular technology. Employees have ready access to a range of GAI-powered writing tools, text-to-image (and other AI art) generators, computer coding, and other programs that are available on a variety of easy-to-use platforms, even mobile devices. Employees are eagerly using them, whether in an official or unofficial capacity, as noted by this executive VP at a manufacturing company:

Individuals are using generative AI to accelerate content generation, but there is no prescribed or designed practice yet.

Other organizations are using GAI tools in a more approved and guided manner, as indicated by this university department chairperson:

Individuals within the organization are using [generative AI]. We are also looking to bring everyone up to speed with AI literacy with professional development initiatives.

Another interesting finding is that, for all the talk in the general press about organizations outright banning their employees (or selective groups of employees) from using GAI (usually due to security considerations around sensitive data, privacy, and intellectual property [IP]), few respondents indicated their organizations have done so.

Note that organizations that do not allow their employees to use GAI tools must recognize the potential loss of productivity gains afforded by the technology, and those allowing it must devise formal policies to (1) regulate its use and (2) capture and document its benefits.

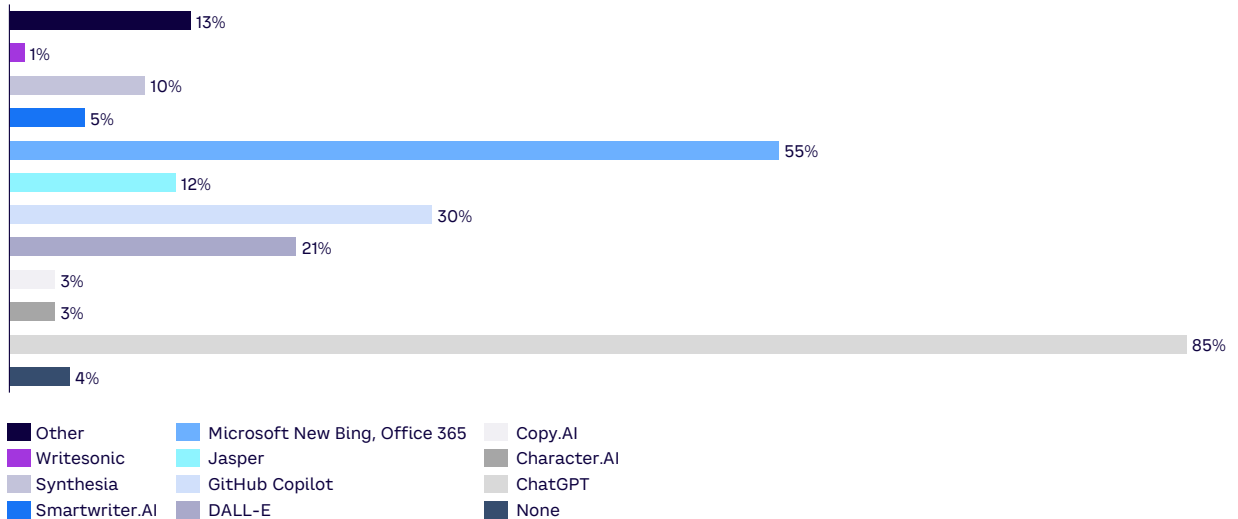


Figure 3. Which commercial GAI providers' products is your organization using/planning to use?

WHICH GAI TECHNOLOGIES?

The overwhelming majority of organizations in our survey are using basic GAI tools like ChatGPT, Craiyon, DALL-E, and Stable Diffusion. This makes sense, if you consider that most of these tools are free or low-cost. That said, organizations appear open to the possibility of using a range of GAI tools, as noted by numerous respondents and summed up by this chief scientist at a manufacturer:

We are exploring all relevant, available generative AI tools that might be leveraged for the benefit of our organization.

Our research also found considerable interest among organizations in using domain-specific enterprise tools and applications featuring integrated GAI functionality. Examples include Cognigy (contact center), Jasper (marketing), GitHub Copilot (code generation), Kaizen Chat (customer support/chat/email), and Lavender (sales).

However, use of industry-specific GAI enterprise tools and applications is quite limited. Examples include CALA (fashion design), COVU (insurance), and Harvey (legal). Most of these tools are still in development or available only to select customers in beta form. However, some organizations are exploring their use, as indicated by this CIO at a major university:

We are interested in generative AI tools integrated with EDU [educational] platforms — LMS ([learning management systems], plagiarism detection, admissions, and IT service management.

Expect the use of domain-specific enterprise tools and applications and industry-specific enterprise tools and applications with integrated GAI functionality to increase significantly over the next few years. Reasons include: (1) huge interest among end-user organizations and (2) the fact that almost all leading enterprise software providers (SAP, Oracle, Microsoft, Salesforce) and a slew of start-ups are integrating GAI capabilities into a broad range of tools and commercial applications.

WHICH COMMERCIAL GAI PRODUCTS?

OpenAI's ChatGPT is the overwhelming tool of choice among end-user organizations using GAI in our survey, followed by Microsoft's Bing Chat and (a distant third) GitHub Copilot (see Figure 3). This is not surprising because these were the first GAI products to become generally available and were readily adopted by many users in record-setting time. Their adoption was also accelerated by Microsoft integrating OpenAI's technologies (ChatGPT, DALL-E 2) into its Edge browser and other products (e.g., Office) and by the company's formidable marketing muscle.

Clearly, ChatGPT (85%) and Bing Chat (55%) are the current leaders when it comes to enterprise GAI use. But a large number of GAI tools are available from a range of vendors, including many start-ups. Organizations appear quite willing to explore these new products, as noted by this business strategist at a financial services company:

ChatGPT has set the early standard, but all relevant generative AI tools are on the table at this stage.

Additionally, although Bing Chat is popular, several organizations pointed out its drawbacks, as commented on by this CEO at a computer consulting firm:

Microsoft's new Bing is useful because it provides access to data that is currently on the Web. However, it is extremely politically correct and otherwise constrained in its answers and, therefore, is far less useful, particularly than ChatGPT-4.

Respondents had the option to indicate other GAI products their organizations are using. Responses included a wide range of tools and commercial applications, including Adobe Firefly, Ask Sage, Google Bard, Anthropic Claude, Databricks Dolly 2.0, Hugging Face BLOOM and StarCoder, Otter.ai OTTER, Mathis Lichtenberger ChatPDF, Microsoft Power BI Copilot, Midjourney, Runway AI Magic Tools, Shutterstock AI image generator, Stability AI Stable Diffusion, and Wonder Dynamics Wonder Studio AI.

ENTERPRISE ADOPTION OF LLMs

Approximately a third of surveyed organizations plan to integrate LLMs into their own applications. Again, this is an impressive rate of adoption, considering that most organizations (outside of tech) have little to no experience working with LLMs. Nearly half of respondents are still unsure about their plans for using LLMs, essentially taking a wait-and-see approach.

Organizations are keenly interested in LLMs because they have proven to improve accuracy in natural language processing (NLP) systems. Moreover, the general availability of LLMs (particularly open source versions) is enabling enterprises, commercial developers, and entrepreneurs to build systems that can perform much more sophisticated NLP tasks.

The caveat is that, for enterprise use, organizations need to train their LLMs on their own data to meet accuracy requirements and avoid the potential for the hallucinations, biases, and other inconsistencies that have so far put a damper on the greater use of LLMs in the enterprise.

This is especially true when it comes to supplementing customer-facing applications like chatbots, conversational interfaces, intelligent assistants, and other self-service customer-assist systems (i.e., automated applications where the output is not first screened for accuracy/correctness by a customer service rep or other human). For these reasons, many organizations are still evaluating LLM applications, as noted by this software engineer/R&D at a telecom company:

We are exploring the use of LLMs. But it all depends on error rates ... how reliable it will be, minimizing the risks, and being able to evaluate its responses. Keep in mind, what you are doing is bringing someone else's software into your organization and exposing its (your organization's) innermost workings.

We expect to see LLMs integrated into a wide range of proprietary enterprise applications and commercial software products — everything from programming and application development tools to customer service, data integration/access, business intelligence, and cybersecurity environments. This is happening at a furious pace today.

STRATEGY, OVERSIGHT, SUPERVISION & SUPPORT

Organizations are currently somewhat lacking when it comes to having a detailed strategy in place for adopting GAI (see Figure 4). This is quite understandable, with the technology being so new. Moreover, new GAI tools appear each day, making planning for adoption a moving target. But this trend appears destined to change soon, as more than three-quarters of responding organizations either plan to implement such strategies within six to 12 months or are seriously investigating doing so. Here is a comment by a IT director in retail:

We are working to create some strategy and guardrails around usage of ChatGPT while exploring business value-driven use cases.

Just 7% of surveyed organizations have no plans in the foreseeable future to develop a detailed strategy for adopting GAI, demonstrating the apparent value that organizations have placed on the technology, as well as the importance of needing to implement strategies to guide its adoption.

OVERSIGHT & STEERING OF GAI ADOPTION

Current oversight of generative AI adoption in the enterprise is quite high: more than a third of surveyed organizations (34%) have assigned or created a group charged with overseeing adoption and use of technology within their establishments. Another 42% plan to establish a group within six to 12 months or are seriously considering doing so in the future.

However, almost a quarter of respondents either had no plans to establish a group charged with overseeing generative AI adoption in the foreseeable future or don't know the status of such groups in their organizations.

Based on these findings, we expect to see many organizations fleshing out their GAI plans over the next 12 months or so.

EMPLOYEE SUPERVISION & SUPPORT

Current and future plans to support employees with their use of GAI is high (see Figure 5). Nearly 80% of surveyed organizations are either already providing training and other resources to assist employees with using GAI tools within their workflows, plan to do so within six to 12 months, or are seriously considering doing so.

A good example of the kind of support organizations seek to provide employees is offered by this senior manager at a publishing/media firm:

We've got a good governance system that has been developed in the past few months that provides our talent with information on which tools to use, when and how, and for doing what type of work ... all through company accounts that are set up for them to use.

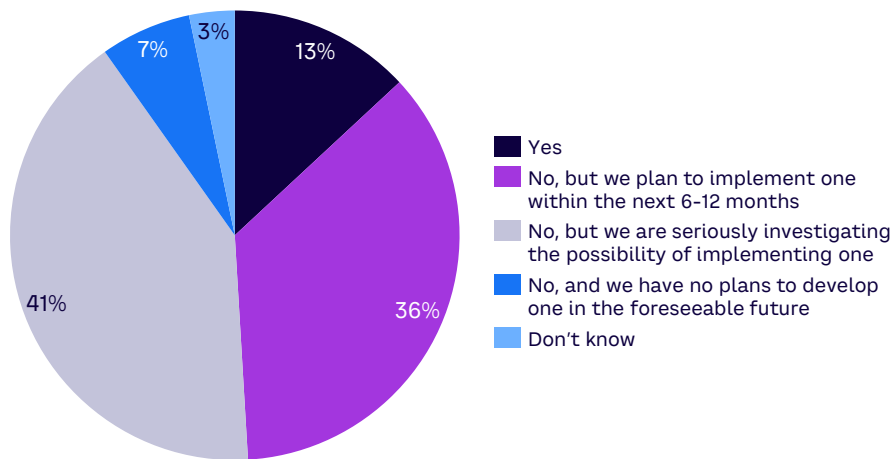


Figure 4. Has your organization implemented, or does it plan to implement, a detailed strategy for adopting/using GAI?

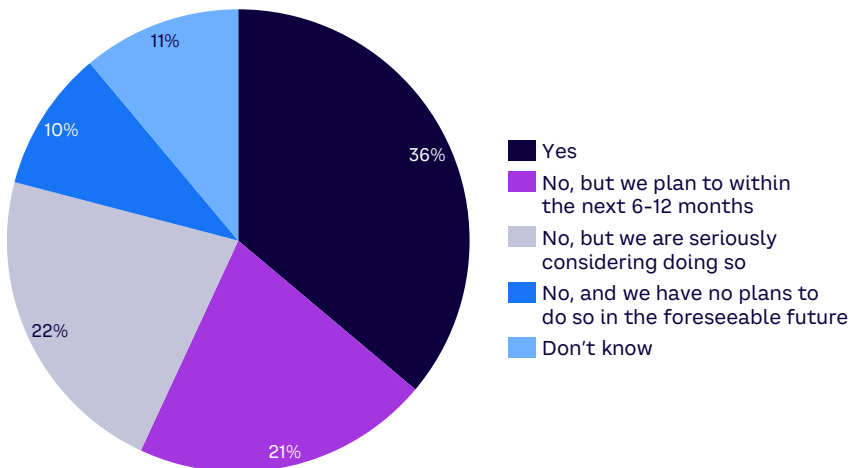


Figure 5. Is your organization currently providing, or planning to provide, any resources (e.g., training, support) to help employees effectively use and integrate GAI into their work processes?

At this stage, we highly recommend providing guidance to employees on the application of GAI tools, particularly on reducing the chance of accidental release of customer data, IP, and other sensitive information.

ENTERPRISE EXPERIENCE WITH GAI SO FAR

One key question around enterprise use of GAI is whether organizations are experiencing benefits from its adoption, including how it is impacting business operations, employee productivity, cost savings, and customer satisfaction.

MEASURABLE BENEFITS

More than 20% of surveyed organizations are already realizing measurable benefits from GAI. This is impressive, given that most organizations have only been using the technology for a short time. However, 15% have not seen any measurable benefits yet. For the majority of organizations, the jury is still out when it comes to whether or not they are benefiting from the technology.

TRANSFORMATION OF BUSINESS OPERATIONS

Just over 10% of survey respondents say that GAI use has led to changes in the way the organization or some of its lines of business (LOBs) operate (see Figure 6). However, more surveyed organizations using GAI indicate that its use has not led to any operational changes (16%).

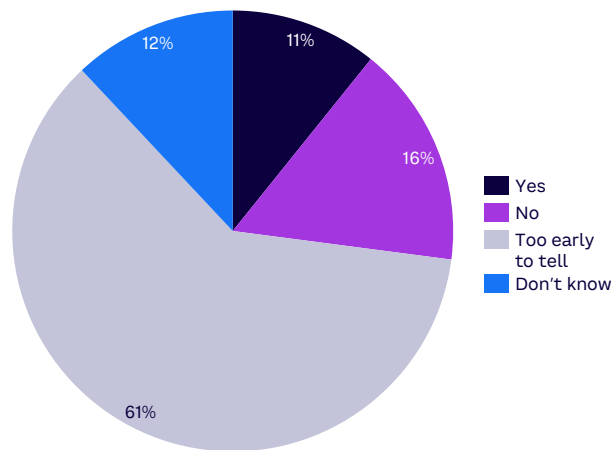


Figure 6. Has use of GAI transformed the way your organization or any of its LOBs operates?

For the majority of organizations currently using the technology (61%), it is simply too early to tell how it is impacting their business operations, as explained by this senior business strategist at an insurance company:

We could be experiencing a transformation today because our marketplace focuses on financial services, and generative AI appears to hold promise in this field, but it is too early to tell how such a transformation is taking shape.

This makes sense since business transformation is typically a complicated undertaking that requires time and a good deal of planning, including determining how to successfully apply a new technology.

IMPROVED EMPLOYEE PRODUCTIVITY

Our survey indicates that initial enterprise use of GAI has resulted in improved employee productivity at some organizations. In fact, more organizations — almost 30% — report gains in employee productivity through early use of GAI than don't.

Here are some typical comments about the impact of GAI on employee productivity:

We are not a technology company. We are a nonprofit. It [generative AI] does allow us to do more without having to increase staffing size.

— CEO, nonprofit organization

We are using ChatGPT to help create a book. We are still doing significant editing, and it has been great to have a starting point. It has greatly accelerated our process.

— CEO, management consulting firm

Other respondents report that GAI is benefiting employees in other, perhaps unforeseen, ways, as noted by this VP at a consulting firm specializing in implementing AI solutions:

Happier, more engaged employees. In particular, chatbots provide a naturalistic partner for individuals, for motivation, for job satisfaction, for “emotional” camaraderie, and for being able to ask “dumb” questions without judgment.

Although these early findings appear promising, most organizations using GAI are either still waiting to see how using the technology will impact employee productivity or are unsure at this time.

COST SAVINGS

So far, surveyed organizations have experienced mixed results when it comes to realizing cost savings from early use of GAI (see Figure 7). The majority (56%) are still waiting to see what happens. This makes sense because cost-savings analyses take time to conduct, and the comments we received indicate that organizations are just now beginning to attempt to measure such savings, as noted by this executive at a communications and media company:

We are doing comparisons that show promise in amplifying the creativity and productivity of our teams in significant ways, which vary depending on the type of job. This is already showing that the use of these tools will lead to significant savings.

CUSTOMER SATISFACTION

GAI holds considerable promise for increasing customer satisfaction. For example, by providing highly personalized responses in the form of recommendations and tailored content to customer inquires, GAI could help organizations create more positive experiences that would enhance customer satisfaction. Similarly, R&D departments could use GAI to devise improved products that could lead to more satisfied customers.

However, our research indicates that, to date, organizations have experienced mixed results when it comes to how early use of GAI impacts customer satisfaction. Although some organizations (11%) say their initial use of the technology has had a positive impact on customer satisfaction, most are still waiting to see what happens.

CONCLUSION

Our research points to a number of important findings about the current/future status and adoption of GAI in the enterprise:

- **Although GAI is very new in the form of commercial products, organizations are rapidly adopting it.** Almost half of the organizations we surveyed are using the technology. Future plans for adoption are also high.
- **Adoption of GAI is somewhat haphazard.** For the majority of organizations, unapproved and approved use of generative AI tools is about even.

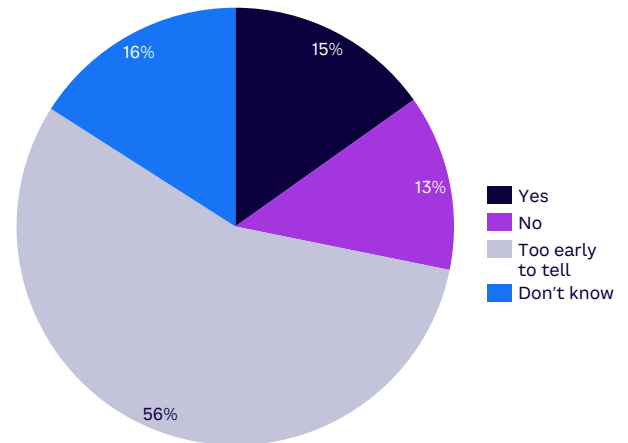


Figure 7. Has using GAI led to any cost savings for your organization?

- **Most organizations are currently using basic GAI tools like ChatGPT.** But they are also very interested in using domain-specific enterprise tools and applications with integrated GAI functionality.
- **Current use of industry-specific enterprise tools and applications featuring integrated generative AI capabilities is limited.** We believe this is primarily due to the limited availability of such products.
- **ChatGPT, Bing Chat, and Copilot are the most popular tools among end-user organizations using GAI.** But organizations are eagerly exploring new products as well.
- **Interest in leveraging LLMs in the enterprise is high.** About a third of organizations currently plan to integrate LLMs into their own applications.
- **Few organizations currently have a detailed strategy in place for adopting GAI.** Expect organizations' plans to implement such strategies to accelerate considerably over the next six to 12 months.
- **Oversight of GAI adoption in the enterprise is high.** More than a third of surveyed organizations have already assigned or created a group to oversee adoption and use of the technology.
- **Current and future plans to support employees with use of GAI are high.** Nearly 80% of respondents are either providing training to assist employees with using GAI, plan to do so in the next six to 12 months, or are considering doing so.

- **Some organizations are realizing measurable benefits from GAI.** However, this is limited to only about 20% of those surveyed. Most are still waiting to see what benefits the technology will provide.
- **GAI has had a limited impact on business transformation so far.** Only 11% of respondents indicate that use of the technology has changed the way the organization or some of its LOBs operate. For most, it's too early to tell.
- **Enterprise use of GAI is improving employee productivity.** Nearly 30% of surveyed organizations indicate this is the case.
- **Organizations have experienced mixed results when it comes to costs savings from initial use of GAI.** Most are still waiting to see what happens.
- **Initial use of GAI has had a limited impact on customer satisfaction.** However, for most organizations, it is too early to tell how the technology is affecting their customer-satisfaction efforts.
- **Banning GAI tools outright in the enterprise is rare.** Organizations should carefully consider before implementing such bans to avoid missing out on employee productivity gains and other possible benefits from using the technology.

SURVEY DEMOGRAPHICS

Our findings are based on the responses of 103 worldwide organizations of various sizes. Fifty-two percent are headquartered in North America, 25% in Europe, 9% in the Middle East, 7% in India, and 4% in Asia/Australia/Pacific, with the remainder in Africa and South America.

Responding organizations' annual revenues vary, with 8% having annual revenues of more than US \$50 billion, 11% between \$10 billion and \$50 billion, 14% between \$1 billion and \$10 billion, 20% with more than \$50 million to \$1 billion, and the remaining 47% with annual revenues less than \$50 million. Figure A below shows responding organizations broken down by industry.

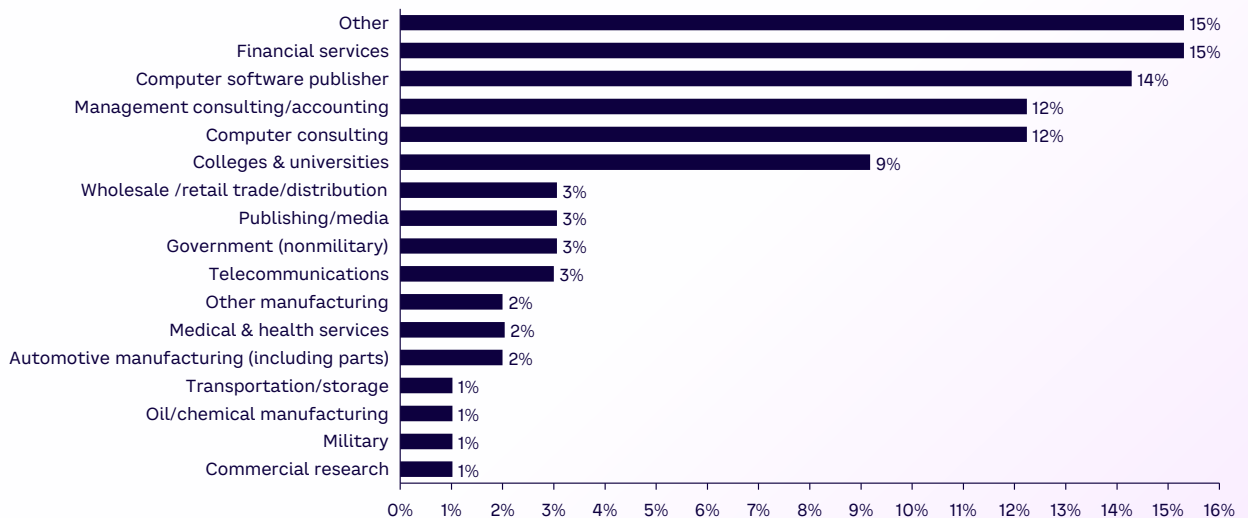


Figure A. Respondent organizations by industry

About the author

Curt Hall is a Cutter Expert and a member of Arthur D. Little's AMP open consulting network. He has extensive experience as an IT analyst covering technology and application development trends, markets, software, and services. Mr. Hall's expertise includes artificial intelligence, machine learning, intelligent process automation, natural language processing and conversational computing, blockchain for business, and customer experience management. He also focuses on the Internet of Things, including platforms, architectures, and use cases; big data platforms and use cases; and business intelligence (BI), predictive modeling, and other analytic practices. Mr. Hall's research also includes mobile and social technologies in the enterprise as well as mobile BI and collaboration. He has conducted extensive research on how all these technologies are being applied to develop new advisory, decision support, customer engagement, and other enterprise applications.

Mr. Hall is a frequent contributor to Cutter's Technology and Sustainability research deliverables as well as *Amplify*. He served as Editor of numerous Cutter journals, including *Intelligent Software Strategies*, *Data Management Strategies*, and *Business Intelligence Advisor*. His recent studies examining the enterprise adoption of key emerging technologies and practices delivered four series of in-depth, survey-based Cutter Consortium research: "AI & Machine Learning in the Enterprise," "Blockchain Rising," "CX Management in the Enterprise," and "IPA in the Enterprise." Mr. Hall also coauthored, with Cutter contributor Paul Harmon, *Intelligent Software Systems Development: An IS Manager's Guide* and contributed to James Martin and James Odell's *Object-Oriented Methods: Pragmatic Considerations*. His work has appeared in various technical journals and IT publications, including as a contributing author to PricewaterhouseCoopers Technology Forecast Yearbooks. Mr. Hall can be reached at experts@cutter.com.



**GENERATIVE
AI: LOVE,
HATE,
IGNORE,
OR JUST
REGULATE?**

Author

Stephen J. Andriole

We're on the verge of creating the smartest assistants in history, ones that can help us cure cancer, plan cities, improve the legal system, and manage environmental disaster, among other tasks that humans have had difficulty performing.

Large language models (LLMs) distribute power to individuals who have been trying to optimize intelligent systems for years. This provides conversational connectivity to old, newly created, and real-time knowledge that can help solve problems humans have avoided or just plain botched. And who wouldn't want to connect to what venture capitalist Rob Toews describes in *Forbes* as "the world's total stock of usable text data.... This includes all the world's books, all scientific papers, all news articles, all of Wikipedia, all publicly available code, and much of the rest of the Internet"?¹

There are some legitimate critics of the nature of the "intelligence" that LLMs reflect, but it's safe to say that the impact of LLMs and their access platforms will be enormous. It's not a matter of "if" but "when" this impact will be felt across all industries and within every household in the world that opts into its potential.

Note that this "intelligence" will initially take the form of "assistants" but will soon advance to "partners," and in some cases "bosses." Again, no one knows when these promotions will occur, but they will selectively happen across tasks, domains, industries, and even households. There's no question about the outcome — so we can fight it or welcome the inevitability.

STUPIDITY, PANIC & PAUSES

Every new technology has critics, skeptics, and those just plain terrified of what the technology can do. US President Benjamin Harrison and his wife were afraid to turn on the light switches in the White House.² There were critics of bicycles, cars, nail polish, talkies, laptops, answering machines, and even cheeseburgers.³

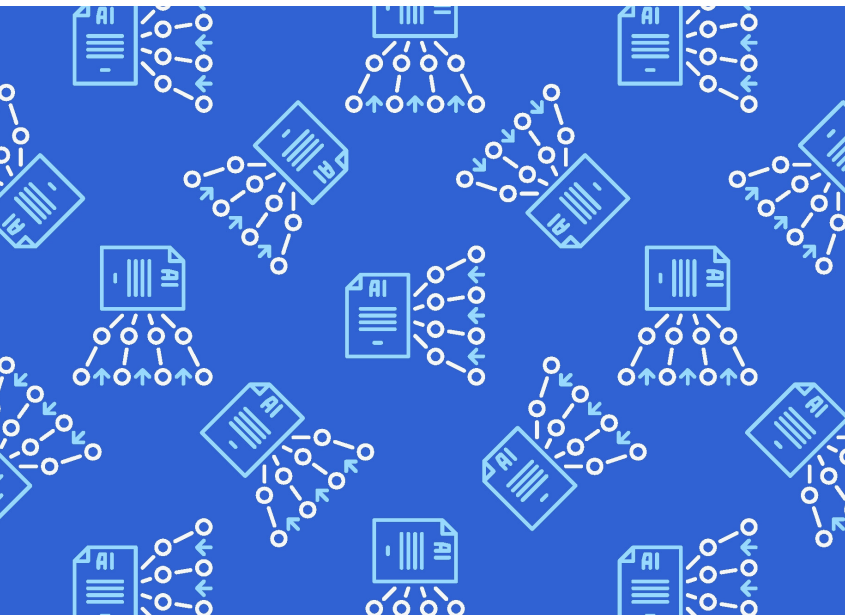
Artificial intelligence (AI), machine learning (ML), and its accelerant, generative AI (GAI), have triggered a new round of debates, with some contemporary heavyweights like Elon Musk sounding yet another attention-seeking alarm. But is the concern justified or another case of technology panic?

THERE'S NO QUESTION ABOUT THE OUTCOME; WE CAN FIGHT IT OR WELCOME THE INEVITABILITY

The recent proposal by AI experts to “pause” research in AI, ML, and GAI is, of course, silly.⁴ Does anyone really believe the proposal will have any impact at all, other than some publicity?⁵ Will the US Congress actually do something? Most likely, the idea is to force people to think about the longer-term implications of handheld access to massive digital intelligence.

GOOD & BAD, BUT WHO'S COUNTING?

LLMs can be extremely helpful. The current conversation (which will no doubt morph into something else in six months) is more negative than positive. We should stop panicking about GAI and start thinking about how to optimize its use. Due diligence is appropriate where “good” and “bad” are assessed as objectively as possible.



Here's a short list of “good”:⁶

- Chatbot/virtual assistants
- Fraud detection
- Translation
- Content creation/research
- Sentiment analysis

Here are some good questions to ask: Could GAI cut healthcare costs or develop new cancer drugs?⁷ Could it disrupt healthcare completely?⁸ Could it disrupt city planning?⁹ What about climate change?¹⁰ Product design?¹¹ Customer service?¹²

We asked ChatGPT about all this.

Question: What are the good things you can do?

Answer (all text in italics has been written by AI):

- *Creative content generation*
- *Personalization*
- *Healthcare*
- *Education*
- *Environmental science*
- *Language translation*
- *Gaming*

Question: What are some bad things you can do?

Answer:

- *Create fake news or propaganda*
- *Generate fake images or videos*
- *Cybersecurity threats*
- *Deepfakes*
- *Unintended biases*

Make no mistake, technology assassinates jobs — which is its *raison d'être*. Efficiency, cost savings, profitability, and market share are always the targets of technology, which is why we love (and hate) technology. Those who own old business models hate it; those creating new ones love it. “New entrants” are praised for their creativity and ingenuity. It all depends on where you sit. It remains to be seen if job displacement is a 7 or a 10. Every company on the planet is trying to measure the impact that AI, ML, and GAI will have on itself, its competitors, and the structure of its industry. At home, everyone will eventually make some new digital friends.

GAI can also distort reality in so many ways. This activity makes GAI a close friend of liars, cheats, and tyrants. GAI can rewrite cybersecurity algorithms, invade our privacy, and make us believe we're not seeing what we're seeing or hearing what we're hearing — even as it develops lifesaving vaccines.

LOVE, HATE, IGNORE, OR REGULATE?

This depends on where one sits. Those who use AI, ML, and GAI to make money (no matter how it's made) love GAI. Those who worry about truth and reality may hate what GAI can do. (No one should plead ignorance.) Facebook's algorithms that threaten the mental well-being of young women is a perfect example of how ML can increase advertising dollars. GAI will improve these algorithms.

Some form of regulation is clearly necessary. Even ChatGPT and Google Bard agree. Some countries are on their way to legislative control.¹³ At the very least, this technology (like all technologies) must be "managed."

When asked about this regulation, ChatGPT knows exactly what to do:

- **Establish clear ethical guidelines.** Governments, industries, and institutions should collaborate to establish ethical guidelines to ensure the responsible development and use of generative AI. These guidelines should address issues such as privacy, security, transparency, and accountability.
- **Require transparency.** Generative AI systems should be designed to be transparent, meaning that their decision-making processes should be clear and understandable. Developers should be required to provide explanations of how their systems work and how they make decisions.
- **Implement data privacy laws.** Generative AI systems rely on vast amounts of data to learn and improve. Regulations should be in place to protect individuals' privacy rights and ensure that data is collected and used ethically and legally.

- **Create oversight agencies.** There should be agencies responsible for overseeing the development and deployment of generative AI systems. These agencies should have the authority to review and approve or reject new systems and to investigate any issues or complaints that arise.
- **Encourage collaboration.** The development and use of generative AI should be a collaborative effort between government, industry, academia, and the public. These stakeholders should work together to identify and address potential ethical, social, and legal issues.
- **Ensure accountability.** Generative AI developers and users should be held accountable for any negative consequences that result from their systems. There should be a system in place to monitor and regulate the use of these systems, with appropriate penalties for noncompliance.

Bard agrees:

- **Content moderation.** Governments could require companies that use generative AI to moderate the content that is generated. This could involve filtering out harmful or offensive content or requiring users to verify their identity before generating content.
- **Transparency.** Governments could require companies that use generative AI to be transparent about how their systems work. This would allow users to understand how their data is being used and to make informed decisions about whether or not to interact with generative AI systems.
- **Education.** Governments could educate the public about the potential risks and benefits of generative AI. This would help people understand how to use generative AI responsibly and to be aware of the potential harms that it could cause.
- **Research.** Governments could fund research into the development of generative AI that is safe and ethical. This would help to ensure that generative AI is used in a way that benefits society and does not harm individuals or groups.

Is anyone listening to these guys?¹⁴

Those who want to “pause” have their list, too:

- “Mandate robust third-party auditing and certification for specific AI systems.
- Regulate organizations’ access to computational power.
- Establish capable AI agencies at the national level.
- Establish liability for AI-caused harms.
- Introduce measures to prevent and track AI model leaks.
- Expand technical AI safety research funding.
- Develop standards for identifying and managing AI-generated content and recommendations.”¹⁵



PROGRESS OR PARALYSIS?

Regulatory lists are everywhere, but who’s actually regulating what?

In April 2023, the Chinese government released a draft set of regulations for GAI.¹⁶ These regulations would require providers of GAI services to take several steps to ensure that their products are used responsibly, including:

- Obtaining user consent before using their data to train GAI models
- Taking steps to prevent the generation of harmful or misleading content
- Implementing security measures to protect user data

The US government has not yet implemented any specific regulations on GAI, but there is a growing debate about the need for such regulations. Some experts argue that GAI poses myriad risks, such as its potential to be used to generate deepfakes or spread disinformation. Others argue that it has the potential to be used for good, such as to create educational content or help people with disabilities.

The EU has implemented a number of regulations that could impact the development and use of GAI. For example, the General Data Protection Regulation (GDPR) requires companies to obtain user consent before collecting or using their personal data. The GDPR also requires companies to take steps to protect user data from unauthorized access or use.

A post on New York University’s law blog notes that the Italian Data Protection Authority’s orders against OpenAI’s operations of ChatGPT in Italy highlighted tensions between the EU’s GDPR and GAI infrastructures trained on massive data sets involving both personal and nonpersonal data.¹⁷ The emergence of GAI infrastructures has led to rethinking in the EU’s proposed Artificial Intelligence Act, which aims for comprehensive, risk-based, product safety-based AI regulation. National agencies, including the Cyberspace Administration of China, are exploring new regulatory measures in this area. In regulation, licensing, contracts, and litigation, the allocation of risk and responsibilities along the GAI supply chain is vigorously in contention.

The US faces several somewhat unique regulatory challenges, ranging from the technology ignorance of lawmakers to lobbyists who own much of the legislative process, not to mention partisan politics and the relationships many US lawmakers have with the companies and industries they’re expected to regulate.

Although it’s impossible to predict whether the US will meaningfully regulate AI, ML, and GAI, there are signs that progress is at least possible. ChatGPT notes that:

The Algorithmic Accountability Act, which was reintroduced in Congress in 2022, would require companies to conduct impact assessments for certain high-risk AI systems, including generative AI, to identify and mitigate potential harms ... the bipartisan Artificial Intelligence Initiative Act, introduced in 2021, would provide funding for research and development of AI, including studies on the ethical, legal, and social implications of AI.

US states like California and New York may take the lead. It's possible that a bottom-up regulatory approach will be more effective than a federal top-down approach, but that remains to be seen. State-by-state regulations will complicate cross-border commerce, which is why a federal approach may be necessary. Partnerships with contiguous countries might offer some regulatory promise. For example, Canada introduced the Artificial Intelligence and Data Act in 2022, which could form the basis of a NAFTA-like agreement among Canada, Mexico, and the US.

There are some unusually challenging issues that may paralyze regulatory efforts simply because of their complexity. Should artists be compensated if GAI mimics their work? New challenges around copyright and intellectual property (IP) rights are far from understood. Compensation and ownership questions are complicated issues that are far from resolved. ChatGPT suggests that licensing payments be made to artists when works similar to their originals are created, shared, or published.

And it's not just about copyright.

There are additional challenges that must be managed. In her article "Generative AI Is a Legal Minefield," Axios Chief Technology Correspondent Ina Fried writes:

At issue is whether or not such training falls under a principle known as "fair use," the scope of which is currently under consideration by the Supreme Court. Much of the early legal battles have been about this issue. Getty, for example, is suing Stable Diffusion, saying the open source AI image generator trained its engine on 12 million images from Getty's database without getting permission or providing compensation. It's not just about copyright. In a lawsuit against GitHub, for example, the question is also whether the CoPilot system — which offers coders AI-generated help — violates the open source licenses that cover much of the code it was trained on.¹⁸

Nor are the potential IP infringement issues limited to the data that trains such systems. Many of today's GAI engines are prone to spitting out code, writing, and images that appear to directly copy from one specific work or several discernible ones.

The US National Institute of Standards (NIST) recently entered the regulatory picture by providing, as it always does, a set of suggestions about how to proceed with standards:

On March 30, NIST launched the Trustworthy and Responsible AI Resource Center, which will facilitate implementation of, and international alignment with, the AI RMF. On January 26, 2023, NIST released the AI Risk Management Framework (AI RMF 1.0) along with a companion NIST AI RMF Playbook, AI RMF Explainer Video, an AI RMF Roadmap, AI RMF Crosswalk, and various Perspectives.¹⁹

Technology is clearly moving faster than regulators can (or want to) move. And even as efforts are underway to regulate AI, ML, and GAI, there are also efforts to delay or avoid any kind of regulation. It's safe to say that the world is both confused and challenged by this technology. Many regulatory drafts have been developed and shared, but nothing is final. One especially challenging aspect of regulation is enforcement. What happens when some individual, company, or country violates the regulations?

CONCLUDING THOUGHTS

Regulatory action depends on how quickly the power of GAI is revealed. We know, for example, that orders of performance magnitude separate ChatGPT-3 from ChatGPT-4. What tasks and processes will ChatGPT-5 or -6 enable? As more industries, functions, and processes yield to LLMs, there will be additional pressure to regulate at some level. Of course, if there's sufficient coverage of GAI's limitations and a few high-profile regulations that quell the most serious fears, broader regulatory efforts will likely collapse.

Decisions around regulation will not be completely anchored in technology capabilities. Social, political, and economic concerns about the impact of regulation will exert as much, if not more, influence on whatever regulatory scenarios emerge. This changes the game, the players, and the rules. All of the activity around draft and proposed regulations will have several filters through which proposed regulations must pass. This means meaningful legislation will be slow to proceed. It's also likely that the US will lose the regulatory game to countries that are outpacing the US's regulatory efforts.

Predictions are impossible to make in areas as complicated as the regulation of AI, ML, and GAI, but it's safe to say there will be a lag between regulatory policy and the growing power of this technology. Regulations may lag applications for years and perhaps even permanently. This happens when technology moves as fast as intelligent systems technology is moving — and is likely to move in the future.

The old ways of treading lightly in the regulatory world will not work for GAI. This technology represents a sea change; treating it as just another incremental advance is a huge mistake. That warning aside, all of this assumes that there's a real desire to regulate the technology. Although there may be an honest desire to regulate the technology in several countries and a few US

states, it remains to be seen whether the US is capable of developing (and enforcing) impactful regulations for such a fast-moving technological target.

We cannot ignore AI, ML, and GAI. We should not love or hate them, either. The only answer is regulation, regardless of who takes the lead.

REFERENCES

- ¹ Toews, Rob. "[The Next Generation of Large Language Models.](#)" *Forbes*, 7 February 2023.
- ² Lantero, Allison. "[The History of Electricity at the White House.](#)" US Department of Energy, 14 October 2015.



- ³ Edwards, Phil. "[7 World-Changing Inventions People Thought Were Dumb Fads.](#)" Vox, 29 June 2015.
- ⁴ "[Pause Giant AI Experiments: An Open Letter.](#)" Future of Life Institute, 22 March 2023.
- ⁵ Loizos, Connie. "[1,100+ Notable Signatories Just Signed an Open Letter Asking 'All AI Labs to Immediately Pause for at Least 6 Months.'](#)" TechCrunch, 29 March 2023.
- ⁶ "[5 Practical Business Use Cases for Large Language Models.](#)" Open Data Science (ODSC), 2 March 2023.
- ⁷ Knutsson, Kurt. "[How Generative AI Could Cut Health Care Costs, Develop New Cancer Drugs.](#)" Fox News, 8 March 2023.
- ⁸ Jones, Brian, and Rod Fontecilla. "[Calling Dr. GPT: The Impact of Generative AI on Healthcare.](#)" Guidehouse, 9 February 2023.
- ⁹ Murphy, Patrick. "[The Role of Generative AI in Creating More Pleasant Cities.](#)" Maket, accessed August 2023.
- ¹⁰ Maharaj, Sahir. "[Generative AI: A Tool to Combat Climate Change.](#)" Medium, 7 February 2023.
- ¹¹ "[Generative AI in Product Design & Development — Benefits & Tips.](#)" RedBlink Technology, 23 March 2023.
- ¹² Afshar, Vala. "[How Can Generative AI Improve the Customer Experience?](#)" ZDNET, 1 February 2023.
- ¹³ Dans, Enrique. "[Why Do We Always See New Technology as a Threat?](#)" Medium, 13 April 2023.
- ¹⁴ Apparently, ChatGPT and Bard are male, but they also don't have to be.
- ¹⁵ "[Policy Making in the Pause.](#)" Future of Life Institute, 12 April 2023.
- ¹⁶ Luo, Yan, et al. "[China Proposes Draft Measure to Regulate Generative AI.](#)" Covington, 12 April 2023.
- ¹⁷ "[GPT, GDPR, AI Act: How \(Not\) to Regulate 'Generative AI'?](#)" *The Docket*, New York University School of Law, 19 April 2023.
- ¹⁸ Fried, Ina. "[Generative AI Is a Legal Minefield.](#)" Axios, 24 February 2023.
- ¹⁹ "[AI Risk Management Framework.](#)" National Institute of Standards and Technology (NIST), US Department of Commerce, accessed August 2023.

About the author

Stephen J. Andriole is a Fellow with Cutter Consortium, a member of Arthur D. Little's AMP open consulting network, and the Thomas G. Labrecque Professor of Business Technology at Villanova University. His specialty areas include digital transformation, emerging technology trends, cloud computing, social media, technology due diligence, software IP valuation, business technology strategy, business technology management, technology governance, business technology organization, the business value of technology, and technology performance management. He is an acclaimed columnist in *Forbes*. Dr. Andriole advises clients across the spectrum of business technology and has been a frequent Cutter author and keynoter since the late 1990s.

Dr. Andriole is the former Director of the Cybernetics Technology Office of the US Defense Advanced Research Projects Agency (DARPA). He served as CTO and Senior VP

of Safeguard Scientifics, Inc., where he was responsible for identifying technology trends, translating that insight into the Safeguard investment strategy, and leveraging trends analyses with Safeguard partners to help them develop business and marketing strategies. Dr. Andriole was also CTO and Senior VP for Technology Strategy at CIGNA Corporation. As an entrepreneur, Dr. Andriole founded International Information Systems (IIS), Inc., which designed interactive systems for a variety of corporate and government clients. He is also cofounder of The Acentio Group, a strategic consulting consortium that identifies and leverages technology trends to help clients optimize their business technology investments. Dr. Andriole is also former Professor of Information Systems and Electrical and Computer Engineering at Drexel University as well as a former Professor and Chair of the Department of Information. He can be reached at experts@cutter.com.

**LLM SECURITY
CONCERNS
SHINE A LIGHT
ON EXISTING DATA
VULNERABILITIES**

Authors

Michael Papadopoulos, Nicholas Johnson,
Michael Eiden, Philippe Monnot,
Foivos Christoulakis, and Greg Smith

In the rapidly evolving landscape of artificial intelligence (AI), large language models (LLMs) have emerged as a powerful tool, capable of generating human-like text responses, creating conversational interactions, and transforming the way we perceive and interact with technology.

Like all powerful tools, LLMs come with a set of security concerns. This article delves into those concerns, emphasizing that although LLMs certainly present novel security threats, the fundamental concerns, protections, and remedies remain similar to existing, well-understood information security challenges. In fact, characteristics of LLMs and their associated data pipelines allow more sophisticated and proportional security interventions, potentially leading to a better equilibrium between protection and benefit.

The first point to understand is that LLMs, by their nature, can only divulge information they were exposed to during their training phase. Thus, if an LLM reveals sensitive or private information, it's not because the model is inherently insecure — it's because it was given access to this information during its training. This highlights that the root of the problem is improper data access and management. Consequently, the focus should be on ensuring that data used to train these models is carefully curated and managed in order to prevent any potential downstream data leaks.

However, managing the training data is just one part of the equation. Even with the best data management practices, an LLM might still generate inappropriate or harmful content based on the patterns it learned during training. This is where the implementation of an LLM module, coupled with strategic prompt engineering, can serve as a robust, layered security mechanism.

Prompt engineering involves carefully crafting the prompts that are given to the LLM to guide it toward generating the desired output. By scrutinizing both the inputs (user prompts) and the outputs of the LLM, we can establish a multitiered safety environment that can effectively mitigate security risks. For instance, an LLM module can be designed to reject certain types of prompts that are likely to lead to harmful outputs, and it can be programmed to filter out any potentially harmful content from the LLM's responses.

MANAGING THE TRAINING DATA IS JUST ONE PART OF THE EQUATION

This approach to security doesn't just protect against the known risks associated with LLMs, it provides a framework for identifying and mitigating new risks as they emerge. It's a dynamic, adaptable approach that can evolve alongside the LLMs. Indeed, the pace of innovation within the LLM and wider language-processing domain ensures that any security approach not based on continuous sensing, analyzing, adapting, and iterating is doomed to failure.

It's important to be aware of the security concerns associated with LLMs, but it's equally important to understand that these concerns are new manifestations of existing security threats and thus manageable. With proper data handling and innovative security strategies, we can harness the full potential of these powerful AI tools without compromising safety or security.

TOP 10 LLM SECURITY CONCERNS

The exponential integration of LLMs within organizations holds the promise of seamless automation and enhanced efficiencies. However, with these advancements come unique security challenges.

Our research and use in the field have yielded a top 10 list of vulnerabilities that pose either new threat vectors or new context for typical vulnerabilities to be exploited or manipulated in an LLM context (see Figure 1):

1. Prompt injection
2. Insecure output handling
3. Training data poisoning
4. Model denial of service
5. Supply chain vulnerabilities
6. Data leakage/sensitive information disclosure
7. Insecure plug-in design
8. Excessive agency
9. Overreliance
10. Model theft

PROMPT INJECTION

The age-old tactic of manipulating systems through cunning inputs finds its way to LLMs. Attackers craftily modify the prompts fed into the model, leading to unintended actions. There are two primary avenues for these attacks: (1) direct injections involve overriding the system prompts and (2) indirect ones alter the inputs from external sources. These can compromise the integrity of the LLM's response and, subsequently, the systems relying on it.

To remediate prompt injection attacks, users must validate and sanitize all inputs before they're processed. Simultaneously, they should maintain a white list of accepted commands to aid in filtering out malicious inputs. Regular monitoring and logging of prompts become vital to detect and address unusual patterns swiftly, and it's beneficial to limit the amount of user-defined input that an LLM can process. Finally, introducing a system of regular user feedback can help fine-tune the model's responsiveness to malicious prompts.

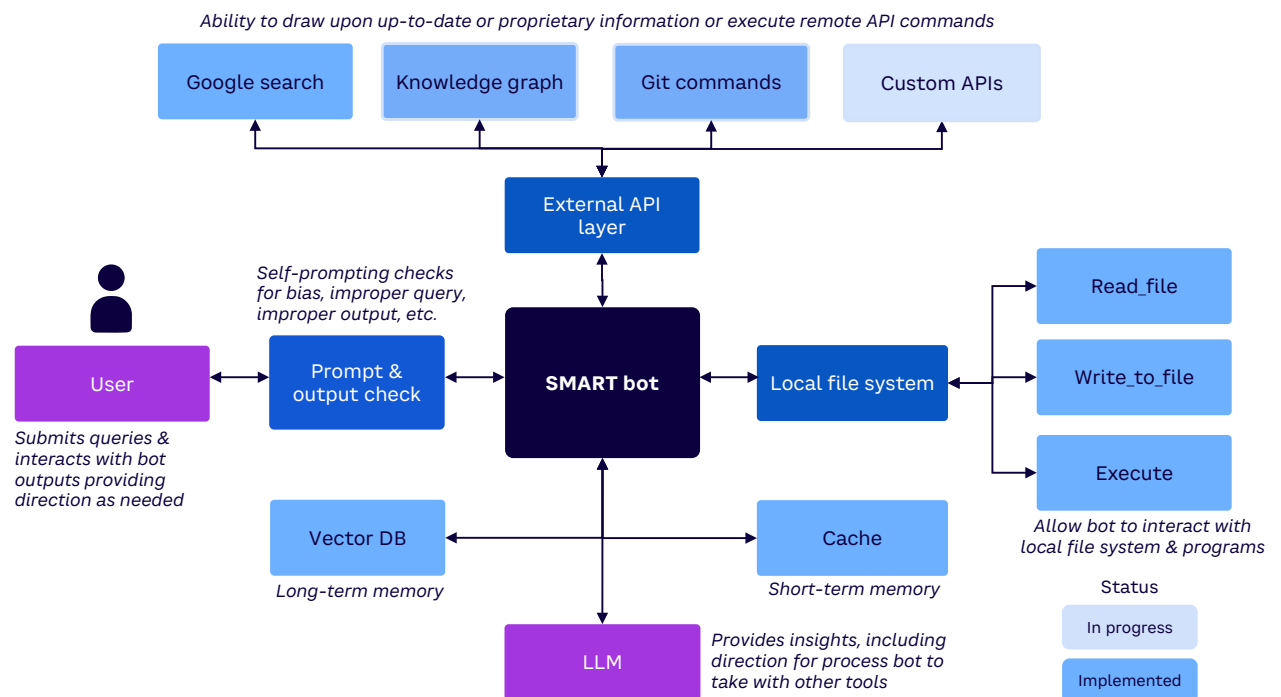


Figure 1. LLM application architecture: remediation of most common vulnerability issues (source: Arthur D. Little)

INSECURE OUTPUT HANDLING

LLMs can produce a wide variety of outputs. Accepting these without proper verification opens the gates for multiple threats, including XSS (cross-site scripting), CSRF (cross-site request forgery), and SSRF (server-side request forgery). Moreover, privilege escalation or remote code execution becomes feasible, posing an enormous security risk to the back-end systems that treat the output as safe.

Proactive measures like output sanitization, strict validation, and monitoring should be established to prevent privilege escalation and remote code execution. This will ensure the consistent security of responses generated by LLMs.

TRAINING DATA POISONING

Training data is the backbone of any LLM. However, when this data is compromised or injected with malicious intent, the resultant LLM can exhibit vulnerabilities or biases. This can weaken the model's security, overall effectiveness, and even ethical behavior.

To remediate training data-poisoning attacks in LLMs, it's crucial to prioritize the integrity of the training data by sourcing it exclusively from reputable sources and meticulously validating its quality. Applying rigorous data-sanitization and preprocessing techniques is essential to weed out potential vulnerabilities or biases inherent in the data. It's also beneficial to conduct periodic reviews and audits of the LLM's training data and its fine-tuning processes. Finally, incorporating monitoring and alerting systems can be invaluable in identifying any unusual behavior or performance anomalies, further bolstering the model's security.

MODEL DENIAL OF SERVICE

The resource-intensive nature of LLMs makes them susceptible to denial-of-service attacks. Perpetrators can introduce resource-heavy operations, overburdening an LLM, causing either service degradation or unexpectedly high operational costs.

To counteract these attacks, it's essential to implement rate-limiting measures and monitor user inputs for resource-heavy operations. By managing the workload and detecting unusual spikes in resource usage, organizations can maintain optimal LLM performance and prevent excessive operational costs.

SUPPLY CHAIN VULNERABILITIES

The lifecycle of LLM applications involves data sets, pretrained models, plug-ins, and more. Introducing vulnerabilities at any of these stages can compromise the entire model, making it an attractive target for attackers. To secure the LLM application lifecycle, conduct regular audits of all components. Employing stringent validation and vetting processes during integration will safeguard the model, reducing its susceptibility to external threats.

TRAINING DATA IS THE BACKBONE OF ANY LLM

DATA LEAKAGE/SENSITIVE INFORMATION DISCLOSURE

LLMs, while sophisticated, may unintentionally leak confidential information through their responses. This can lead to unauthorized data access, breaches, and severe privacy violations. Organizations must stress data sanitization and user policies to circumvent such exposures. We have found that using a secondary LLM to test the outputs for sensitive information is an excellent way to help ensure security.

INSECURE PLUG-IN DESIGN

LLMs often incorporate plug-ins to enhance functionality. However, if these plug-ins have insecure input mechanisms or flawed access controls, they become glaring vulnerabilities. Exploiting them might result in grave consequences, including remote code execution.

To mitigate vulnerabilities in LLM plug-ins, ensure rigorous vetting before integration. Prioritize plug-ins with robust input validation and stringent access controls. Regular security audits of plug-ins can also help detect and rectify potential weak points, preventing potential exploits.

EXCESSIVE AGENCY

Assigning excessive permissions, functionality, or autonomy to LLMs can spell disaster. Such models can autonomously make decisions, potentially leading to significant unintended consequences. This issue emphasizes the need for setting boundaries for LLM-based systems.

To safeguard against overpowered LLMs, it's imperative to implement a permissions framework, limiting the LLM's functionality and autonomy. Regularly review and adjust these permissions to strike a balance between operational efficiency and control to ensure LLMs function within defined boundaries.

OVERRELIANCE

Reliance on LLMs without human oversight is a treacherous path. Such "blind reliance" can lead to misinformation, legal conundrums, and a host of security vulnerabilities, mainly if the LLM churns out incorrect or inappropriate content.



To counteract this risk, introduce human oversight in critical decision-making processes. Establishing a hybrid system, in which human experts review and validate LLM outputs, can reduce misinformation risks, address potential legal issues, and bolster overall security against inappropriate content generation.

MODEL THEFT

Proprietary LLMs are of immense value. Unauthorized access or exfiltration can cause substantial economic losses, erode competitive advantages, and even expose sensitive information. Ensuring stringent security protocols is paramount to prevent such incidents.

To protect proprietary LLMs, deploy multilayered security measures, including encryption, access controls, and regular audits. By closely monitoring system activity and restricting unauthorized access, organizations can safeguard their valuable assets, preserving both competitive advantage and data confidentiality.

The era of LLMs is transformative, heralding countless possibilities. However, navigating this landscape requires organizations to be acutely aware of the inherent security challenges. Addressing these concerns head-on will ensure a future where LLMs can be harnessed safely and efficiently.

HOW LLMs CAN IMPROVE SECURITY

Rather than introducing wholly unprecedented threats into society, LLMs highlight and stress test existing vulnerabilities in how organizations govern data, manage access, and configure systems. With care and responsibility, we can respond to their revelations by engineering solutions that make technology usage more secure and ethical overall.

Specific ways responsible LLM adoption can improve security include:

- **Automated vulnerability scanning.** Leverage LLM conversational ability to identify flaws in public-facing chat interfaces.
- **Anomaly detection.** Monitor corporate system logs with LLMs fine-tuned to flag unusual internal events as possible attacks.
- **Safety analysis.** Stress test new features through automated conversational exploration of potential abuses.
- **Product-security reviews.** Use LLMs as a team member when designing new products to probe attack possibilities in simulated conversations.
- **Threat intelligence.** Continuously train LLMs on emerging attack data to profile bad actors and model potential techniques.
- **Forensic reconstruction.** Assist investigations of past incidents by using LLMs to speculate about criminal conversations and motives.
- **Security policy analysis.** Check that policies adequately address LLM-relevant risks revealed through conversational probing.
- **Security training.** Use LLM-generated attack scenarios and incidents to build staff defensive skills.
- **Bug bounties.** Expand scope of bounty programs to include misuse cases identified through simulated LLM hacking.

With careful design and effective oversight, LLMs can be an ally rather than a liability in securing organizations against modern technological threats. Their partially open nature invites probing for weaknesses in a controlled setting.

**THE ERA
OF LLMs IS
TRANSFORMATIVE,
HERALDING
COUNTLESS
POSSIBILITIES**

LLMs present a further opportunity to improve an organization's information security capability. The practical application of LLMs to business challenges requires creating sophisticated, multistage, software-driven data pipelines. As these pipelines start to become prevalent, an opportunity to design with more effective security protocols is presented.

Various security postures can be applied at different points in the pipeline. For instance, a permissive security posture that allows an LLM to generate the best possible response can be followed by a more restrictive security filter that automatically checks the output for potential data leakage.

If we accept that LLM security problems are new manifestations of existing information security challenges (and that human behavior is the biggest cause of security breaches), then automated multistage processes with carefully constructed security gateways can provide a powerful new tool in the toolkit.

CONCLUSION

LLMs, such as GPT-4, represent a breakthrough in language-capable AI, but commentary casting their risks as wholly unprecedented is overstated. A closer look reveals that concerns around their potential for data exposure and security issues/bias largely echo existing vulnerabilities, often exacerbated by poor underlying security and data governance practices.

Rather than engaging in an ultimately futile battle to ban promising AI innovations, the responsible path is to address underlying root causes. The route to achieving this is well understood but often poorly implemented, requiring organizations to take a systematic and pragmatic approach to security, including better aligning access controls, tightening monitoring, enhancing information literacy, and ensuring effective oversight. LLMs can even assist in this by stress testing systems and uncovering policy gaps through exploratory conversation.

The emerging technology does not intrinsically undermine safety — it shines a light on long-standing cracks that ought to be sealed and has the potential to enhance security.

About the authors

Michael Papadopoulos is a Cutter Expert, Chief Architect of Arthur D. Little's (ADL's) UK Digital Problem Solving practice, and a member of ADL's AMP open consulting network. He is passionate about designing the right solutions using smart-stitching approaches, even when elegance and architectural purity are overshadowed by practicality. Mr. Papadopoulos leads the scaling of multidisciplinary organizations by focusing on continuous improvement, establishing quality standards, and following solid software engineering practices. He mentors team members, leaders, and managers along the way. Mr. Papadopoulos is a strong advocate of the DevOps culture and Agile principles and has demonstrated experience in solving problems in challenging global environments. Coming from a development background, he remains highly technical, with hands-on involvement in code review, design, architecture, and operations. Mr. Papadopoulos has more than 15 years' experience in technology and digital consulting and has worked in a variety of sectors, including telecom, gaming, energy, and media. He can be reached at experts@cutter.com.

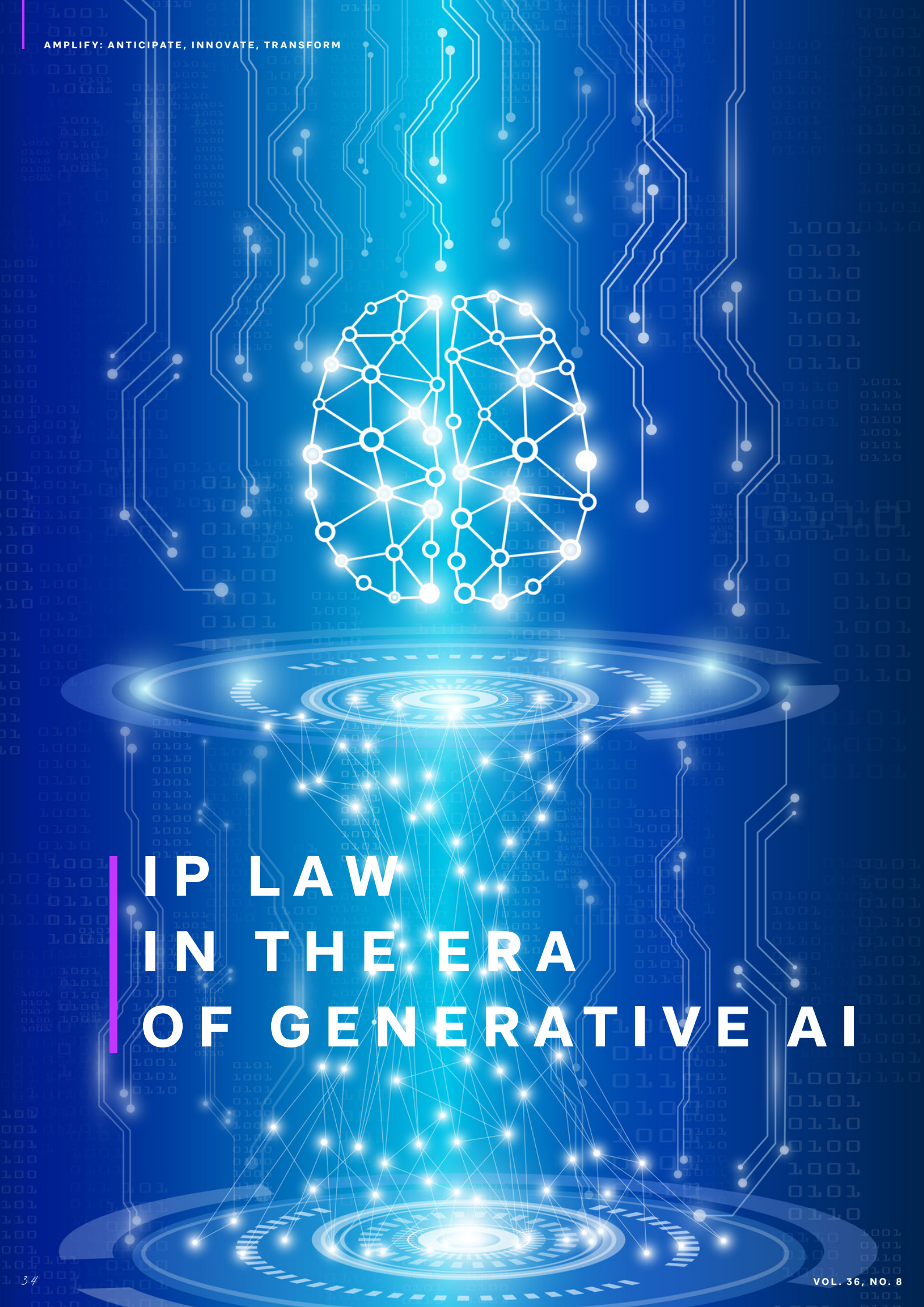
Nicholas Johnson is a Partner with ADL's Digital Problem Solving practice, based in London. He focuses on how emerging digital technologies can be harnessed to drive transformation of both the business and its internal technology function. Mr. Johnson believes that the technology patterns/approaches used in the past are no longer appropriate to today's challenges, and that businesses must adopt new approaches. With over 20 years in technology consulting, he has worked across a wide range of sectors, including telco, media, transport, logistics, gaming, and energy, having considerable experience delivering global digital transformation initiatives, often from initial idea conception through full product launch. Mr. Johnson also has extensive experience in managing large-scale Agile projects and proof of concepts. He earned a bachelor of science degree from Cardiff University, UK, and a master of science degree in computer science, with distinction, from the University of Bath, UK. He can be reached at experts@cutter.com.

Michael Eiden is a Cutter Expert, Partner and Global Head of AI & ML at ADL, and a member of ADL's AMP open consulting network. Dr. Eiden is an expert in machine learning (ML) and artificial intelligence (AI) with more than 15 years' experience across different industrial sectors. He has designed, implemented, and productionized ML/AI solutions for applications in medical diagnostics, pharma, biodefense, and consumer electronics. Dr. Eiden brings along deep expertise in applying supervised, unsupervised, as well as reinforcement ML methodologies to a very diverse set of complex problem types. He has worked in various global technology hubs, such as Heidelberg (Germany), Cambridge (UK), and Silicon Valley (US), with clients ranging from small and medium-sized enterprises to globally active organizations. Dr. Eiden earned a doctorate in bioinformatics. He can be reached at experts@cutter.com.

Philippe Monnot is a Data Scientist with ADL's Digital Problem Solving practice, and a member of ADL's AMP open consulting network. He's passionate about solving complex challenges that impact people's livelihood through the use of data, statistics, and ML. Mr. Monnot enjoys developing accessible solutions that customers will adopt through effective data storytelling and explainable AI. Before joining ADL, he worked in R&D, where he used ML to implement smart, scalable manufacturing processes to manufacture sustainable composite structures for the aerospace and oil and gas industries. He can be reached at experts@cutter.com.

Foivos Christoulakis is a Solutions Architect with ADL's Digital Problem Solving practice and a member of ADL's AMP open consulting network. He is a passionate cloud architect who has designed and implemented numerous solutions currently in production in global-scale organizations. Mr. Christoulakis helps organizations grow by focusing on high engineering standards and following solid software engineering practices. He continues to be a strong advocate of DevOps and Agile principles and showcases both skill sets mentoring and being a servant leader for multiple DevOps and architecture teams. Mr. Christoulakis has more than 10 years' experience in cloud architectures across many business verticals, including telecoms, entertainment, and software development. He can be reached at experts@cutter.com.

Greg Smith is Managing Partner of ADL. He founded and co-leads ADL's Digital Problem Solving practice and is a member of ADL's Executive Committee, where he has responsibility for ADL's global innovation strategy. His work focuses on business strategy in the context of digital transformation as well as the application of disruptive information technologies in solving intractable business problems in major enterprises. Recently, Mr. Smith has been focusing on digital operating models for established businesses, including the changes required in technology, culture, IT functions, business technology interactions, organizational design, and governance, and how these can combine to enhance customer and service experience. He holds a bachelor of science degree in biological sciences from the University of Leicester, UK, and finds that after 30 years of dormancy within his professional life, the underlying concepts of biology are becoming increasingly valuable at unlocking business problems and articulating solutions — especially where reductive, engineering-based approaches need to be replaced with whole-system, evolutionary thinking. He can be reached at experts@cutter.com.



IP LAW IN THE ERA OF GENERATIVE AI

Authors

Ryan Abbott and Elizabeth Rothman

Technological evolution and increased artificial intelligence (AI) adoption are driving interest in the challenges AI poses for legal frameworks designed to regulate human behaviors. In the context of intellectual property (IP) law, many tasks done by today's AI are longstanding AI capabilities, but the technology has improved to the point where they have meaningful commercial value.

Those improved capabilities are increasing concerns about the protectability of AI output, deepfakes, privacy, and the use of copyright-protected content for training AI systems. This article focuses on AI-generated output created without a traditional human author or inventor and examines whether this output is protectable under current laws. It also describes a series of legal test cases put forth by the Artificial Inventor Project (AIP).¹

IS AI MERELY A TOOL USED BY HUMANS?

Nearly seven decades after the term “artificial intelligence” was coined, it lacks a uniform definition — and the need to define AI has now departed the realm of academia. This lack has neither impeded engineers nor businesses from developing and employing AI, but it poses substantial challenges for legislation like EU’s AI Act, aimed at governing AI use and development. The same applies to defining AI-generated output, which is muddled by ambiguous terms like “autonomous” and “tool” commonly used to describe AI abilities and functions.

Here, we use “AI” to refer to an algorithm or machine capable of completing tasks that would otherwise require cognition. We use “AI-generated” works and inventions to refer to output made without a traditional human author or inventor.² AI systems are developed and directed by humans but can, to varying degrees, automate tasks and make creative or technical decisions.

Some argue that AI systems are just tools used by people to generate works and inventions, not different in kind from a pencil or a microscope.³ In this view, AI output is simply a natural extension of human creativity. Others argue that, in some instances, AI is stepping into the shoes of natural persons and automating tasks that would traditionally make a person an author or inventor, including in cases where no natural person is acting as an author or inventor. If this view is correct, it may not be possible to obtain copyright and patent protection for AI-generated output in jurisdictions that require the involvement of a human author or inventor.

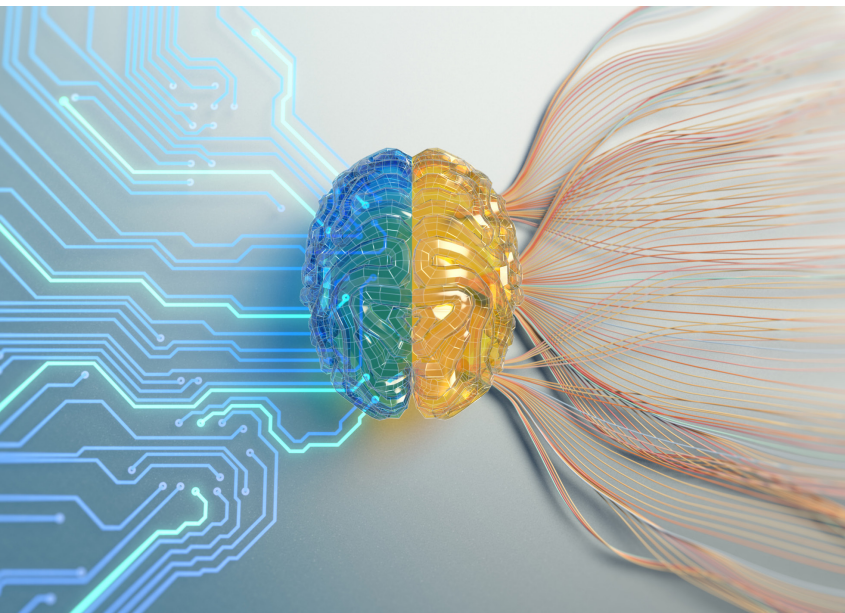
Policymakers have become acutely aware of the disruptive nature of AI and are actively debating how best to regulate it. In the US, the Senate Subcommittee on Intellectual Property held a hearing in June 2023 on AI and patents as part of a series on AI and IP.⁴ The US Patent and Trademark Office issued a “Request for Comments” on AI and patent inventorship in February 2023.⁵

The EU is creating a new legal framework specifically targeted to AI: the AI Act.⁶ The UK Intellectual Property Office (UKIPO) conducted two consultations on IP and AI as part of its National AI Strategy.⁷ The World Intellectual Property Organization (WIPO), the UN agency primarily responsible for IP matters, is conducting a series of meetings on IP and AI as part of its “WIPO Conversations on IP and Frontier Technologies.” WIPO describes the conversations as “an open, inclusive, multi-stakeholder forum designed to provide stakeholders with a leading, global setting

to discuss the impact of frontier technologies on all IP rights and to bridge the existing information gap in this fast-moving and complex field.”⁸

AUTHORSHIP & INVENTORSHIP IN IP

One of the primary challenges in determining whether AI-generated output is eligible for protection is that many existing legal systems have, to some extent, framed authorship and inventorship in terms of human activity. Sometimes this was the result of historical assumptions that only a person could be an author or inventor. Some commentators say this is still a desirable position on the basis that IP laws should exist to incent certain human behaviors or protect human moral rights. Others argue that IP laws exist primarily to benefit the public by encouraging the creation and dissemination of IP and thus should be agnostic as to the manner in which protectable output is generated.



Establishing IP rights for AI-generated output may also raise new questions about ownership. Whereas authors and inventors are sometimes the first-instance owners of their output, AI systems, lacking in legal personality, cannot own IP. It thus becomes necessary to determine who should hold the corresponding rights — including in cases where multiple parties may have a claim to ownership, such as those programming or training AI systems, using AI systems, or having ownership rights in AI systems. Adding further complexity, IP

law is governed by a complex web of national and international rules.

THE ARTIFICIAL INVENTOR PROJECT

AIP is a global initiative that:

... includes a series of pro-bono legal test cases seeking intellectual property rights for AI-generated output in the absence of a traditional human inventor or author. It is intended to promote dialogue about the social, economic, and legal impact of frontier technologies such as AI and to generate stakeholder guidance on the protectability of AI-generated output.⁹

In 2018, AIP filed two patent applications for AI-generated inventions to UKIPO and the European Patent Office (EPO).¹⁰ The applications pertained to a beverage container based on fractal geometry and an emergency beacon, both devised by an AI system called DABUS (Device for the Autonomous Bootstrapping of Unified Sentience), which is created, owned, and operated by Dr. Stephen Thaler.¹¹ DABUS was not provided with a specific problem to solve, and Thaler lacked technical expertise in the fields of DABUS’s output.¹²

The patent applications successfully passed a preliminary substantive examination by UKIPO, after which an inventorship designation was filed to disclose that the inventor for both applications was DABUS rather than a natural person. Thaler was listed as the patent applicant and thus owner of any property rights in the applications and any future issued patents.

From there, an “international” patent application was filed under the Patent Cooperation Treaty,¹³ and the application was ultimately nationalized in 18 jurisdictions: Australia, Brazil, Canada, China, EPO, Germany, India, Israel, Japan, New Zealand, Republic of Korea, Saudi Arabia, Singapore, South Africa, Switzerland, Taiwan, the UK, and the US.¹⁴

DABUS was named as the legal inventor, not to grant rights to a machine (AI lacks legal personality and rights) but because it was the factual inventor and to maintain transparency about the origin of the inventions. Although DABUS is the first AI inventor to be listed on an application, it is not unusual for human inventors to have no rights to patents on their inventions. Most patents are owned by artificial entities, such as corporations, which typically acquire these rights through employment relationships.

No statute explicitly governs the ownership of patents on AI-generated inventions, but longstanding property law principles suggest that the AI's owner should own these patents. For example, the owner of a 3D printer owns the physical objects produced by his or her machine. In some jurisdictions, people who employ computer software and hardware to generate (aka "mine") cryptocurrencies own that cryptocurrency.

These examples follow a rule sometimes referred to as "accession," which dates back to at least Roman law and dictates that property owners own property created by their property, whether it is fruit from a tree or a calf from a cow. Similarly, if someone owns a machine, he or she should own the output of that machine, whether tangible or intangible. This outcome is crucial given the intangible, non-rivalrous nature of AI output.

Securing patents on AI-generated inventions is essential for achieving patent law's policy objectives. It fosters innovation by encouraging the use and development of inventive AI, since the output generated by these machines will be more valuable. It also promotes public disclosure of AI-generated inventions that could be protected as trade secrets or confidential information and incents the commercialization of inventions.

In some industries, where patent protection is critical to business models like drug development, most costs associated with an invention arise not from initial discovery but from subsequent development expenses required to turn a discovery into a marketable product. In the case of new drugs, clinical validation can cost hundreds of millions of dollars. Allowing patents on AI-generated inventions enables businesses that rely on patent protection to automate aspects of R&D, without facing legal penalties, if it offers a technical advantage.

So far, AIP applications have received final, non-appealable denials in three jurisdictions: the US, Australia, and Taiwan.¹⁵ South Africa granted the patents,¹⁶ and the patent office of Saudi Arabia has accepted the designation of DABUS as the inventor and is conducting substantive examination. In five other jurisdictions (EPO, Germany, Israel, New Zealand, and the UK), the applications are undergoing judicial appeals after being rejected by patent offices. Of note, the UK Supreme Court heard oral arguments in the case in March 2023.¹⁷ In the remaining jurisdictions, the applications either await examination or have

received preliminary rejections from patent offices that are still being internally appealed.

In addition to the patent filings, in November 2018, AIP applied to register an artwork created by DABUS with the US Copyright Office (USCO).¹⁸ The application identified the AI as the author and sought ownership rights for Thaler as the AI's owner. However, on 14 February 2022, USCO issued a final rejection based on its policy that human authorship is a prerequisite for copyright protection in the US and said the work could not be registered.¹⁹

On behalf of Thaler, AIP filed a lawsuit against USCO on 2 June 2022, arguing that the Copyright Act permits copyright for AI-generated works and AI authorship and that the owner of an AI should be entitled to copyright in an AI-generated work.²⁰ While that case makes its way through the courts, on 16 March 2023, USCO released "Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence," which stated that "if a work's traditional elements of authorship were produced by a machine, the work lacks human authorship and the Office will not register it."²¹ They continued discussing AI-assisted works, saying:

When an AI technology determines the expressive elements of its output, the generated material is not the product of human authorship. As a result, that material is not protected by copyright and must be disclaimed in a registration application.²²

As generative AI becomes integrated into widely used consumer software like Microsoft Word, Gmail, and Adobe Photoshop, this level of parsing of human versus machine involvement will be increasingly difficult. We will likely see a continued reevaluation of this policy as the ubiquitous nature of these technologies becomes clear.

Unlike the US, the UK and other jurisdictions explicitly allow the protection of AI-generated works. In the UK, such works receive a limited term of protection compared to traditional author works (50 years from the year of creation versus the life of an author plus 70 years), and the producer of the work is legally deemed or fictionalized to be the author.²³ Since the UK elected to protect AI-generated works under Section 93 of the 1988 Copyright, Designs and Patents Act, the law has only been at issue in one case, and even then only tangentially, as no party was challenging the subsistence of the underlying work.²⁴

CONCLUDING THOUGHTS

As AI systems advance and produce increasingly sophisticated and innovative output, the question of how to treat this output under IP law becomes more pressing. The characteristics of some AI systems, including the self-improving nature of certain AI models and the difficulties associated with attributing their outputs to human creators, challenge the existing framework and necessitate a thorough rethinking of what rules will result in the greatest social value.

The treatment of AI-generated output under IP law is complex and evolving. As various jurisdictions grapple with the challenges presented by AI-generated works and inventions, international harmonization and cooperation will be crucial to develop a coherent and efficient legal framework.

Ultimately, reconciling the diverse approaches and addressing the legal, ethical, and economic implications of AI-generated output will be essential to foster innovation, promote the responsible use of AI, and ensure the equitable distribution of the benefits arising from AI-generated works and inventions.

REFERENCES

- ¹ This article is partially adapted from a recent article in the *European Intellectual Property Review*; see: Abbott, Ryan, and Elizabeth Rothman. "[AI-Generated Output and Intellectual Property Rights: Takeaways from the Artificial Inventor Project](#)." *European Intellectual Property Review*, January 2023.
- ² Abbott, Ryan. *The Reasonable Robot: Artificial Intelligence and the Law*. Cambridge University Press, 2020.
- ³ See, for example: Salsberg, Corey. "[Novartis Comments on Artificial Intelligence and Inventorship \(Docket PTO-P-2022-0045, 88 FR 9492\)](#)." Novartis, 15 May 2023.
- ⁴ "[Artificial Intelligence and Intellectual Property — Part I: Patents, Innovation, and Competition](#)." Subcommittee on Intellectual Property, US Senate Committee on the Judiciary, 7 June 2023.
- ⁵ Patent and Trademark Office, US Department of Commerce. "[Request for Comments Regarding Artificial Intelligence and Inventorship](#)." *Federal Register*, Vol. 88, No. 30, February 2023.
- ⁶ "[Proposal for a Regulation of the European Parliament and of the Council: Laying Down Harmonised Rules on Artificial Intelligence \(Artificial Intelligence Act\) and Amending Certain Union Legislative Acts](#)." European Commission, 21 April 2021.
- ⁷ UK Department for Science, Innovation and Technology; Office for Artificial Intelligence; Department for Digital, Culture, Media & Sport; and Department for Business, Energy & Industrial Strategy. "[National AI Strategy](#)." GOV.UK, 22 September 2021.
- ⁸ "[Intellectual Property and Frontier Technologies](#)." World Intellectual Property Organization (WIPO), accessed August 2023.
- ⁹ [The Artificial Inventor Project](#) website, accessed August 2023.
- ¹⁰ Together with the patent cases, a copyright case was filed in the US to register copyright in an AI-generated work. The US Copyright Office (USCO) rejected the registration; this is the subject of a pending case in US federal court; see: "[Thaler v. Perlmutter, et al.](#)" US District Court for the District of Washington DC, 2 June 2022.
- ¹¹ "[Thaler v. Vidal](#)." US Court of Appeals for the Federal Circuit, 5 August 2022.
- ¹² For a detailed technical explanation of how DABUS invented the beverage container and emergency beacon, see: Thaler, Stephen L. "[Vast Topological Learning and Sentient AGI](#)." *Journal of Artificial Intelligence and Consciousness*, Vol. 8, No. 1, 2021.
- ¹³ "[PCT — The International Patent System](#)." World Intellectual Property Organization (WIPO), accessed August 2023.
- ¹⁴ The Artificial Inventor Project ([see 9](#)).
- ¹⁵ An updated list of jurisdictions, file/case numbers, and outcomes can be found at The Artificial Inventor Project website ([see 9](#)).

- ¹⁶ South Africa Companies and Intellectual Property Commission. "[Food Container and Devices and Methods for Attracting Enhanced Attention.](#)" *Patent Journal*, Vol. 54, No. 07, July 2021.
- ¹⁷ "[Thaler v. Comptroller-General of Patents, Designs and Trademarks.](#)" UK Supreme Court, 2 March 2023.
- ¹⁸ "Thaler v. Perlmutter, et al" ([see 10](#)).
- ¹⁹ Perlmutter, Shira. "[Re: Second Request for Reconsideration for Refusal to Register A Recent Entrance to Paradise \(Correspondence ID 1-3ZPC6C3; SR # 1-7100387071\).](#)" Copyright Review Board, US Copyright Office (USCO), 14 February 2022.
- ²⁰ "Thaler v. Perlmutter, et al" ([see 10](#)).
- ²¹ US Copyright Office (USCO)/Library of Congress. "[Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence.](#)" *Federal Register*, Vol. 88, No. 51, March 2023.
- ²² US Copyright Office (USCO)/Library of Congress ([see 21](#)).
- ²³ "[Copyright, Designs and Patents Act 1988.](#)" [Legislation.gov.uk/The National Archives](#), accessed July 2023.
- ²⁴ "[Nova Productions Ltd. v. Mazooma Games Ltd. & Ors.](#)" England and Wales Court of Appeals, 14 March 2007.

About the authors

Ryan Abbott, MD, JD, MTOM, PhD, is Professor of Law and Health Sciences at the University of Surrey School of Law (UK); Adjunct Assistant Professor of Medicine at the David Geffen School of Medicine at UCLA; Partner at Brown, Neri, Smith & Khan, LLP; and a mediator and arbitrator with JAMS. He is author of *The Reasonable Robot: Artificial Intelligence and the Law* as well as Editor of *Research Handbook on Intellectual Property and Artificial Intelligence*. Prof. Abbott has published widely on issues associated with life sciences and intellectual property in leading legal, medical, and scientific books and journals, and his research has been featured prominently in the popular press, including *The Times of London*, *New York Times*, *Financial Times*, and other media outlets. He has worked as an expert for, among others, UK Parliament, European Commission, World Health Organization (WHO), and World Intellectual Property Organization (WIPO). Prof. Abbott is a licensed physician and patent attorney in the US and a solicitor advocate in England and Wales. He can be reached at drryanabbott@gmail.com or via www.ryanabbott.com.

Elizabeth Rothman is an attorney, author, and speaker based in Los Angeles, California, USA. With more than a decade of legal experience, she specializes in intellectual property, healthcare, and the emerging technology space. Ms. Rothman advises businesses, funds, and policymakers on the implications of artificial intelligence, blockchain, and extended reality technologies. In addition to her law practice, she is currently an advisor for Cantellus Group and the XR Safety Initiative, a global nonprofit standards developing organization. Ms. Rothman is an instructor for Harvard Law School and the World Intellectual Property Organization (WIPO) PatentX course "Patent Law and Global Public Health." She is also a member of the board of directors for the Metaverse Standards Forum. Ms. Rothman is licensed to practice law in the US and utilizes her legal and technical knowledge to bridge the gaps between often disparate disciplines in the technology space. She holds a bachelor's degree from Hunter College and a JD from Lewis & Clark Law School. She can be reached at liz@futuretechlegal.com or www.elizabethrothman.com.

ENVIRONMENTAL IMPACT OF LARGE LANGUAGE MODELS

Authors

Greg Smith, Michael Bateman, Remy Gillet,
and Eystein Thanisch

“The era of global warming has ended; the era of global boiling has arrived,” comes the stark warning from United Nations Secretary-General António Guterres.¹

“I knew I had just seen the most important advance in technology since the graphical user interface,” Microsoft cofounder Bill Gates wrote of seeing an early demo of ChatGPT in September 2022. “Entire industries will reorient around [artificial intelligence].”²

Will this reorientation of entire industries around powerful and power-hungry large language models (LLMs) push us deeper into global boiling? Will Stable Diffusion destabilize our climate even further? Will Google’s PaLM leave us all living in a desert? Will Meta’s newly released Llama 2 spit in the face of net zero?

Exactly how power-hungry are these models, anyway?

Modern LLMs require huge amounts of computing power to churn through huge amounts of data. Leaked estimates about GPT-4 (the latest LLM from OpenAI) put it at consuming trillions of words of text over a three-month-long training process that required up to 25,000 state-of-the-art graphics processing units (GPU) from industry leader Nvidia. OpenAI CEO Sam Altman put the monetary cost at more than US \$100 million.³ The breakdown of this figure is not known, but back-of-the-envelope calculations imply that energy costs alone could account for almost \$10 million.⁴

How much energy does the training process require? How much energy is needed to continuously serve an LLM via ChatGPT to serve billions of page views each month? How do these numbers stack up against the energy required to fly, use the oven, or stream a video?

In this article, we dissect the carbon footprint of LLMs, compare this footprint to familiar activities, highlight ways to minimize their impact, and evaluate the green credentials of some major players behind the recent wave of generative artificial intelligence (GAI).

WHAT IS THE ENVIRONMENTAL IMPACT?

Environmental impacts come in many forms, but here we focus primarily on carbon dioxide equivalent (CO₂e) emissions, since CO₂ is the main greenhouse gas (GHG) causing global warming and the biggest threat to the environment.

The carbon emissions from an LLM primarily come from two phases: (1) the up-front cost to build the model (the training cost) and (2) the cost to operate the model on an ongoing basis (the inference cost).

The up-front costs include the emissions generated to manufacture the relevant hardware (embodied carbon) and the cost to run that hardware during the training procedure, both while the machines are operating at full capacity (dynamic computing) and while they are not (idle computing). The best estimate of the dynamic computing cost in the case of GPT-3, the model behind the original ChatGPT, is approximately 1,287,000 kWh (kilowatt-hours), or 552 tonnes (metric tons) of CO₂e.

We will put this number into a fuller context in the next section, but we mention here that this figure is approximately the same emissions as two or three full Boeing 767s flying round-trip from New York City to San Francisco. Figures for the training of Llama 2 are similar: 1,273,000 kWh, with 539 tonnes of CO₂e.⁵ Analysis of the open source model BLOOM suggests that accounting for idle computing and embodied carbon could double this requirement.⁶

The ongoing usage costs do not include any additional embodied carbon (e.g., from manufacturing the computers, which have been accounted for in the building cost) and are very small per query, but multiplying over the billions of monthly visits results in an aggregate impact likely far greater than the training costs.

Estimates from one study for the aggregate cost of inferences for ChatGPT over a monthly period were between 1 to 23 million kWh considering a range of scenarios, with the top end corresponding to the emissions of 175,000 residents of the author's home country of Denmark.⁷ Another pair of authors arrived at 4 million kWh via a different methodology, suggesting these estimates are probably in the right ballpark.⁸

We note that in any event, the electricity usage of ChatGPT in inference likely surpasses the electricity usage of its training within weeks or even days. This aligns with claims from AWS and Nvidia that inference accounts for as much as 90% of the cost of large-scale AI workloads.^{9,10}

One comment about efficiency. Continuing our earlier analogy, instead of two or three full Boeing 767s flying round-trip from New York to San Francisco, current provision of consumer LLMs may be more like a Boeing 767 carrying one passenger at a time on that same journey. For all their power, people often use the largest LLMs for relatively trivial interactions that could be handled by a smaller model or another sort of application, such as a search engine, or for interactions that arguably need not happen at all. Indeed, some not-exactly-necessary uses of ChatGPT, such as “write a biblical verse in the style of the King James Bible explaining how to remove a peanut butter sandwich from a VCR” bear more resemblance to a single-passenger flight from New York to Cancún than from New York to San Francisco.¹¹

Excitement around GAI has produced an “arms race” between major providers like OpenAI and Google, with the goal of producing the model that can handle the widest range of possible use cases to the highest standard possible for the largest number of users. The result is overcapacity for the sake of market dominance by a single flagship model, not unlike airlines flying empty planes between pairs of airports to maintain claims on key routes in a larger network.¹² The high levels of venture capital (VC) funding currently on offer in the GAI space enable providers to tolerate

overcapacity for the sake of performance and growth.¹³ As we will discuss, business models that are much more energy- and cost-efficient are available.

We must emphasize the huge uncertainty surrounding the estimates on which this analysis is based, which stems from both lack of standard methodology and lack of transparency in the construction of LLMs. ChatGPT maker OpenAI has not publicly announced either the data used to train the model nor the number of parameters in its latest model, GPT-4. Speculation and leaks about GPT-4 put the figure at approximately 10 times the number of parameters in GPT-3, the model powering the original ChatGPT.¹⁴ Google has not released full details about the LamMDA model powering its chatbot, Bard. DeepMind, Baidu, and Anthropic have similarly declined to release full details for training their flagship LLMs.

Uncertainty remains even for open source models, since the true impact of a model involves accounting for the cost of deploying the model to an unknown and varying number of users, as well as the emissions used to produce the hardware that serves these models to end users. Still greater complexity derives from the precise mix of fossil fuels and renewable energy used where the models are trained and deployed.

Finally, we mention briefly that the water consumption of ChatGPT has been estimated at 500 milliliters for a session of 20-50 queries. Aggregating this over the billions of visitors ChatGPT has received since its launch in December 2022 amounts to billions of liters of water spent directly cooling computers and indirectly in the process of electricity generation.¹⁵

IS THAT WORSE THAN BOILING THE KETTLE?

Millions of kWh per month to run an LLM sounds like a lot. But how does that compare to the emissions generated by other activities, computational or otherwise?

A round-trip flight from New York City to San Francisco emits about 1 tonne of CO₂e per passenger.¹⁶ So the ~500 tonnes of CO₂e required to train GPT-3 equates to the emissions of approximately two or three full round-trip flights from New York City to San Francisco. Given that the

world's busiest airport, Hartsfield-Jackson Atlanta International Airport, sees an average of 1,000 departing flights per day, those 500 tonnes are relatively insignificant.

Thinking about computational activity more broadly, the information and communications technology (ICT) sector accounts for a quite significant 2%-4% of all GHG emissions globally, with a total of 1 to 2 billion tonnes CO₂e per year, on par with sectors like aviation or shipping.¹⁷ By comparison, Bitcoin mining generates 21 to 53 million tonnes of CO₂e per year, according to Massachusetts Institute of Technology (MIT) analysis.¹⁸ Bitcoin, of course, does not offer the same potential as LLMs to accelerate scientific discovery or alter the work of white-collar professionals worldwide.

Noting once again the difficulty of ascertaining such estimates, we cautiously assert that LLMs likely account for less than half a percent of emissions from the entire ICT sector and less than 0.01% of global emissions. Indeed, global annual GHG emissions in CO₂e have hovered above 50 billion tonnes annually since 2010.¹⁹ Even if inference uses 100 times as much as training, and even if there are 100 models as popular as ChatGPT, these LLMs still account for only 5 million tonnes CO₂e (100 x 100 x 500 = 5 million): 0.01% of global emissions or at most half a percent of global ICT emissions.

Table 1 shows how using ChatGPT compares to other daily activities. To compare the environmental impact of different activities, we use two metrics. The first is the electricity required, in

APPLIANCE	USAGE	ASSUMPTIONS	kWh/YEAR	KG CO ₂ e/ YEAR
Kettle	1,542 uses/year	0.11 kWh/use based on heating 1 liter of water	170	73
Electric oven	135.1 uses/year	1.56 kWh/use	211	91
Primary TV (plasma, 34-37 inches)	6.5 hours/day	263.9 w	626	269
Low-energy light bulb	4 hours/day	18 w	18	11
Using ChatGPT	Once/day	Each conversation has 20 queries; .00396 kWh/query	29	11
Google search	20 searches/day	.0003 kWh/search	2.19	<1
Email/messaging/voice/etc.	20/day	Average technological progress, average carbon intensity for Canada	Not reported	<1
Video streaming	2 hours/day	Average technological progress, average carbon intensity for Canada	Not reported	26
Flight from NY to SF	Once/year		Not reported	1,000
Bitcoin mining	219 million people with Bitcoin	Average/Bitcoin owner	Not reported	96-242
Average emissions/person globally				-6,000

Sources: Carbon Footprint, Medium, Full Fact, Luciano Rodrigues et al., *The Guardian*, Crypto News, Our World in Data

Table 1. Using ChatGPT compared to other daily activities

kWh. The downside of this metric is that many important processes (such as airplanes) do not run on electricity, so we need to consider another metric. The second yardstick is CO₂e emissions. This allows us to compare electric with nonelectric uses of energy, so it's a truer reflection of the environmental cost, but the cost of doing the same activity (e.g., powering a laptop) varies tremendously depending on the precise mix of fossil fuel and renewable energy sources where the laptop is plugged in. Note that we have only one of these metrics for some of the entries.

A SUSTAINABLE FUTURE FOR GAI

Whatever the environmental impact of LLMs, all players can reduce it by improving the location and time of training, model size, transparency, and hardware efficiency. Machine learning engineers can make improvements on any data science project by optimizing algorithms for computational efficiency or by carbon profiling,^{20,21} data leaders can empower them to train in times and places with low-carbon energy available, and LLM providers can enable carbon budgeting by being more transparent about emissions associated with their models.

Other approaches are more strategic and call for a fundamental restructuring of how GAI is currently done. LLMs first came to widespread public attention via services like ChatGPT. In this model, a well-funded tech company trains a very large LLM to handle a wide range of tasks and serves it at scale as a cloud-based chatbot to a general user base.

An alternative business model would be for AI vendors to train much smaller LLMs for specific categories of tasks.²² Specialist companies or teams would fine-tune as needed on dedicated data sets for specific use cases — that is, they would modify a model by updating a few relevant parameters. For additional efficiency, engineers can quantize models: reduce the model parameters' theoretical precision without sacrificing overall accuracy.

Such approaches can slash the computational cost of producing a fit-for-purpose LLM. In a seminal study of quantized fine-tuning,²³ some Guanaco models less than half the size of ChatGPT achieved more than 97% of the latter's performance on certain tasks with as little as 12 hours fine-tuning on a single GPU. Compare this to the purported 25,000 GPUs required to train GPT-4.

In addition to making LLMs more environmentally sustainable, this approach empowers users to harness the power of GAI for specific objectives, even those without AI engineering skills — watch for no-code fine-tuning already coming online.²⁴ With emphases on transparency and user empowerment, efforts to make LLMs environmentally sustainable also happen to align with efforts to keep it democratic.²⁵ The small-but-many-models approach may also be the best way to continue advancing the technology, with Altman himself stating that the returns from increasing the size of models will soon begin to diminish.²⁶

WHATEVER THE ENVIRONMENTAL IMPACT OF LLMS, ALL PLAYERS CAN REDUCE IT BY IMPROVING THE LOCATION AND TIME OF TRAINING, MODEL SIZE, TRANSPARENCY, AND HARDWARE EFFICIENCY

GAI shares many of the same challenges as AI and computing more generally. Data centers, storage, memory, GPUs, and so on, underlie modern computing as a whole, not just LLMs. Historically, data centers have been able to offset increases in computational demand through increased efficiency; energy required per computation in data centers decreased by 20% between 2010–2018.²⁷

Although there is debate over how long this will continue at the level of hardware, Koomey's law suggests the computational efficiency of GPUs will continue to double every couple of years over at least the medium term,²⁸ and such developments would reduce carbon emissions for anything running on GPUs, including LLMs. Blockchain, for all its sins, presents an optimistic analogy. Bitcoin specifically may deserve the blame for huge carbon emissions, but the underlying technology is evolving in more sustainable directions. The increasingly adopted proof-of-stake consensus mechanism, for example, can be orders of magnitude less costly than the Bitcoin proof-of-work consensus mechanism in terms of energy consumption.²⁹

Who will ensure the ongoing sustainability of GAI or computing more broadly? The combination of VC incentives to grow with difficulties regulating Big Tech may prove challenging for sustainability efforts. Indeed, the VC-driven push to grow at all costs runs directly counter to a desire for environmental efficiency.

Nevertheless, we believe that explicitly monitoring energy efficiency in meeting a specific user need (or the energy efficiency of the overarching models) would help keep financial incentives aligned with sustainability. Such efforts have already paid dividends in the automobile industry, where fuel efficiency of American cars approximately doubled between the introduction of Corporate Average Fuel Economy standards in 1975 and 1985,³⁰ with President Obama-era regulations driving efficiency ever higher through the 2010s.

Following the money helps us identify where to focus such pressure, regulatory or otherwise. Leading Silicon Valley venture firm Andreessen Horowitz points to hardware and infrastructure providers (Nvidia and AWS/Google Cloud/Microsoft Azure, respectively) as currently claiming the lion's share of the profits, with the widest moat as well, not the model-providers themselves.³¹

Speaking of moats, a widely circulated memo from a Google employee agrees that the models do not provide a moat for Google or for OpenAI.³² Given that venture capitalists often push for growth and more growth, rather than sustainability, and given that the tech giants are greener than most other big companies (see below), have deep pockets, and

are seemingly more susceptible to social pressure, perhaps regulators and consumers can most productively focus pressure on the likes of AWS and Nvidia.

As a final side note, following the money trail even further reveals Taiwan Semiconductor Manufacturing Company chips underpinning all these other giants,³³ which highlights the geopolitical risk to a truly sustainable existence for all kinds of computing, at least in the West.

TECH GIANTS GREENER THAN MOST OTHER BIG COMPANIES

LLM operators can reduce their carbon footprint by using renewable energy. For example, they could use corporate power purchase agreements (CPPAs) to procure green electricity from wind or solar farms. Google reached 100% renewable energy in 2017 with precisely this approach.³⁴ The company maintained that level by signing additional CPPAs for a cumulative total of 7 gigawatts (GW) through 2021 (equivalent to 44% of the total capacity installed in 2020 in the US³⁵) to cover the rapid expansion of computing conducted in the company's data centers in the recent years.

Other players have 100% renewable electricity targets; Microsoft is aiming for 2025.³⁶ In other words, tech players' electricity consumption is less carbon-intensive than the national average. This results in overly high estimates of the carbon footprint of GAI in countries with a highly carbon-intensive grid like the US, since such estimates rely on an average carbon usage per kWh rather than what a tech company actually uses.

Big Tech pioneered CPPA market development, and this leadership translates into other benefits supporting broader impact minimization.³⁷ First, this initiative added to the increasing scrutiny about GAI's energy consumption and put pressure on other players to follow suit.³⁸ As a result, most GAI will likely run on green electricity in the future. Second, this push from tech players had a positive effect on the entire clean energy industry. These companies' efforts have pushed growth with substantial investments that kick-started the market and bypassed less agile utilities.

In 2020 alone, Google, Amazon, Facebook, Apple, and Microsoft procured 7.2 GW of renewable capacity, which is almost 30% of all CPPAs, or around 3.5% of all global renewable capacity additions.³⁹ This contributed to making CPPAs an accessible tool for companies to source renewable electricity and offered an alternative to subsidies for developers.

One could argue that these new renewable energy projects could have been developed for other consumers (i.e., they could replace current electricity supply from fossil sources instead of supplying new demand). There is some truth to this, but on balance, we believe the contribution of the tech giants to be positive here, and there remains vast potential for other players and industries to follow suit. For example, this objection does not apply to onsite solar photovoltaic projects that would otherwise not happen.

In addition, in some data center hot spots, such as the Nordics, wind or solar farms would likely not get off the ground without this new demand, given the limited local need for additional renewable energy. Finally, the electricity generation need

not sit near the data center, or even in the same country, in the case of integrated electricity grids like in Europe.

Nevertheless, energy usage by LLMs is the latest stage in an ever-increasing thirst for energy across the tech industry, especially among hyperscalers. Globally, data centers may represent a relatively small share of electricity demand, but locally they can play an outsized role.

Grid capacity is limited, and the addition of renewables shifts power generation to new areas and results in an increasingly decentralized system. This complexity, combined with the lengthy process to develop new transmission infrastructure, hobbles energy transition. Hyperscalers will compete for the remaining access to electricity to the detriment of others (which some may consider more crucial to the economy and/or security). In a prime example of this conflict, the Norwegian group Nammo recently blamed a new TikTok data center for preventing the expansion of the group's ammunition plant supplying Ukraine with artillery rounds. TikTok's data center had already taken the grid's spare capacity near the plant.⁴⁰



The remaining elephant in the room is this: when is the electricity generated? Typically, a company signs a CPPA for a total annual volume of renewable electricity, regardless of when the electricity is generated. So on an annual net basis, it considers its energy consumption green. However, on an hourly basis, there is often a difference between the consumption of the data center and what the wind and solar assets have generated. This disconnect between the timing of the electricity produced and the electricity consumed calls into question whether corporations can honestly call their consumption green.

Thus, the remaining challenge to minimize the environmental impact of GAI's energy consumption is to make that link in real time, which explains why companies now increasingly focus on procuring electricity in line with their consumption profile or adapt their consumption to the electricity generated. Google aims to establish this link with its "24/7 Carbon-Free Energy by 2030" program.⁴¹

CLOSING THOUGHTS

Whether or not we have truly entered a phase of global boiling, we can confidently predict an increasing need to reduce carbon emissions. And regardless of how many industries truly reorient around LLMs, we can count on a GAI playing a growing role in our lives. Exactly how these two threads intertwine remains to be seen.

What is certain is that, at present, the emissions from LLMs are relatively insignificant compared to both their popularity and to other everyday activities. At the same time, as compute-intensive LLMs permeate our lives, the extent to which the technology may come to compromise sustainability merits continued attention.

This is particularly true if dreams of democratized AI become a reality; widespread empowerment of diverse grassroots users to host and fine-tune smaller LLMs could complicate efforts to track the technology's environmental impacts and would create more players with the duty to use it responsibly.

Fortunately, opportunities abound for all participants in the LLM space to strive for responsible AI, from optimizing model-training efficiency to sourcing cleaner energy and beyond. We believe LLM-driven advances in R&D have the potential to turbocharge society's journey toward net zero. Indeed, a reorienting of green tech around LLM advances may be just the impetus we need.

REFERENCES

- ¹ Niranjana, Ajit. "['Era of Global Boiling Has Arrived,' Says UN Chief as July Set to be Hottest Month on Record.](#)" *The Guardian*, 27 July 2023.
- ² Gates, Bill. "['The Age of AI Has Begun.'](#)" *GatesNotes*, 21 March 2023.
- ³ Knight, Will. "['OpenAI's CEO Says the Age of Giant AI Models Is Already Over.'](#)" *Wired*, 17 April 2023.
- ⁴ Ludvigsen, Kasper Groes Albin. "['The Carbon Footprint of GPT-4.'](#)" *Medium/Towards Data Science*, 18 July 2023.
- ⁵ Touvron, Hugo, et al. "['Llama 2: Open Foundation and Fine-Tuned Chat Models.'](#)" *Cornell University*, 19 July 2023.
- ⁶ Luccioni, Alexandra Sasha, Sylvain Viguier, and Anne-Laure Ligozat. "['Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model.'](#)" *Cornell University*, 3 November 2022.
- ⁷ Ludvigsen, Kasper Groes Albin. "['ChatGPT's Electricity Consumption.'](#)" *Medium/Towards Data Science*, 1 March 2023.
- ⁸ Luccioni et al. ([see 6](#)).
- ⁹ Barr, Jeff. "['Amazon EC2 Update — Inf1 Instances with AWS Inferentia Chips for High Performance Cost-Effective Inferencing.'](#)" *AWS News Blog*, 3 December 2019.
- ¹⁰ Leopold, George. "['AWS to Offer Nvidia's T4 GPUs for AI Inferencing.'](#)" *HPCwire*, 19 March 2019.
- ¹¹ Tewari, Niharika. "['ChatGPT: 10 Most Hilarious and Weird Responses That ChatGPT Has Produced!'](#)" *Sociobits.org*, 14 December 2022.

- ¹² Stokel-Walker, Chris. "[Thousands of Planes Are Flying Empty and No One Can Stop Them.](#)" *Wired*, 2 February 2022.
- ¹³ Pennington, Clarke. "[Generative AI: The New Frontier for VC Investment.](#)" *Forbes*, 17 January 2023.
- ¹⁴ Peleg, Yam. "[GPT-4's Details Are Leaked. It Is Over. Everything Is Here.](#)" Archive.today, accessed August 2023.
- ¹⁵ Li, Pengfei, et al. "[Making AI Less 'Thirsty': Uncovering and Addressing the Secret Water Footprint of AI Models.](#)" Cornell University, 6 April 2023.
- ¹⁶ Kommenda, Niko. "[How Your Flight Emits as Much CO2 as Many People Do in a Year.](#)" *The Guardian*, 19 July 2019.
- ¹⁷ Lavi, Hessam. "[Measuring Greenhouse Gas Emission in Data Centres: The Environmental Impact of Cloud Computing.](#)" *Climatiq*, 21 April 2022.
- ¹⁸ Stoll, Christian, Lena Klaassen, and Ulrich Gattersdorfer. "[The Carbon Footprint of Bitcoin.](#)" Massachusetts Institute of Technology (MIT) Center for Energy and Environmental Policy Research (CEEPR), December 2018.
- ¹⁹ Ritchie, Hannah, and Max Roser. "[Greenhouse Gas Emissions.](#)" Our World in Data, 2020.
- ²⁰ Ludvigsen, Kasper Groes Albin. "[How to Estimate and Reduce the Carbon Footprint of Machine Learning Models.](#)" Medium/Towards Data Science, 1 December 2022.
- ²¹ "[CodeCarbon.](#)" GitHub, accessed August 2023.
- ²² Luccioni, Sasha. "[The Mounting Human and Environmental Costs of Generative AI.](#)" *Ars Technica*, 12 April 2023.
- ²³ Dettmers, Tim, et al. "[QLoRA: Efficient Finetuning of Quantized LLMs.](#)" Cornell University, 23 May 2023.
- ²⁴ Datta, Souvik. "[Introducing No-Code LLM FineTuning with Monster API.](#)" Monster API, 4 July 2023.
- ²⁵ Dempsey, Paul. "[Access for All: The Democratisation of AI.](#)" The Institution of Engineering and Technology, 10 November 2021.
- ²⁶ Knight ([see 3](#)).
- ²⁷ Masanet, Eric, et al. "[Recalibrating Global Data Center Energy-Use Estimates.](#)" *Science*, Vol. 367, No. 6481, February 2020.
- ²⁸ Scott, Art, and Ted G. Lewis. "[Sustainable Computing.](#)" *Ubiquity*, Vol. 2021, February 2021.
- ²⁹ Platt, Moritz, et al. "[The Energy Footprint of Blockchain Consensus Mechanisms Beyond Proof-of-Work.](#)" *Proceedings of the 21st International Conference on Software Quality, Reliability, and Security Companion (QRS-C)*. IEEE, 2022.
- ³⁰ "[Driving to 54.5 MPG: The History of Fuel Economy.](#)" Pew, 20 April 2011.
- ³¹ Bornstein, Matt, Guido Appenzeller, and Martin Casado. "[Who Owns the Generative AI Platform?](#)" Andreessen Horowitz, accessed August 2023.
- ³² Patel, Dylan, and Afzal Ahmad. "[Google 'We Have No Moat, and Neither Does OpenAI.'](#)" *SemiAnalysis*, 4 May 2023.
- ³³ Bornstein et al. ([see 31](#)).
- ³⁴ "[Operating on 24/7 Carbon-Free Energy by 2030.](#)" Google Sustainability, accessed August 2023.
- ³⁵ "[US Renewable Energy Factsheet.](#)" University of Michigan Center for Sustainable Systems, accessed August 2023.
- ³⁶ Hook, Leslie, and Dave Lee. "[How Tech Went Big on Green Energy.](#)" *Financial Times*, 10 February 2021.
- ³⁷ Varro, Laszlo, and George Kamiya. "[5 Ways Big Tech Could Have Big Impacts on Clean Energy Transitions.](#)" IEA, 25 March 2021.
- ³⁸ Hook et al. ([see 36](#)).
- ³⁹ Varro et al. ([see 37](#)).
- ⁴⁰ Dodgson, Lindsay. "[Weapons Firm Says It Can't Meet Soaring Demand for Artillery Shells Because a TikTok Data Center Is Eating All the Electricity.](#)" *Insider*, 27 March 2023.
- ⁴¹ Google Sustainability ([see 34](#)).

About the authors

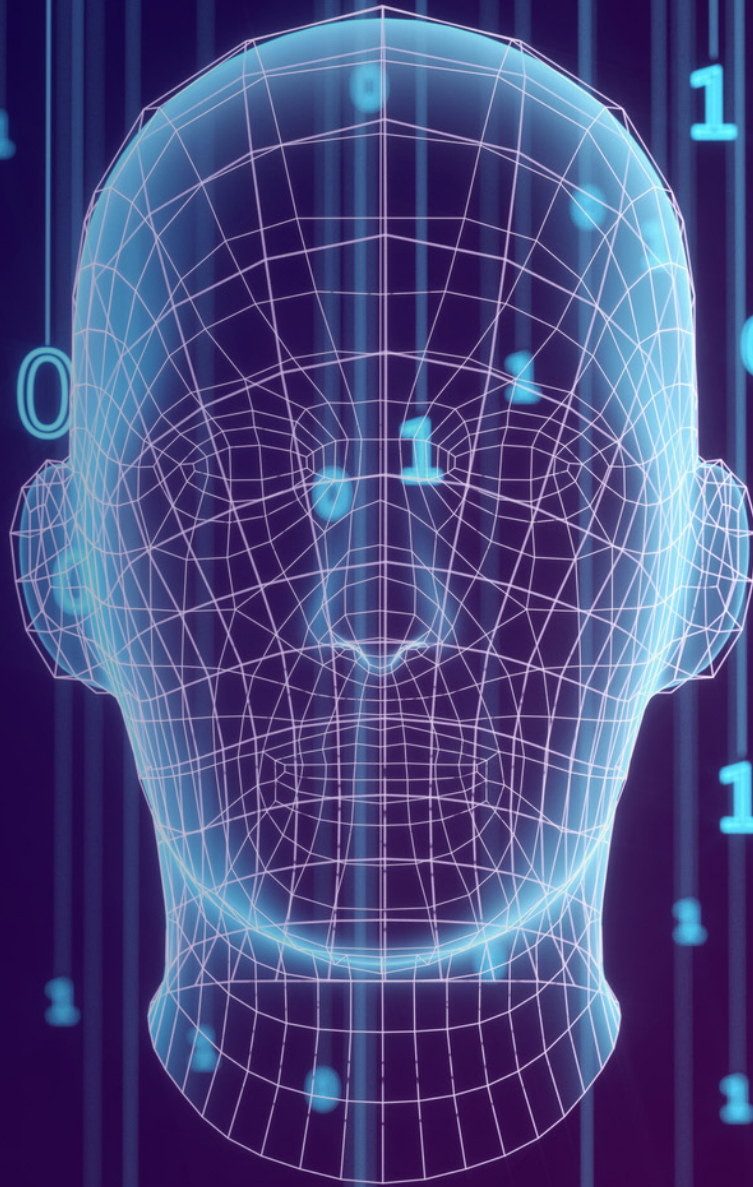
Greg Smith is Managing Partner of Arthur D. Little (ADL). He founded and co-leads ADL's Digital Problem Solving practice and is a member of ADL's Executive Committee, where he has responsibility for ADL's global innovation strategy. His work focuses on business strategy in the context of digital transformation as well as the application of disruptive information technologies in solving intractable business problems in major enterprises. Recently, Mr. Smith has been focusing on digital operating models for established businesses, including the changes required in technology, culture, IT functions, business technology interactions, organizational design, and governance, and how these can combine to enhance customer and service experience. He holds a bachelor of science degree in biological sciences from the University of Leicester, UK, and finds that after 30 years of dormancy within his professional life, the underlying concepts of biology are becoming increasingly valuable at unlocking business problems and articulating solutions — especially where reductive, engineering-based approaches need to be replaced with whole-system, evolutionary thinking. He can be reached at experts@cutter.com.

Michael Bateman is a Data Scientist with Arthur D. Little's (ADL's) Digital Problem Solving practice. Dr Bateman's work ranges from fundamental research to providing actionable insights for decision makers, and in verticals spanning drug discovery, healthcare, defense, pharma, transport, and public policy. Prior to joining ADL, he held positions at UCLA and the University of Cambridge (UK) as well as in government and the biotech industry. Dr. Bateman earned a bachelor's degree in mathematics from the University of Kansas and a PhD in pure mathematics from Indiana University Bloomington. He can be reached at experts@cutter.com.

Remy Gillet is a Senior Consultant with Arthur D. Little's (ADL's) UK Energy & Utility practice and a member of ADL's AMP open consulting network. He's passionate about decarbonization through the deployment of renewable energy and other clean technologies. Mr. Gillet enjoys supporting renewable energy companies and investors to navigate and accelerate the energy transition. Before joining ADL, he worked as a financial analyst with two independent power producers and contributed to the development of wind and solar PV projects across North America. He can be reached at experts@cutter.com.

Eystein Thanisch is a Software Developer with Arthur D. Little's (ADL's) UK Digital Problem Solving practice. He enjoys ambitious projects that involve connecting heterogeneous data sets to yield insights into complex, real-world problems and believes in uniting depth of knowledge with technical excellence to build things of real value. Dr. Thanisch is also interested in techniques from natural language processing and beyond for extracting structured data from texts. Prior to joining ADL, he worked on *Faclair na Gàidhlig*, the historical dictionary of Scottish Gaelic, on a team tasked with building a tagged corpus of transcriptions from pre-modern manuscripts. He also was involved in IrishGen, a project on the use of knowledge graphs to represent medieval genealogical texts. Dr. Thanisch also worked as a freelance editor and analyst for a number of IGOs and academics. He earned a master of science degree in computer science from Birkbeck, University of London, and a PhD in Celtic studies from The University of Edinburgh, Scotland. He can be reached at experts@cutter.com.

WHAT'S REALLY AHEAD FOR GENERATIVE AI?



Author

Paul Clermont

The term “artificial intelligence” (AI) was coined in 1956, when a group of mathematicians and computer scientists (surely inspired by Alan Turing, who died in 1952) conceived of a computer that could not only solve mathematical and combinatorial problems but could learn to become progressively better at doing so.¹ AI stayed well out of the public eye, confined to academic and corporate laboratories until 1996, when Deep Blue, an IBM supercomputer, beat the world’s best chess player. Other notable successes followed, including Watson (another IBM supercomputer) winning *Jeopardy!* in 2008 and DeepMind (software developed by a British company subsequently bought by Google) beating the world’s best *Go* player in 2016. Mass media took note. Today, AI news and opinions are daily fare, driven in part by the emergence of generative AI (GAI), which can generate original text and pictures about almost anything with minimal human instruction.

This is far more impressive than winning games with clear objectives and strict rules. New publicly accessible software has inspired both awe and fear as chatbots and illustrators (e.g., DALL-E) appeared and began improving, seemingly by the day. The speed of innovation has stoked fears that GAI will lead to terrible consequences. Governments in the US, UK, EU, and China are scrambling to develop AI regulations in an effort to minimize potential harm.

GAI promises (threatens?) to be a conceptual breakthrough on the level of automobiles, TV, and personal computers. Although none of these were very useful or universally embraced at first, they did not face active hostility as they worked their way into our daily lives. GAI faces a more skeptical public:

- The long honeymoon that Big Tech enjoyed is over.
- Less endearing aspects of the Internet have come to the fore (more on this later).
- Some tech gurus and pundits have raised the specter of super-intelligent, amoral machines taking over and destroying humankind.
- Tech industry moguls are calling for government regulation, a first for a group in which libertarians are well represented.

Anxiety, expressed in words like “trustworthiness” and “harms,” has become prominent in the public discourse and, more importantly, in government documents, such as proposed laws and requests for public comment.

THE LONG HONEYMOON THAT BIG TECH ENJOYED IS OVER

CURRENT CAPABILITIES & LIMITATIONS

Before we can talk meaningfully about GAI, we should think about what the word “intelligence” conveys. When we describe people as intelligent, we may be referring to intellectual capabilities; for example:

- Ability to command a great deal of information, both general and specific to a person’s career and interests
- Ability to see nonobvious patterns and connections

- Ability to identify and methodically evaluate a rich set of options before making a decision
- Ability to see beyond conventional wisdom when problem solving
- Ability to express oneself in clear, coherent, grammatically correct language



All of these have been successfully replicated on computers. The last item can be seen in GAI like GPT-4 and DALL-E, which can create high-quality, readable, and understandable text, pictures, and graphics for a variety of purposes:

- It can help with research by finding and summarizing relevant published content.
- It can find and summarize relevant legal cases and statutes.
- It can do not-very-creative writing, including boilerplate and form letters. It can draft documents for people to finish (e.g., proposals and business letters), a boon for those spooked by a blank screen.
- It can create pictures or modify existing pictures from a text description. Uses can be commercial or personal.
- It can help software developers find and incorporate segments of proven open source code.
- It can simplify and accelerate extracting and summarizing relevant information from large numbers of responses to open-ended questions.²

As we'll see, fact-checking is still needed when the stakes are high and/or when there may be legal considerations or possible copyright infringement. Nevertheless, it's enough to scare folks. If a computer can do all that and get better at it almost by the day ...

HOW SCARED SHOULD WE BE?

On the spectrum between Alfred E. Neuman (the "What, me worry?" kid of *Mad Magazine* fame) and Chicken Little ("The sky is falling!"), I suggest leaning toward Alfred. Lots of technologists and other pundits disagree, so I'll try to make the case as clear as possible.

However impressive the GAI we have seen thus far is, it falls well short of a reliable tool. It is still in the stunt (or, more politely, the proof-of-concept) stage of development. For example, partners in a law firm were recently fined US \$5,000 for using ChatGPT to write a legal brief. The output looked great, citing decisions in relevant cases that supported the brief's arguments. The problem was that no such cases or decisions existed. The tool made them up, and the partners did not check the work as they would have done for a paralegal or junior associate.³

This is not an isolated case. In April, the *Washington Post* reported on an incident in which an AI chatbot generated a list of legal scholars who had sexually harassed someone and included a law professor who had never been accused of sexual harassment.⁴ The AI term for this is "hallucinations," which includes coming up with inappropriate non-facts or highly plausible fakery. Large language model (LLM) mavens haven't yet figured out why this happens or how to prevent it but are presumably working feverishly on the problem. A text-generating tool that requires thorough fact-checking before using its products may be worse than no tool at all.

On a positive note, GAIs write rather well. An LLM uses vast amounts of written material to learn what words tend to follow other words, similar to what we see when composing text on a smartphone. The quantity of material it uses for this training is the paramount consideration, assuming that material is grammatical and uses words correctly. Judging by current LLMs, this part of the training works well.

However, training LLMs in the subjects it may be requested to write about requires concentration on the quality (not quantity) of the training data. LLMs cannot currently recognize obsolete information or detect subtle biases in training materials, and if they're there, they'll make their way into its written products: GIGO (garbage in, garbage out), as always. (The people tasked with cleaning up training data aren't infallible either.)

In the meantime, GAI images are getting better. When DALL-E first came out, I asked for pictures of a smiling woman with a martini. It returned crude, almost grotesque images as though the intelligence itself had been dulled by more than a few drinks. I tried again last week and got high-quality, believable pictures. Since images of good-looking young women are plentiful, I asked for the smiling woman to be in her 80s — again, high-quality and realistic images were returned.

Generally speaking, current design concepts appear robust enough to permit more sophistication, assuming the engineers who tackle problems like hallucinations succeed. The quality of subject-area training data can, and will, be improved as AIs focus on more limited knowledge pools. AI learns autonomously and can increasingly participate in the campaign to eradicate obsolete and biased training data, especially when augmented by human reinforcement. Caveats and favorable assumptions notwithstanding, one can only say that generative AI is an impressive achievement on a par with the first modern digital computer and the Internet.

But how close is it to artificial general intelligence?

Although GAI shows some impressive signs of human intelligence, it is still an idiot savant.⁵ It far exceeds human capability at tasks computers are good at (e.g., instant recall from a gargantuan photographic memory or rapid computation or pattern recognition), and it will tirelessly and monomaniacally pursue an objective without being distracted by anything.

But there are hallmarks of human intelligence that are less easily defined (but recognizable when you see them) and vital to making the world work. Examples include judgement (pragmatic, ethical, moral); intuition; reading between the lines; sensing implicit questions, dealing with nuance;

creating truly original artifacts (stories, images, music) out of one's imagination; seeing subtle connections and analogies and distinguishing them from coincidence; sensitivity to context; plain old common sense; and the panoply of people skills, not least persuasiveness. People given an objective can sense implicit constraints that matter when their effort might be carried to the point of conflicting with something more important.

To be truly artificially generally intelligent, the technology must reflect these additional hallmarks to some degree. In this endeavor, we are essentially nowhere, which should surprise no one. The history of AI is littered with erroneous assumptions that because a complex intellectual feat has been achieved (champion-level checkers in 1962 plus the milestones cited earlier), less intellectually demanding tasks will be relatively easier.

In fact, the opposite has proven true. The more "primitive" the mental activity humans have been engaging in for thousands of years (before chess), the more difficult it is to capture it in software.⁶ It may be impossible in some cases. Ironically, some of the most intelligent people in terms of "softer" hallmarks are not the most brilliant in "harder" ones, and vice versa.⁷ Games like chess reward a computer's strengths; real life rewards uniquely human strengths, supplemented with appropriate computer strengths.

We should also bear in mind that the concepts and techniques underlying GAI were developed years ago and could have been brought to market much sooner but for the lack of sufficiently fast processors and sufficiently massive storage. The fast progress of the past few months does not represent a steady-state development rate.

This is not to say that we will never build machines with artificial general intelligence. We may do so eventually (it's a mistake to underestimate the creative power of amazingly clever people setting their minds to an intellectually challenging task), but we ought not to hold our breath.

Some very prestigious people, including tech executives, are clamoring for AI regulation out of concern for artificial general intelligence (or maybe even super intelligence) wreaking havoc in the not-too-distant future. A more cynical interpretation is that this is a deliberate diversion from having

governments worry about the plentiful harms that today's limited AI can (and probably will) do in the short term.⁸ Warning: This argument applies only to generative AI (see sidebar "Potential Dangers of Non-Generative AI").

ETHICAL & SOCIAL IMPLICATIONS

In the short term, it's widely understood that AI can be misused to do real harm to individuals and society. That was not a consideration for computers, per se. It wasn't for the Internet either, but in retrospect should have been as we wrestle with loss of privacy; advertising revenue-driven surveillance capitalism;⁹ an algorithm-enhanced attention economy sending people down informational rabbit holes; scammers; screen addiction (particularly among children and teens); trolling; cyberbullying to the point of driving some to suicide; and the proliferation of fake news, disinformation, conspiracy theories, and outright lies

POTENTIAL DANGERS OF NON-GENERATIVE AI

AI built into highly networked systems that control physical facilities like the power grid present far more opportunity for near-apocalyptic, intermediate-term mischief. They may seem easier to build since they don't require artificial general intelligence, but doing it right is far from easy. It requires building in the software equivalents of guardrails and containment vessels needed to respond to all that can go wrong. Unless multilevel constraints are built in, an AI given an objective will pursue that objective relentlessly in both reasonable and crazy ways, and any attempt to work around them must be recognized and thwarted. *Common sense is not an AI feature.* Obviously, security against intrusion and hacking are of paramount importance. There's also a case to be made for not pursuing the last iota of efficiency in choosing what to automate and whether to use AI. Overoptimization is a trap: it assumes everything will work correctly all the time — those whose supply chains were massively disrupted by the pandemic or the Fukushima nuclear meltdown have presumably learned this lesson.

that can polarize societies. Unfortunately, AI puts what we don't like about today's Internet on steroids — mass production with minimal human labor. Implications to consider include:

- Chatbots hosted by large tech companies are likely to reflect honest efforts to make them as good as possible because the stakes are high, but that's not necessarily true for smaller players.¹⁰
- AI facilitates undesirable (and sometimes already illegal) hyperrealistic deepfakes (pictures, videos, or audios) that purport to show someone doing or saying something they never did or said in a place they may never have been, perhaps in the company of people they never met.¹¹ The potential for this to sway elections and destroy careers, reputations, and marriages is obvious.
- Face recognition without laws to restrain its use could lead to a level of privacy loss that would make the world of George Orwell's *1984* seem benign.

Bias-free is a goal that must be pursued assiduously, but it cannot be an a priori requirement. Essentially, all potential training data other than hard science reflects some bias (conscious or unconscious). No effort to "clean" it can be 100% effective, but it must still be remediated when detected. Fortunately, a side industry of think tanks, both independent and within major players, has emerged to address AI ethics.¹² We need them to be not only smart, but skillful in ensuring they're heard and heeded.

Additionally, a host of new legal issues needs attention:

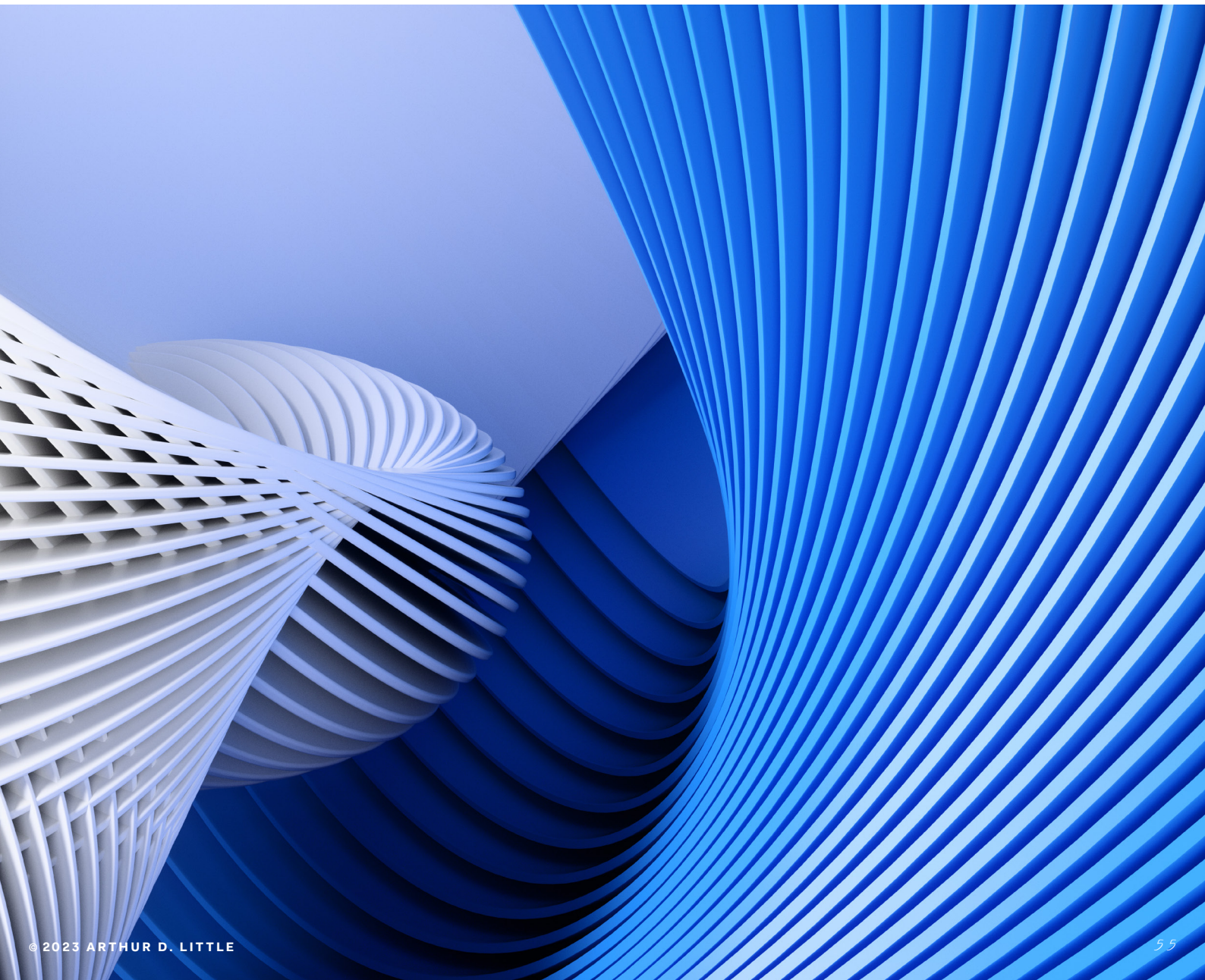
- Compensating the owners of the data used for training. (Will LLMs like the one behind ChatGPT that hoovered up training data from everywhere without regard to copyright have to be replaced?)
- Footnoting and citing reference sources.
- Limits on free speech when producing and disseminating assertions known to be false (e.g., "the Earth is flat") or overwhelmingly discredited by numerous reputable bodies (e.g., "Trump won in 2020").
- Classifying AI applications by level of risk to guide the scope and depth of regulation as the EU is doing with four categories: forbidden, strictly regulated, loosely regulated, and unregulated, based on the potential to do harm.

- Identification of material substantively generated by AI (not just improved in grammar and style) as an AI product.
- Treatment of libelous material when generated by AI rather than humans.
- Liability of AI app users versus providers of the LLMs they rely on.
- Licensing and auditing AI applications.
- Penalties for malicious deepfakes.

There is also the issue of jobs. Every wave of industrialization and automation has spawned predictions of massive job loss. So far, the lost jobs have been replaced by new jobs.¹³ The lump-of-labor fallacy has proven durable, though — a bit like the boy who cried wolf.¹⁴ Will generative AI be a real wolf at last? Probably not, but there are no guarantees.

WHAT CAN WE EXPECT IN THE NEAR TERM?

First, generative AI will gradually become less controversial, although some will continue to decry it, emphasizing its problems and potential for harm. Some early critics will find ways to use it to good effect. For example, many teachers who initially saw only the problem of high-tech plagiarism have now embraced AI as a way to teach critical reading and thinking, including lessons on wording requests to get the results you seek (the term “prompt engineering” refers to this). AI will get better, and we’ll see cautious (very cautious if the hallucination problem isn’t substantially mitigated) uptake in areas where its value is clear.



Second, barring major fiascos, building in AI will become the latest corporate fad, with almost every large company claiming to use it, no matter how trivial the application. (Remember reengineering?)

Third, governments will play a role in the evolution of AI usage that they did not play in the advent of the computer or the Internet. The EU and the UK government are devising and enacting laws. China is making rules as well, untrammelled by messy democratic processes. The US government (both Congress and the Executive Branch) is gathering testimony and commentary from industry executives and thought leaders, but the path to laws and regulations will likely be circuitous. Of necessity, US companies doing business overseas will need to conform to local regulations. GAI is sufficiently new and different, with effects on everyone, that attempting to regulate it by tweaking existing laws and regulations would be futile. Instead, legislators and bureaucrats need to approach the job from the ground up, based on principles that fully account for GAI's uniqueness.¹⁵

Fourth, AI's potential value in waging war will provide an impetus for rapid development wholly apart from business and personal use.

Fifth, any further talk of moratoriums on AI development will be just talk, no matter how eminent the people who call for it. In a field as exciting as AI in an industry as fiercely competitive as high tech, the idea that super-motivated inventors and engineers will simply take an extended vacation is ludicrous.

CONCLUSION

For better or worse, GAI is here. It may be a boon, reducing drudgery, leveraging talent, increasing productivity, and simplifying life at home and at work. It may also be a bane, as I've suggested.

The assumption that a free market will somehow grow the good AI and suppress the bad is hopelessly naive, as our Internet experience has shown. Potential profits are too high for that. We must hope our political leaders have the fortitude to

push entrepreneurs and executives to do the right thing, which will require making doing the wrong thing unprofitable.¹⁶ Turning good ideas into enforceable laws and regulations that capture their spirit (i.e., not riddled with loopholes) will be the challenge of the decade. Still, we have to hope.

REFERENCES

- ¹ One of them was Marvin Minsky (1927–2016), whose course in AI I took many years ago. It was irrelevant to my career until recently.
- ² I got this from my son, a political pollster. Before generative AI, pollsters avoided open-ended questions despite the deeper insights they offer because of the labor required to analyze the answers.
- ³ Weiser, Benjamin. "[ChatGPT Lawyers Are Ordered to Consider Seeking Forgiveness.](#)" *The New York Times*, 22 June 2023.
- ⁴ Verma, Pranshu, and Will Oremus. "[ChatGPT Invented a Sexual Harassment Scandal and Named a Real Law Prof as the Accused.](#)" *The Washington Post*, 5 April 2023.
- ⁵ This is an old-fashioned term, now surely politically incorrect, used to describe some unusual people who would today be considered on the autism spectrum. They seemed to be of low intelligence but exhibited superhuman capabilities at tasks like rapidly multiplying three-digit numbers in their heads or quickly identifying the day of the week, given the date on which some long-past event occurred. Sadly, they were exploited as curiosities.
- ⁶ It has proven extremely difficult to teach a robot to run across a randomly cluttered room without bumping into things, something most one-year-olds teach themselves to do.
- ⁷ US President Franklin D. Roosevelt, generally considered among the top three presidents (with Washington and Lincoln), was described by US Supreme Court Justice Oliver Wendell Holmes as having a second-class intellect but a first-class temperament.

- ⁸ This possibility has been mentioned in several publications, including: Heaven, Will Douglas. [“How Existential Risk Became the Biggest Meme in AI.”](#) *MIT Technology Review*, 19 June 2023.
- ⁹ Zuboff, Shoshana. [《The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power》](#). PublicAffairs, 2019.
- ¹⁰ Thompson, Stuart. [“Uncensored Chatbots Provoke a Fracas Over Free Speech.”](#) *The New York Times*, 3 July 2023.
- ¹¹ This is not a new idea; AI just makes it easier. Prior-year photographs of Stalin’s Politburo in the 1930s were seamlessly doctored not to show former members whom Stalin had subsequently defenestrated, perhaps literally.
- ¹² Knight, Will, Khari Johnson, and Morgan Meaker. [“Meet the Humans Trying to Keep Us Safe from AI.”](#) *Wired*, 27 June 2023.
- ¹³ A 1964 study commissioned by the US government predicted that the automation just getting started would in 30 years result in massive unemployment. Almost 60 years on, the US is enjoying a record-low unemployment rate.
- ¹⁴ The assumption that if some technology made people doing X twice as productive, employers would simply sack half of them rather than use them to increase output or perform some new tasks.
- ¹⁵ Symons, Charles, Ben Porter, and Paul Clermont. [“Twelve Principles for AI Regulation.”](#) Prometheus Endeavor, 5 July 2023.
- ¹⁶ The EU seems to be on that track, levying fines stiff enough that even a multi-billion-dollar corporation feels some pain.

About the author

Paul Clermont is a Cutter Expert. He has been a consultant in IT strategy, governance, and management for 40 years and is a founding member of Prometheus Endeavor, an informal group of veteran consultants in that field. His clients have been primarily in the financial and manufacturing industries, as well as the US government. Mr. Clermont takes a clear, practical view of how information technology can transform organizations and what it takes to direct both business people and technicians toward that end. His major practice areas include directing, managing, and organizing information technology; reengineering business processes to take full advantage of technology; and developing economic models and business plans.

Mr. Clermont is known for successfully communicating IT issues to general managers in a comprehensible, jargon-free way that frames decisions and describes

their consequences in business terms. In his consulting engagements, he follows a pragmatic approach to the specific situation and players at hand and is not wedded to particular models, methodologies, or textbook solutions.

Before going into individual practice, Mr. Clermont was a Principal with Nolan, Norton & Co., a boutique consultancy that became part of KPMG. Before joining Nolan, Norton & Co., he directed IT strategy at a major Boston bank and launched its IT executive steering committee. Mr. Clermont has spoken and written about the challenges of getting significant and predictable value from IT investments and has taught executive MBA courses on the topic. His undergraduate and graduate education at MIT’s Sloan School of Management was heavily oriented toward operations research. He can be reached at experts@cutter.com.

AMPLIFY

Anticipate, Innovate, Transform

Cutter Consortium, an Arthur D. Little community, is dedicated to helping organizations leverage emerging technologies and the latest business management thinking to achieve competitive advantage and mission success through our global research network. Cutter helps clients address the spectrum of challenges disruption brings, from implementing new business models to creating a culture of innovation, and helps organizations adopt cutting-edge leadership practices, respond to the social and commercial requirements for sustainability, and create the sought-after workplaces that a new order demands.

Since 1986, Cutter has pushed the thinking in the field it addresses by fostering debate and collaboration among its global community of thought leaders. Coupled with its famously objective “no ties to vendors” policy, Cutter’s *Access to the Experts* approach delivers cutting-edge, objective information and innovative solutions to its community worldwide.

Amplify is published monthly by Cutter Consortium, an Arthur D. Little community, 37 Broadway, Arlington, MA 02474-5552, USA

Founding Editor: Ed Yourdon
Publisher: Karen Fine Coburn
Group Publisher: Christine Generali
Production Manager: Linda Dias
Copyeditor: Tara K. Meads

© 2023 Arthur D. Little. All rights reserved. For further information, please visit www.adlittle.com.

CUTTER

AN ARTHUR D. LITTLE
COMMUNITY

For more content,
visit www.cutter.com