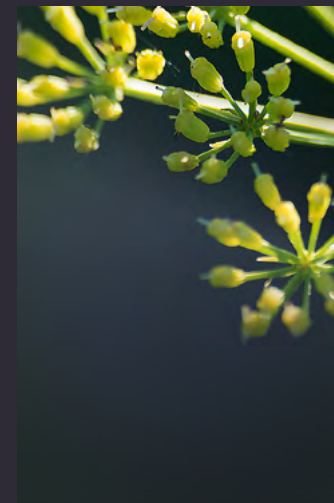
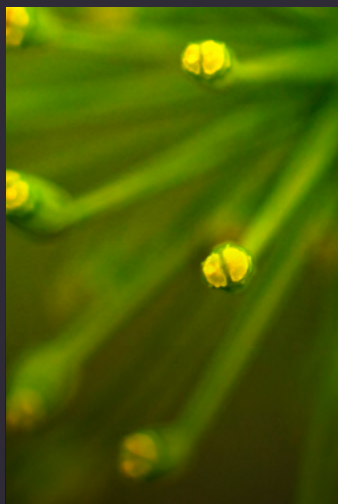




Developing *sustainable Gen AI*

Table of contents





Executive summary

Generative AI has a significant environmental impact:

Generative AI (Gen AI) relies on processing huge volumes of data, which necessitates enormous computational power, rendering it an energy-intensive technology. The production of graphics processing units (GPUs), integral to the functioning of Gen AI, requires rare earth metals, the mining of which contributes to greenhouse gas (GHG) emissions. Additionally, the frequent hardware upgrades that Gen AI requires put a great deal of stress on natural resources, as well as further polluting the environment. Recent estimates suggest that, Gen AI could create between 1.2 to 5.0 million metric tons of e-waste by 2030, which is around 1,000 times more e-waste than was produced in 2023.¹

Estimates suggest that training a GPT-3 model (which includes 175 billion parameters) consumes an amount of electricity equivalent to the annual consumption of 130 US homes.² Moving to the next model size up, GPT-4 (with 1.76 trillion parameters), power consumption of training is estimated to be equivalent to yearly power consumption of 5,000 US homes.³ After training, which is a one-time event in the model's lifecycle, the inferencing phase, where models

perform real-time functions, requires an equal or greater amount of energy. Data centers also consume a huge amount of energy and water for cooling purposes. Running an inference of 20–50 queries on an LLM uses about 500 ml of water each time.

Gen AI has contributed to increased GHG emissions:

In our research, we surveyed executives from 2,000 organizations that have Gen AI initiatives underway.

- 48% of executives believe that their use of Gen AI has driven a rise in GHG emissions.
- Organizations that currently measure the environmental impact of their use of Gen AI project the share of Gen AI-driven emissions as a proportion of total organizational carbon emissions to rise, on average, from 2.6% to 4.8% over the next two years.
- 42% of executives have had to relook at their climate goals due to Gen AI's growing footprint.

Executive summary

However, many organizations continue to ignore Gen AI's growing carbon footprint:

Just 12% of the executives in our research confirm that their organizations measure Gen AI footprint. Furthermore, just 20% rank “environmental footprint of Gen AI” among the top five factors when selecting or building Gen AI models, with performance, scalability, and cost dominating their consideration process. Only 27% of executives say they compare energy consumption levels of Gen AI models. Organizations are currently taking only a partial view of costs, effectively ignoring the energy costs of model deployment and inferencing. As Gen AI models become more complex and pervasive, careful management of both financial and environmental costs will be crucial to scalability.

Organizations are just beginning to incorporate sustainability measures into the Gen AI lifecycle, with 31% of executives saying their organization has taken steps to this end. As many as 74% of executives find measuring Gen AI's environmental impact challenging due to limited transparency from hyperscalers and model providers. They expect the tech sector to lead efforts to normalize and streamline measurement and transparent reporting of the environmental impact of Gen AI.

How to create responsible Gen AI for sustainable business value?

Gen AI has the potential to accelerate sustainable business priorities, control costs, and drive growth. However, organizations must weigh these benefits against the technology's environmental costs. We suggest the following approach to make informed, responsible business decisions:

- **Identify the right technology that addresses your business needs:** It is important to note that Gen AI is just one element of the broader tech landscape, and solving most business problems requires a combination of different techniques. “Hybrid intelligence” – a convergence of traditional AI, Gen AI, and technologies such as automation, robotic process automation (RPA), etc. – can unlock new levels of ingenuity and efficiency. Nearly three-quarters (74%) of executives believe that choosing the appropriate technology (be it AI, Gen AI, analytics, or a combination of different technologies) that addresses your business needs is crucial to reducing the environmental footprint of Gen AI and harnessing its full potential.

Executive summary

- **Assess and mitigate Gen AI's environmental impact:**
 - **Build the business case for sustainable Gen AI:** In addition to being more responsible from an environmental perspective, prioritizing sustainable Gen AI initiatives offers significant business advantages, notably cost reductions and acceleration of work. It is important to showcase the business case to top leadership and take into account the incurred environmental cost.
 - **Evaluate Gen AI partners and models on sustainability parameters:** More than half (55%) of executives believe that including sustainability as a key criterion in vendor selection for all Gen AI-related requirements would reduce environmental footprint. It is crucial that organizations select the most appropriate model. For example, when comparing large and small models, organizations must decide on the optimum balance between performance and energy consumption. Decisions on using prebuilt or custom models also impact computational power and carbon footprint, with the former demanding greater resources.
 - **Monitor and report your Gen AI footprint:** Analysis and accurate measurement, monitoring, and tracking are paramount. Organizations should also communicate their sustainability intentions clearly to stakeholders, disclosing emissions levels, detailing progress transparently, and setting definite goals. Of those measuring the environmental footprint of Gen AI (12% of our survey sample), only 28% disclose it, and only 24% have set targets to reduce it.
 - **Implement sustainable practices throughout Gen AI's lifecycle:**
 1. **Hardware-related measures:** Prioritize partners with more energy-efficient and recyclable hardware specifically designed for AI/Gen AI.
 2. **Model architecture and algorithm-related measures:** Use smaller, task-specific pre-trained models. Consider optimizing model size through techniques such as model compression, pruning, quantization, and knowledge distillation to significantly lower cost and energy consumption.

Executive summary

- 3. Sustainable-infrastructure measures:** Consider providers which use low-carbon energy sources such as solar, wind, geothermal, and nuclear to power AI/Gen AI infrastructure and data centers. Select cloud providers that utilize energy-efficient green data centers. Choose a region for your server that ensures smaller environmental impact. Additionally, consider utilizing edge computing devices to reduce data transfer and distribute associated energy usage.
 - 4. Sustainable usage:** Tracking and quantifying the carbon footprint of Gen AI applications is critical to eliminating unnecessary usage. Consider implementing batch processing and prompt optimization techniques such as prompt caching or concise chain of thought (CCOT) to ensure efficient processing.
- **Tap into Gen AI's potential by investing in the right use cases to accelerate sustainable business value:** One-third (33%) of executives say they have already started using Gen AI for sustainability initiatives – with half of

them at the pilot stage currently. An additional 37% say they have started to explore its potential for sustainability. Moreover, two in three (66%) executives say they expect a reduction of more than 10% in GHG emissions in the next 3–5 years as an output of Gen AI-led sustainable business initiatives. However, this assumption needs to be taken with caution, given the limited number of organizations which measure the environmental footprint of their Gen AI use.

When it comes to using Gen AI to accelerate progress toward business and sustainability goals, it is key to identify suitable use cases. Organizations should carefully identify and prioritize the most appropriate sustainability use cases for Gen AI based on financial and environmental costs and the expected sustainability and business benefits.

We looked at more than 100 use cases across functions and sectors and assessed them across two dimensions: the complexity of implementation and potential to create a sustainable business impact. A few quick-wins emerged from our analysis, including ESG reporting, sustainable product design, life cycle assessment (LCA), supplier

Executive summary

sustainability reporting, virtual assistance, sustainable IT governance, and ESG scenario planning among others. These use cases need to be weighed on environmental cost (energy, water, carbon) of using Gen AI and the business and environmental benefits the technology offers. Our analysis also reveals the potential of Gen AI to accelerate progress on UN Sustainable Development Goals (SDGs). It should be noted that in many cases organizations plan to deploy a combination of technologies to achieve a more comprehensive approach to problem solving and innovation.

- **Develop the right data and technology foundations:** Only 37% of executives claim their organization has the right data-management tools and technologies for Gen AI, and only one-third (33%) evaluate and monitor data quality for Gen AI. Building the right data foundations and developing the required skillsets are the keys to deriving maximum benefits from Gen AI. Organizations can also evaluate the potential of AI agents to create sustainable business value in ESG reporting and compliance-related areas.

- **Govern for sustainable Gen AI:** Implementing a governance model for safe, transparent, sustainable, and ethical usage is also imperative. Most executives (62%) believe robust guardrails and governance can effectively mitigate Gen AI's environmental impact. Nearly half (49%) also rank the lack of clear governance models among the top five challenges of implementing Gen AI for sustainability. Partner with technology partners, startups, research institutions, sustainability experts, and governments to share best practices, develop sustainable Gen AI standards, and harness Gen AI to accelerate sustainable business goals.

42%

of executives have had to relook at their climate goals due to Gen AI's growing environmental footprint.

Who should read this report and why?

Who?

This report should speak to technology and business leaders across functions, but **data, digital, and sustainability leaders** will find it particularly helpful.

Gen AI models are resource-hungry – with a substantial carbon, energy, water, and material footprint. It is imperative that, as the technology develops, it remains within the guardrails of environmental sustainability.

Why?

This report explores the environmental impact, as well as the potential of Gen AI to drive sustainable business value. We analyze organizational awareness and priorities and look at potential use cases. We recommend a series of actions for organizations to minimize Gen AI's environmental impact and maximize its sustainability potential, while managing cost and performance, and sustaining the impetus of technological change.

The report presents a detailed five-step approach to developing a sustainable Gen AI. It draws on comprehensive research building on our internal expertise and a survey of 2,000 senior executives (director level and above) at organizations that have

annual revenue above \$1 billion and are already working on Gen AI initiatives. These organizations are based in 15 countries: US, Canada, Brazil, UK, France, Germany, Italy, Spain, the Netherlands, Norway, Sweden, India, Australia, Singapore and Japan. The survey spans 12 key industries and sectors: aerospace and defense, automotive, banking and capital markets, consumer products, energy and utilities, insurance, life sciences, manufacturing, public sector/government, retail, telecom, and high tech. The report also includes qualitative findings from industry leaders.

01

**Gen AI has a
significant
environmental impact**

Generative AI (Gen AI) has rapidly gained traction in the business and consumer world. Our recent report on consumer trends highlights that 58% of consumers in 2024 have replaced traditional search engines by Gen AI tools for product recommendations – up from 25% in 2023.⁴ In the business world as well, Gen AI can replicate – and, in some respects, surpass – human thought processes, synthesizing tailored content with far-reaching implications for driving innovation, enhancing customer experience (CX), streamlining operational efficiency, and boosting growth. Our recent research reveals that organizations recognize the vast potential of Gen AI: 80% have increased investment in the technology in the past 12 months. Moreover, while only 6% had integrated Gen AI across their business functions and locations as of the end of 2023, that figure had risen to 24% as of October 2024.⁵

However, these valuable advantages come at a cost that goes beyond the monetary. It is important to recognize and address the energy consumption, carbon footprint, water usage, and e-waste entailed in the implementation of Gen AI throughout its lifecycle.

80%

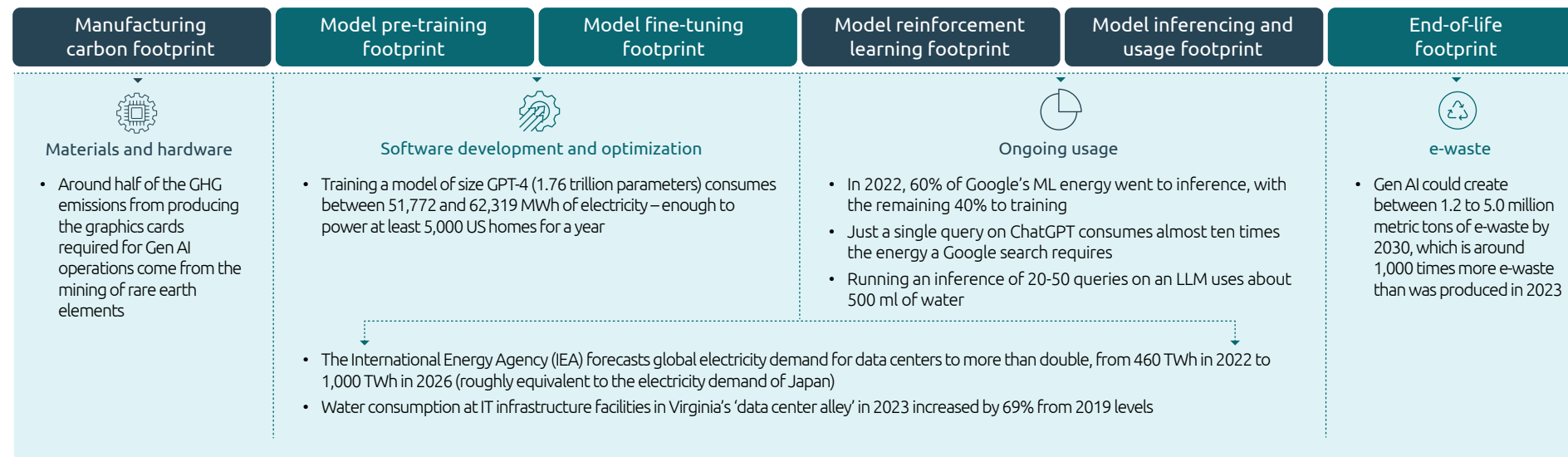
of the organizations have increased investment in Gen AI in the past 12 months.

Throughout its lifecycle, Gen AI has a considerable environmental impact

From manufacturing (encompassing materials and hardware impact), model training, and usage (including data centers' energy, water, and carbon impact) to end-of-life (e-waste), Gen AI consumes vast quantities of resources and leaves notable financial and environmental footprints. Figure 1 highlights Gen AI's environmental impact throughout its lifecycle, based on secondary sources.

Figure 1.

Gen AI's environmental impact across its lifecycle



Source: Capgemini Research Institute analysis, Harvard Business Review, "How to make generative AI greener," July 2023, IEA, Electricity 2024: Analysis and forecast to 2026, January 2024, Financial Times, "US tech groups' water consumption soars in 'data centre alley'," August 2024, Vox, "AI already uses as much energy as a small country. It's only the beginning," March 2024, OECD, "How much water does AI consume? The public deserves to know," November 2023, ARXIV, "The carbon footprint of machine learning training will plateau, then shrink," April 2022, Frontline Magazine, "E-waste from AI computers could 'escalate beyond control': study," October 2024.

Gen AI's hardware requirements put a strain on natural resources and habitats

Gen AI is consuming significant amounts of energy and resources in data centers around the world. For example, it uses thousands of graphics processing units (GPUs). GPU chips require 10–15 times more power to operate than a traditional central processing unit (CPU).⁶ (We note, however, that this usage could be at least partially offset by the greater energy efficiency of GPUs due to their ability to perform many more calculations simultaneously.)⁷

GPU chips and the other hardware that Gen AI requires are often made from copper, cobalt, tungsten, lithium, germanium, palladium, lead, chromium, cadmium, mercury, and other earth metals. Around half of the greenhouse gas (GHG) emissions in the production of the graphics cards required for the operation of Gen AI comes from the mining of earth metals.⁸

Gen AI typically requires a wider array of hardware than other types of computing and cycles through that hardware at a faster rate, requiring more frequent replacement.⁹ These shorter use phases will naturally accelerate the harmful effects on habitats and more rapidly deplete resources.¹⁰

Gen AI models are energy-hungry

Srini Koushik, President of AI, Technology, and Sustainability at Rackspace Technology, a US-based multi-cloud solution provider across apps, data, and security, says: *“As it exists today, AI and sustainability take you in opposite directions. AI consumes a lot of power, whether it’s training large language models (LLMs) or running inference. And this power consumption is growing exponentially.”*¹¹ Larger models (that include more parameters and therefore require more training data) generally consume more energy and generate more carbon in the **training process**:

- Estimates suggest that training a GPT-3 model with 175 billion parameters consumes nearly 1,300 MWh of electricity, roughly the same amount of power consumed by 130 US homes in a year.¹²
- Moving to the next model size up, GPT-4, with 1.76 trillion parameters, power consumption of training is estimated at between 51,772 and 62,319 MWh – enough to power 5,000 US homes for a year (at a conservative estimate).¹³

Following the training process (a one-time event in the model’s lifecycle) is the **inference phase** – essentially

where the model is deployed. This demands as much or more energy than the training phase, with the energy requirement for inferencing expected to increase exponentially as higher numbers of people become regular users of Gen AI. As per estimates, 60% of Google’s machine learning (ML) energy use in 2022 went to inference, with the remaining 40% to training.¹⁴ According to the International Energy Agency (IEA), just a single query on ChatGPT consumes 2.9 watt-hours of electricity – almost ten times what a Google search requires. If we assume around 9 billion daily searches (the estimated total daily searches conducted on Google), running these searches on ChatGPT would require an additional 10 terawatt-hours of electricity annually.¹⁵ This is equivalent to the annual electricity consumption of 1.5 million EU citizens.¹⁶

Reinforcement learning methods, such as reinforcement learning from human feedback (RLHF) and reinforcement learning from AI feedback (RLAIF), also leave a considerable environmental footprint.

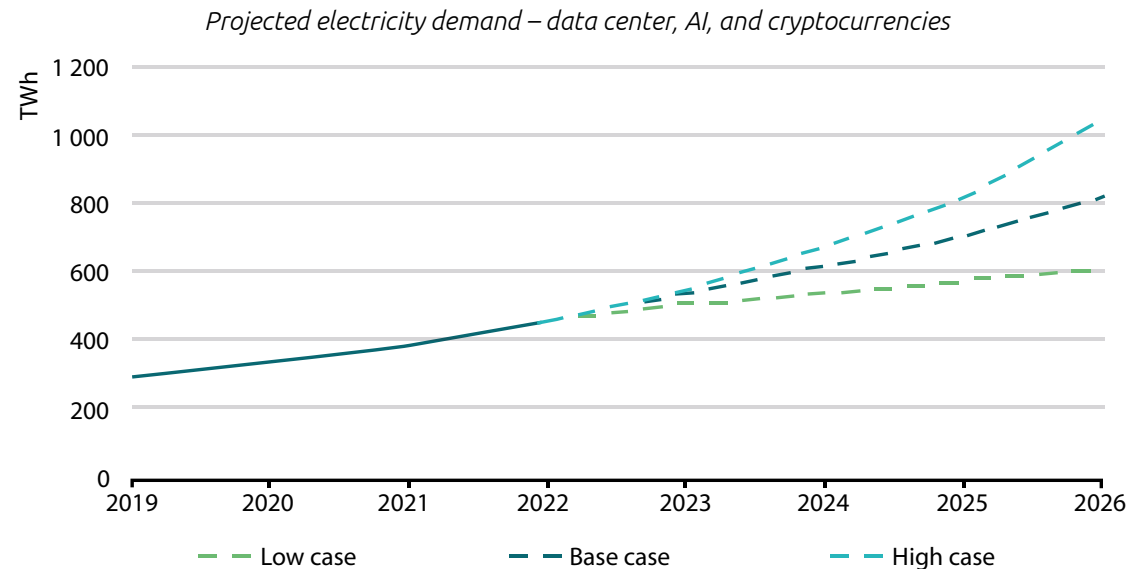
Gen AI training and inferencing drive up energy requirements in data centers

IEA forecasts global **electricity demand** for data centers to more than double, from 460 terawatt-hours in 2022 to 1,000 terawatt-hours in 2026 (roughly equivalent to the electricity demand of Japan), driven primarily by AI (see figure 2).¹⁷ Goldman Sachs estimates that the share of power demand of data centers will form 3–4% of the global power demand by 2030.¹⁸

This surging electricity demand for AI workloads in data centers is already impacting the **GHG emissions** levels of hyperscalers. Microsoft reported a 31% increase in Scope 3 emissions since 2020, primarily due to the expansion of data centers.¹⁹ Google also reported a 48% increase in GHG emissions from 2019 levels, owing to rising data-center energy consumption and supply chain emissions.²⁰ To try to sate the ever-growing energy demand of their data centers, organizations are planning to focus on nuclear energy projects.²¹ Over the last year, Google Cloud,²² AWS,²³ and Microsoft²⁴ announced plans to use small modular reactors (SMRs) to power their data centers.

Figure 2.

Global electricity demand is on a steep upward curve



Note: Includes traditional data centers, dedicated AI data centers, and cryptocurrency consumption; excludes demand from data-transmission networks.

Source: IEA.

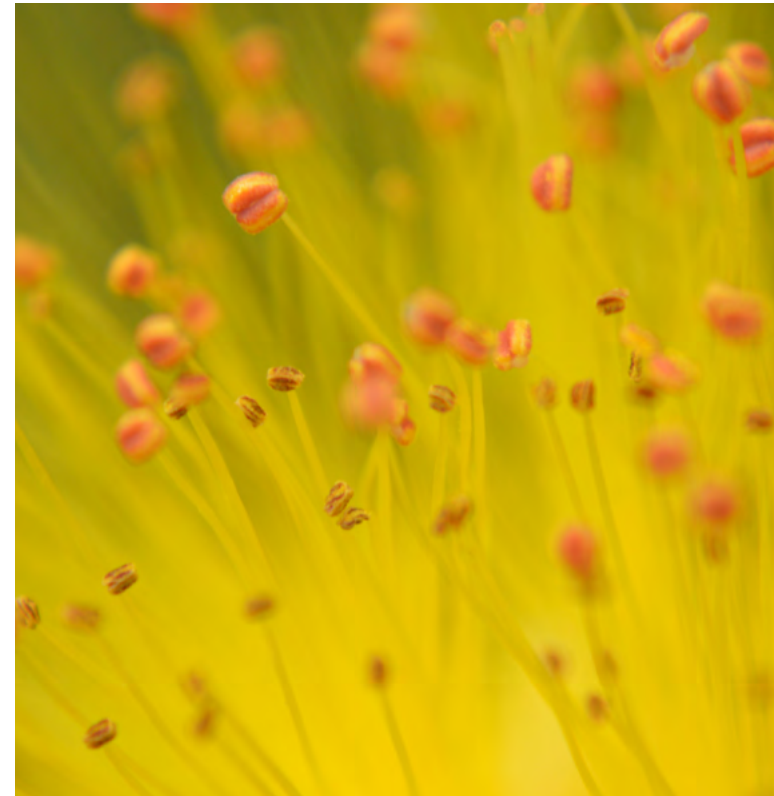
Gen AI models are also water-thirsty

Another consequence of the growth of data center operations is the huge increase in fresh, clean **water** demand to prevent overheating. Water consumption at IT infrastructure facilities in Virginia’s “data center alley” in 2023 increased by 69% from 2019.²⁵ Running an inference of 20–50 queries on an LLM such as GPT-3 uses about 500 ml of water each time.²⁶ Our previous research shows that, as of 2023, around 11% of consumers had replaced traditional search engines with Gen AI tools²⁷— a trend that we expect to grow. If GPT-3 took over all 9 billion daily Google searches,²⁸ it would require 4.5 billion liters of water daily to cool the ensuing data-center operations. This anticipated daily water requirement looks quite substantial, considering almost half the world’s population may face severe water stress as soon as 2030.²⁹ With Gen AI model training and inferencing pushing up the focus on nuclear energy, the amount of water required for cooling nuclear power plants should also be taken into consideration. Moreover, estimates suggest that the production of a single microchip, which is extensively used within the Gen AI landscape, requires approximately 2,200 gallons (8,328 liters) of ultra-pure water (UPW).³⁰ Approximately 10 million gallons (39 million liters) of ultra-pure water is used per day by an average chip manufacturing facility, which is equivalent to water used by 33,000 US households every day.³¹

Widespread Gen AI adoption will see e-waste levels shoot up

Mark Kidd, EVP and General Manager at Iron Mountain Data Centers, says: *“E-waste is one of the fastest-growing waste streams in the world. Annual e-waste production is on track to reach a staggering 75 million metric tons by 2030. Just 17% of global e-waste is documented to be collected and properly recycled each year.”*³²

As the use of Gen AI becomes more widespread and profound, the e-waste challenge and the cost associated with it will grow correspondingly. The limited lifespan of Gen AI hardware will fuel this issue. Some estimates suggest that Gen AI could create between 1.2 to 5.0 million metric tons of e-waste by 2030, which is around 1,000 times more e-waste than was produced in 2023.³³ E-waste largely ends up in landfills, where harmful chemicals such as mercury, lead, bromine, and arsenic leach out from the electronics, polluting soil and consequently endangering the health of wildlife, livestock, and people in the surrounding area. The adoption of circular economy measures among hardware manufacturers and data center operators is imperative to tackle this.

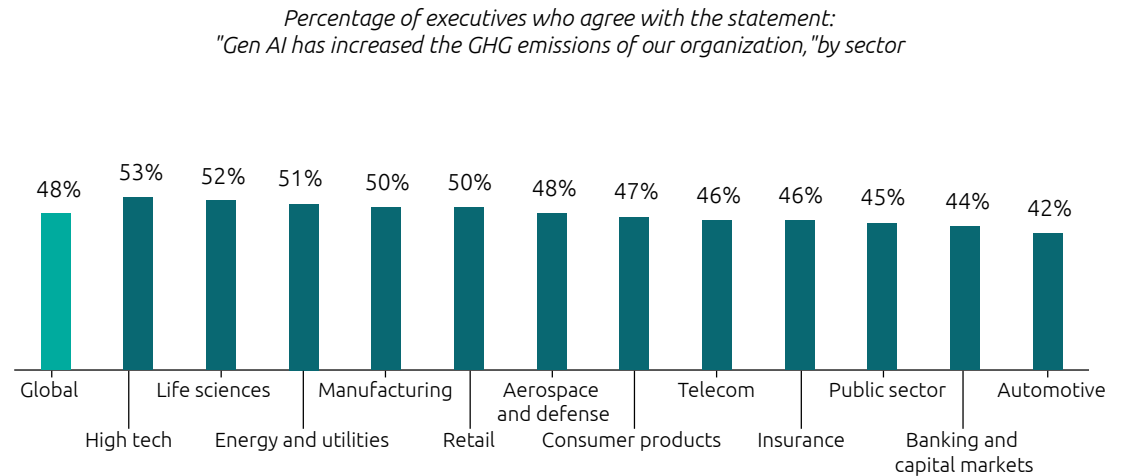


Gen AI is one of the reasons for the rise in GHG emissions in nearly half of organizations

In our research, we surveyed executives from 2,000 organizations that are already working on Gen AI initiatives. A majority (72%) of the executives agree that the environmental impact of Gen AI is higher than traditional AI models. Nearly half (47%) say their organization's GHG emissions have increased in the past 12 months by nearly 6% on average. Executives believe that Gen AI is one of the reasons for this rise in emissions. As figure 3 shows, a similar proportion of executives say Gen AI has driven a rise in their organization's overall GHG emissions.

Figure 3.

Half of executives agree that Gen AI has increased their organization's GHG emissions



Source: Capgemini Research Institute, Gen AI and Sustainability survey, August-September 2024, N = 2,000 executives from organizations that are working on Gen AI initiatives.

Among the organizations that measure the environmental impact of Gen AI, more than half (51%) say Gen AI is one of the reasons for the rise of GHG emissions of their organizations. They also expect the technology's emissions as a proportion of carbon emissions from internal operations to increase from 2.6% to 4.8% in the next two years.

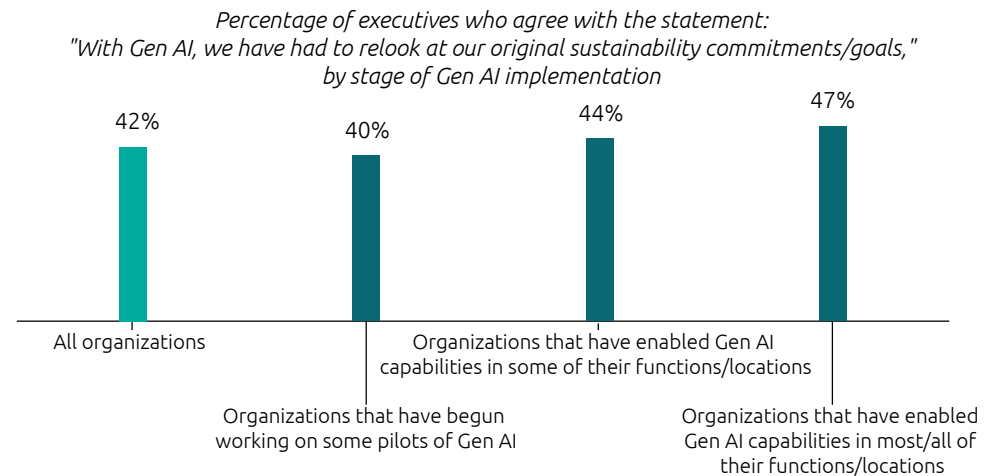
Consequently, organizations that have already started working with Gen AI are re-evaluating their climate goals; 42% of executives in our research agree. Within organizations advanced in Gen AI implementation (those implementing Gen AI across most functions/locations), nearly half (47%) have had to relook at their sustainability commitments (see figure 4). Google's Chief Sustainability Officer, Kate Brandt, explained in a press interview: ***"Reaching the net zero goal by 2030 is extremely ambitious. It will require us to navigate a lot of uncertainty, including around the future of AI's environmental impacts."***³⁴

48%

of the executives say that Gen AI is one of the reasons for rise in GHG emissions of their organization.

Figure 4.

Nearly half of executives from advanced organizations say they have had to relook at their sustainability commitments due to Gen AI



Source: Capgemini Research Institute, Gen AI and Sustainability survey, August-September 2024, N = 2,000 executives from organizations that are working on Gen AI initiatives, N = 1,236 executives from organizations working on Gen AI pilots, N = 636 executives from organizations enabling Gen AI capabilities in some of their functions/locations and N = 128 executives from organizations enabling Gen AI capabilities in most of their functions/locations.

02

The sustainability of Gen AI remains a low priority

Most organizations don't measure the impact of Gen AI

Our research confirms that only a few executives are currently aware of the extent of the overall environmental impact of Gen AI. For example, only 28% of executives in our research were aware that, on average, a Gen AI query requires nearly 10 times as much electricity to process as a Google search. Only 31% were aware that training an LLM at a US-based data center consumes around 700,000 liters of fresh water. Moreover, only 38% claim to be aware of the environmental impact of Gen AI they use.

Mauli Tikkiwal, a board member at UK-based Orchard Hill College and Academy Trust, says: ***“First, you must identify the impact so you can track and reduce it.”*** However, only one in ten executives (12%) says that their organization actively measures their Gen AI footprint, while a majority (82%) plan to start in the next 12–24 months (see figure 5). They give a range of reasons for this omission:

74%

of executives say a lack of transparency from Gen AI providers makes measurement challenging

64%

cite the complexity of tracking energy consumption across various applications

58%

say driving efficiency is more important than measuring impact



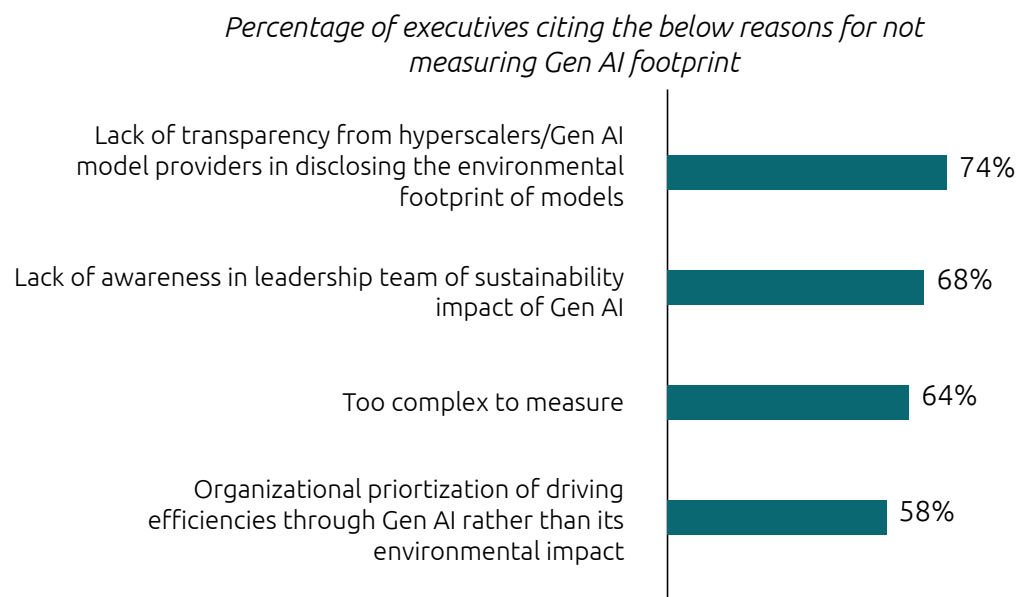
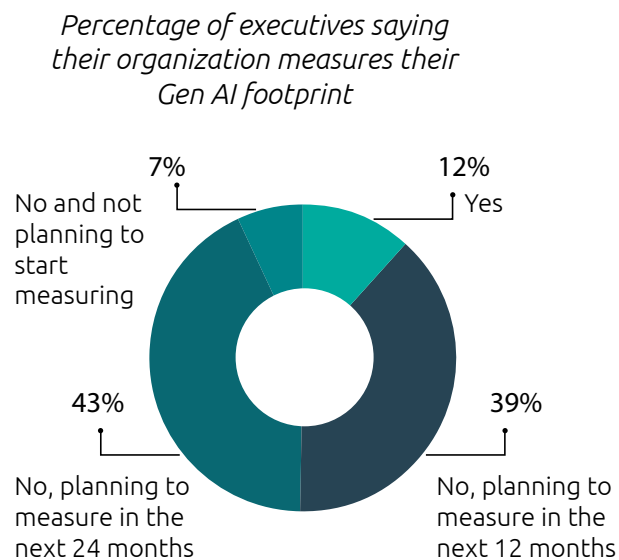
“First, you must identify the impact so you can track and reduce it.”

Mauli Tikkiwal

Board member at UK-based Orchard Hill College and Academy Trust

Figure 5.

Only 12% of organizations measure the environmental impact of Gen AI



Source: Capgemini Research Institute, Gen AI and Sustainability survey, August-September 2024, N = 2,000 executives from organizations that are working on Gen AI initiatives.

Source: Capgemini Research Institute, Gen AI and Sustainability survey, August-September 2024, N = 1,767 executives from organizations that are currently not measuring their Gen AI footprint.

Organizations look to the tech sector to drive sustainable Gen AI

A majority (78%) of the executives in our research say their organization is using pre-trained Gen AI models, and only 4% have built their own models from scratch. Among those using pretrained models, 63% contract them as a service through hyperscale cloud providers, giving rise to a reliance on tech providers for measurement and tracking. However, as figure 5, above, highlights, nearly three-quarters (74%) of executives cite lack of transparency in disclosure and reporting of Gen AI's environmental impact by Gen AI providers as an obstacle to accurate measurement and reporting. Clearly, organizations expect transparency from model providers on Gen AI's footprint. Niklas Sundberg, Chief Digital Officer and Senior Vice President at Swiss global transport and logistics company Kuehne+Nagel, says: *"You should be able to ask Copilot or ChatGPT what the carbon footprint of your last query is but none of the tools will give you a response to that question at the moment."*³⁵

Organizations also expect the tech sector to develop innovative mitigation measures. Eszter Haberl, Sustainability Business Strategy Director at India-based auto component manufacturing company, Motherson Group, says: *"The landscape of interaction between generative AI, sustainability, and ESG goals is very complex, particularly in manufacturing. While generative AI offers transformative potential for optimizing production processes and reducing waste, it also presents significant energy challenges. Focusing on balancing these trade-offs by implementing energy-efficient AI systems that help optimize production line efficiency and reduce material waste, while simultaneously working with tech partners to minimize the computational footprint of these systems through innovations in green computing and energy-efficient data centre operations is key."*

74%

of executives cite lack of transparency in disclosure and reporting of Gen AI's environmental impact by Gen AI providers as an obstacle to accurate measurement



"While generative AI offers transformative potential for optimizing production processes and reducing waste, it also presents significant energy challenges."

Eszter Haberl

Sustainability Business Strategy
Director at Motherhood Group

Technology firms are working to make Gen AI more sustainable

Technology players can play a key role to contain the environmental impact of Gen AI. Below we highlight a few initiatives taken by tech firms – from energy-efficient chips, smaller models, right-sizing AI/Gen AI workloads to investing in low-carbon energy sources.³⁶



Hardware

Nvidia's new Blackwell chip comes with 30 times improved performance for LLM workloads and 25 times lower energy consumption than the preceding iteration.

Rahul Sareen, General Manager, Cross-Industry and Sustainability Solutions, at **AWS**, says: *"Trn1 instances, powered by AWS Trainium chips, are purpose built for high-performance deep learning (DL) training of generative AI models. These are up to 25% more energy efficient than comparable accelerated computing EC2 instances. We will be launching Trainium2 instances that are designed to deliver up to 4x faster training performance, while improving energy efficiency (performance/watt) up to 2x than first generation Trainium chips."*



Model and algorithm optimization

Google's TensorFlow and **Hugging Face** have incorporated **quantization techniques** to reduce the size of models, thereby trimming power and resource requirements.

Liquid AI's Founders have developed a more adaptive, less energy-hungry neural network. Ramin Hasani, co-Founder and CEO of **Liquid AI**, expands: *"We have designed new algorithms that reduce time and memory complexity of foundation models, resulting in considerable energy and carbon-footprint savings."*



Data

AWS offers SageMaker Training, which enhances quality of training while **reducing reliance on high volumes of data**.

Meta has developed AI Research Store (AIRStore) to help **optimize data transfers** to minimize cross-region data traffic with end-to-end encryption.



Model training and inferencing

Microsoft's ML-based Project Forge software helps **schedule AI model training and inferencing workload during timeframes when the hardware has available capacity**, enhancing utilization rate by 80-90% across scale. The company also uses a phase-splitting tool, Splitwise, to increase inferencing efficiency by 2.35x.

Rahul Sareen from **AWS** says: *"There are other mechanisms as well. For example, customers can use SageMaker distributed libraries such as data parallelism and pipeline parallelism to run parallel computing across multiple GPUs or instances. And some of these approaches help maximize GPU utilization by splitting the training batches into micro batches."*

Salesforce also deploys **model training in data centers, with around 70% lower carbon emissions** than global average electricity use.

"continued to the next page"



Model hosting/Data center

Amazon, Microsoft, and Google are planning to focus on nuclear energy projects to power their data centers.

Meta has signed a deal with Sage Geosystems to purchase **geothermal energy** to power its US data centers. **AWS** offers its customers the option to train and deploy models in AWS regions where grids are run on 100% **renewable energy**.

Meta is working on next-gen AI-optimized data centers with denser racks for large-scale AI clusters, **liquid-cooled AI hardware** and network infrastructure, and dry-cooling technology.



Monitoring and reporting

Vishal Singhvi, From **Microsoft**, says: *"In the LLMs, we have introduced carbon monitoring or energy monitoring, and that's already available to the customer. They can see real-time carbon and energy monitoring and [use that] to take decisions. This is the carbon cost, which you can see in the model cards and, once you take into account all this, you can actually do cost-benefit trade-offs."*

Salesforce has implemented collation and publication of the **environmental impact data for AI solutions**, which includes training time, energy, and carbon.



Collaborations

Folake Fabunmi, Managing Executive, Strategic Partnerships at IBM, adds: *"We have launched open-source InstructLab in collaboration with Red Hat, which facilitates knowledge sharing and can be used to customize models and promote resource optimization. It gives communities the tools to create and merge changes to LLMs without retraining the model from scratch."*

Source: Capgemini Research Institute analysis, company websites, press articles.



"Customers can use SageMaker distributed libraries such as data parallelism and pipeline parallelism to run parallel computing across multiple GPUs or instances. And some of these approaches help maximize GPU utilization by splitting the training batches into micro batches."

Rahul Sareen

General Manager, Cross-Industry and
Sustainability Solutions at AWS

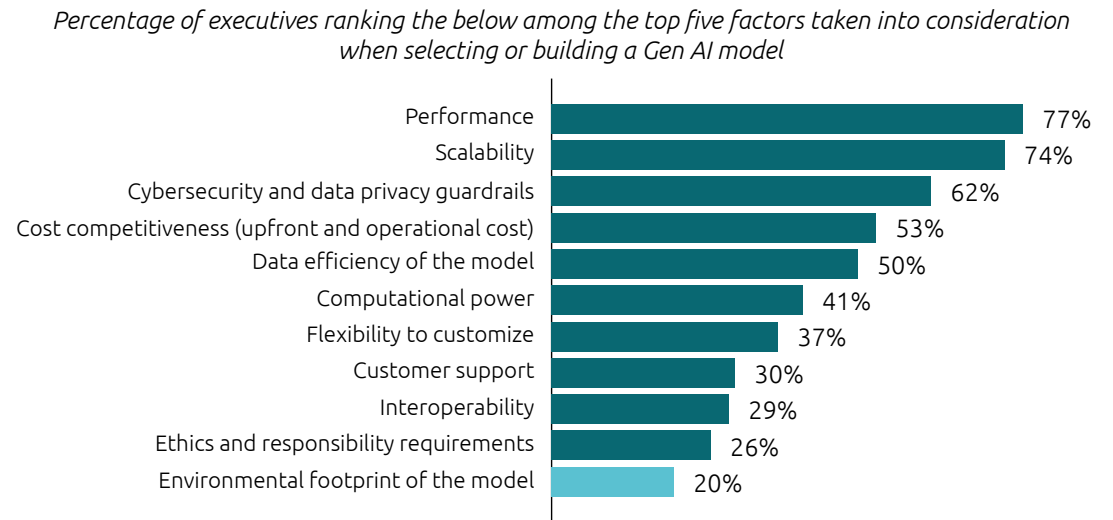
Sustainability isn't a consideration during Gen AI model evaluation

As with any new technology, organizations assess the business case for Gen AI prior to adoption across areas such as CX, productivity, operational efficiency, cost, revenue, and innovation. Factors such as the model's performance model and scalability potential across the organization are important considerations. However, only 20% of executives in our research rank "environmental footprint of Gen AI model" among the top five factors when selecting or building Gen AI models (see figure 6).

Almost six in ten executives (58%) agree that their organizations put a premium on the cost and quality benefits of Gen AI, with environmental footprint only being a minor consideration. This is especially true among executives from the high-tech sector (64%). A VP from a telecom company based in The Philippines elaborates: ***"Our (Gen) AI strategy hinges on five objectives. Number one is to increase revenue. Second is to enhance customer experience. The third is to improve employee productivity. Fourth is to improve network operations. And fifth is to reduce costs. Based on those five objectives, we have defined KPIs for any Gen AI use case. We don't yet have a KPI that measures sustainability elements of Gen AI. This is an area we need to look at."***

Figure 6.

The environmental footprint is given the least consideration when developing or choosing a Gen AI model



Source: Capgemini Research Institute, Gen AI and Sustainability survey, August-September 2024, N = 2,000 executives from organizations that are working on Gen AI initiatives.

Organizations are taking only a partial view of Gen AI costs

As figure 6 shows, while organizations are prioritizing financial costs in decision-making for selecting or building Gen AI models, the cost of these models from an energy and resources perspective is often overlooked. Many organizations assume that licensing, infrastructure, and data costs drive the cost of Gen AI models, but ignore some important factors:

- As Gen AI models become more complex and require more computational power, the energy requirements – and, consequently, the cost – will grow exponentially.
- As Gen AI models become more pervasive, the financial and energy costs for running inferences can act as a barrier to scalability of Gen AI.

Samuel Young, AI Practice Manager at Energy Systems Catapult, a UK-based independent research and technology organization, adds:

“Energy usage influences both cost and carbon impact. When implementing at scale, organizations quickly become sensitive to inference costs. They therefore have an incentive to adopt less energy intensive models, which can reduce carbon impact. However, beyond this incentive for cost efficiency they do not typically have a direct incentive to prefer clean energy sources to carbon intensive ones.”

ROI calculations of Gen AI should consider energy costs as a relevant cost factor. In addition to tapping smaller, fine-tuned models, using a multiple model approach also offers a range of latency, accuracy, and carbon footprint solutions. Using agentic AI also optimizes cost as well as lower energy use.³⁷

In the long run, organizations that manage to innovate and control the cost and energy usage of Gen AI will be better positioned for success. Cost and carbon impact go hand in hand, and addressing both will drive innovation, improve economics, and help make Gen AI more sustainable for the future.



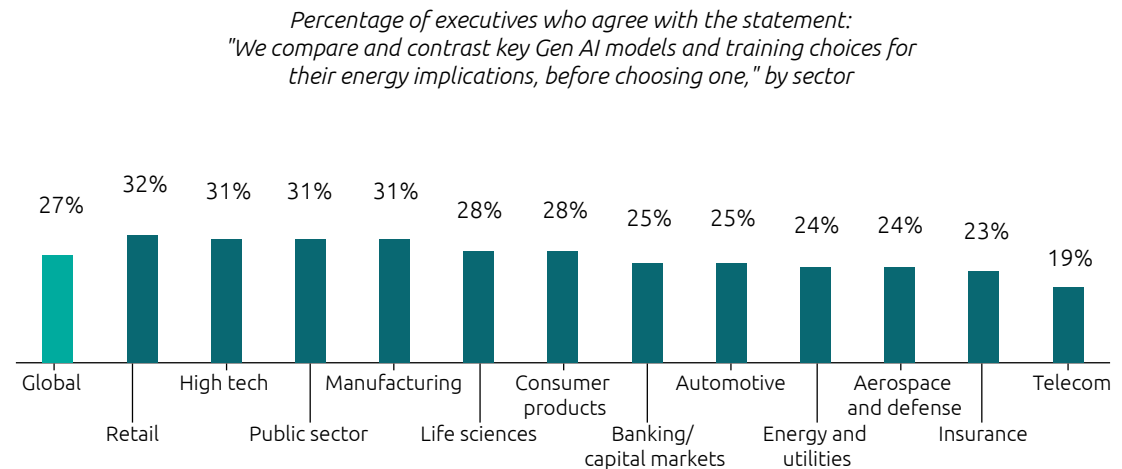
Only 27% of executives say they compare Gen AI models on the energy implications before selecting one (see figure 7). Vincent Caldeira, CTO, APAC at US-based software company Red Hat, adds: ***"Integrating sustainability impact assessments (SIAs) to evaluate the environmental impact of AI models during their development [is crucial]. This includes comparing different models and design choices based on their estimated carbon footprints and choosing the most sustainable options within the scope of acceptable performance."***³⁸

27%

of executives say they compare Gen AI models on the energy implications before selecting one.

Figure 7.

Only one in four organizations compares energy implications of Gen AI models and training choices



Source: Capgemini Research Institute, Gen AI and Sustainability survey, August-September 2024, N = 2,000 executives from organizations that are working on Gen AI initiatives.



"Energy usage influences both cost and carbon impact. When implementing at scale, organizations quickly become sensitive to inference costs. They therefore have an incentive to adopt less energy intensive models, which can reduce carbon impact."

Samuel Young

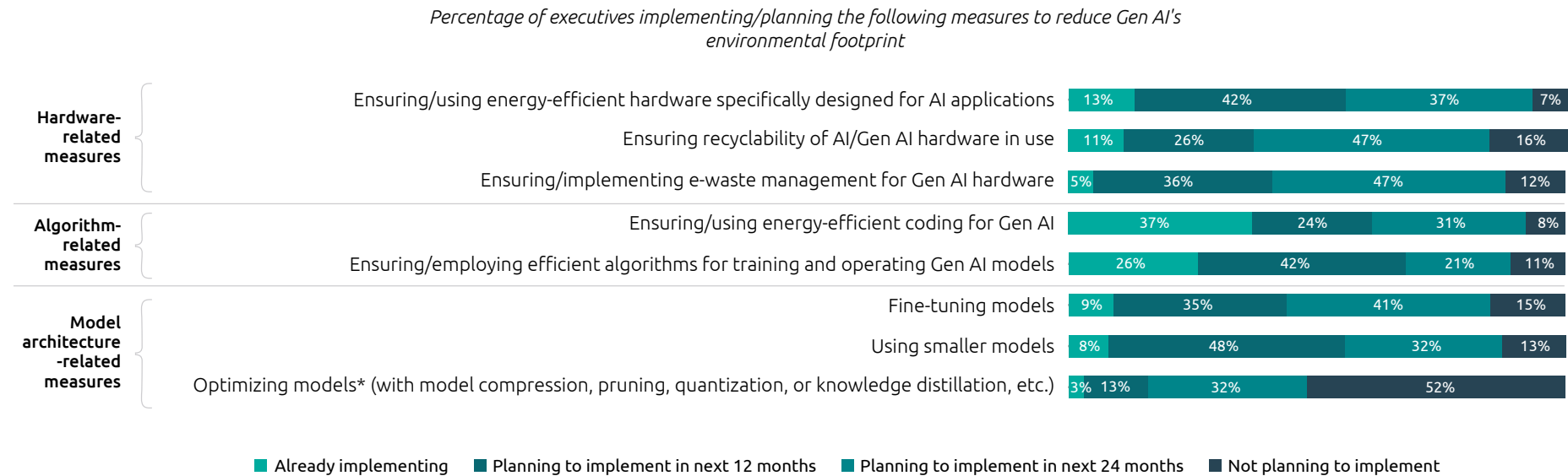
AI Practice Manager at
Energy Systems Catapult

Organizations are just beginning to incorporate sustainability measures throughout the Gen AI lifecycle

Many organizations fail to recognize Gen AI's environmental footprint as a significant issue and consequently fail to address it. Only 31% of executives say their organization has started to do so. Figure 8 shows the implementation maturity of selected measures that could reduce Gen AI's environmental impact. It should be noted that for some measures (such as hardware, data center, etc.), organizations will need to work with the partners and ensure that their partners are taking steps to contain the technology's environmental impact.

Figure 8.

Organizations are at the early stages of reducing Gen AI's environmental impact

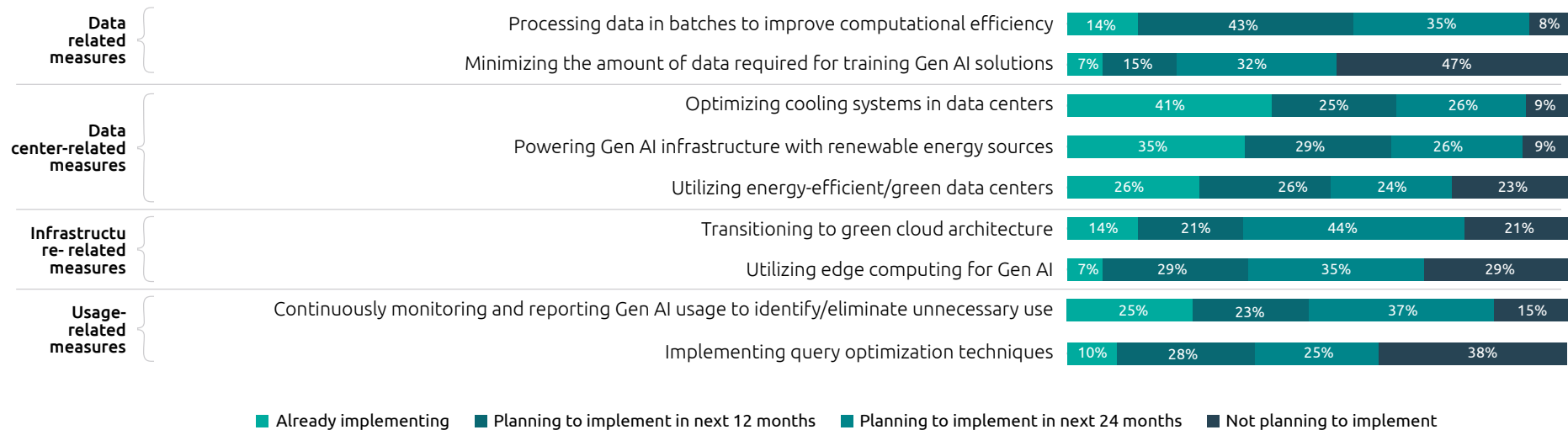


*Note:

- Model compression is a technique that reduces the size and complexity of large neural networks, making them more suitable for deployment on resource-constrained devices such as mobile phones.
- Model pruning is removing unimportant parameters from a neural network model to reduce the model size.
- Quantization is a model size reduction technique that converts model weights from high-precision floating-point representation (32-bit FP) to low-precision FP or integer representations (8-bit).
- Knowledge distillation is the process of transferring knowledge from a large model to a smaller one.

"continued to the next page"

Percentage of executives implementing/planning the following measures to reduce Gen AI's environmental footprint



Source: Capgemini Research Institute, Gen AI and Sustainability survey, August-September 2024, N = 660 technology executives from organizations that are working on Gen AI initiatives.

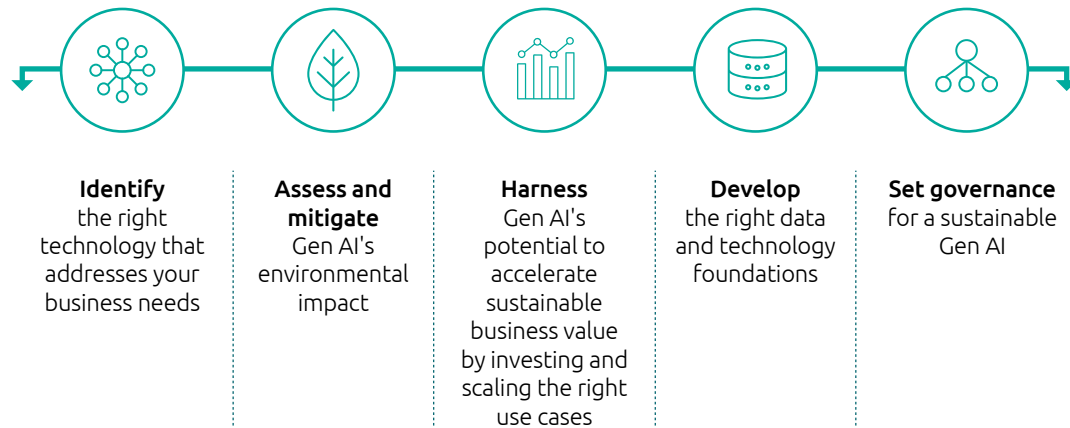
03

A roadmap to responsible Gen AI for sustainable business value

Gen AI has the potential to accelerate sustainable business initiatives, optimize costs, drive efficiencies and propel growth. Nevertheless, this potential must be weighed against the technology's environmental cost. Drawing on our survey analysis and interviews, as well as our experience in this area, we suggest the following roadmap:

Figure 9.

A framework to create responsible Gen AI for sustainable business value



Source: Capgemini Research Institute analysis.



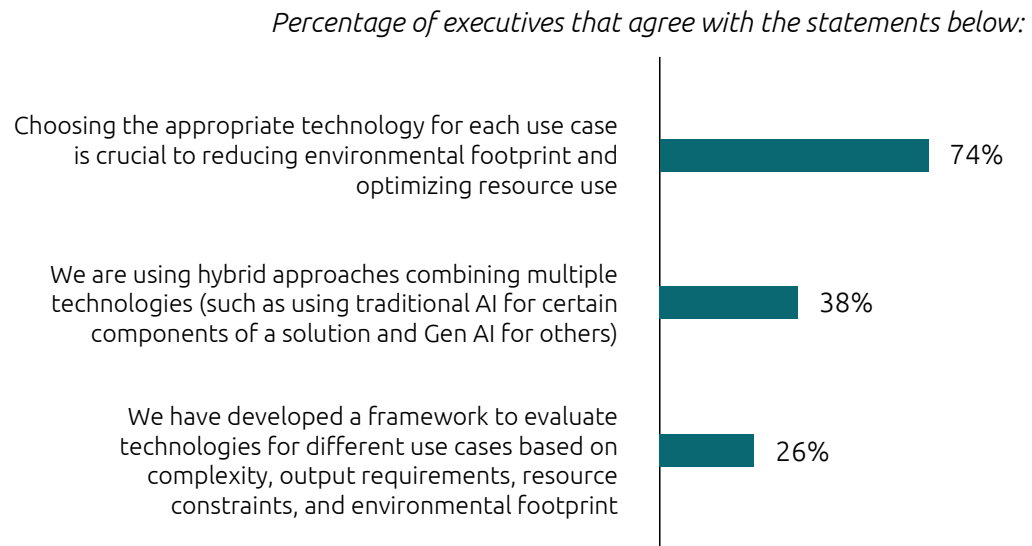
Identify the right technology that addresses your business needs

Gen AI is not a silver bullet. It is just one aspect of the broader tech landscape, and most business solutions require a combination of techniques. A convergence of traditional AI, Gen AI, and technologies such as automation, robotic process automation (RPA), etc., can unlock new levels of ingenuity and efficiency while retaining rigor and creativity, and with a significantly lighter environmental footprint. We call this convergence “hybrid intelligence.”

It is important for organizations to start with the problem, not the technology. Our survey reveals that nearly three-quarters (74%) of executives believe that choosing the appropriate technology (be it AI, Gen AI, analytics, or a combination of various technologies) that addresses your business needs is crucial to reducing the environmental footprint of Gen AI and harnessing its full potential responsibly. Moreover, 38% are already using hybrid approaches that combine multiple technologies. Nearly one-quarter have also developed frameworks to evaluate each technology before assigning it to a use case (see figure 10).

Figure 10.

Identifying the appropriate technology for each use case is vital



Source: Capgemini Research Institute, Gen AI and Sustainability survey, August-September 2024, N = 2,000 executives from organizations that are working on Gen AI initiatives.

Before selecting a solution for any use case, organizations should ensure they understand the trade-off between the incremental value gained in outputs versus additional resource utilization associated with each option. Vishal Singhvi, Director, Strategic Initiatives, Generative AI at Microsoft, elaborates: *“Everybody wants to do something with generative AI, but a lot of times you don't even need it. You can do this very well with your traditional AI, which consumes significantly less compute power and workloads.”*



“Everybody wants to do something with generative AI, but a lot of times you don't even need it. You can do this very well with your traditional AI, which consumes significantly less compute power and workloads.”

Vishal Singhvi

Director, Strategic Initiatives,
Generative AI at Microsoft



Assess and mitigate Gen AI's environmental footprint

Build the business case for sustainable Gen AI

In addition to being more responsible from an environmental perspective, prioritizing sustainable and responsible Gen AI initiatives can offer significant business advantages. The costs (energy or water costs) associated with taking on a Gen AI

model can be daunting. However, by adopting a sustainable mindset, organizations can reduce these costs and improve the economic viability of AI applications. Initiatives such as using small language models (SLMs, models trained on smaller, more specific datasets than those used to train LLMs and optimized for specific tasks or domains), energy-efficient hardware and code, green data centers, and optimized cooling systems not only promote sustainability but also bring significant cost savings. Arthur Mensch, CEO of Mistral AI, highlights the cost-effectiveness, resource-efficiency and environmental friendliness of sustainable Gen AI SLMs: *"Smaller models mean the applications are*

less costly to run and, more importantly, if you have a model that is 100 times smaller, you can call it 100 times more for the same cost, bringing a little more intelligence to your application with each call. We call this 'compressed knowledge.'"³⁹

Furthermore, Sujata Kukreja, General Counsel and Chief Compliance Officer for enterprise connectivity solutions specialist Expereo says: *"Companies that prioritize sustainable AI practices are even more likely to draw and retain top talent in a competitive job market, as they will attract conscientious and purpose-driven experts in their field."*⁴⁰

Evaluate Gen AI partners and models on sustainability parameters

Sustainability must be a key criterion in selecting partners and models

As figure 11 indicates, a majority (55%) of executives believe that including sustainability as a key criterion in vendor selection for all Gen AI-related requirements would reduce environmental footprint. However, only 15% of executives are currently fulfilling this.

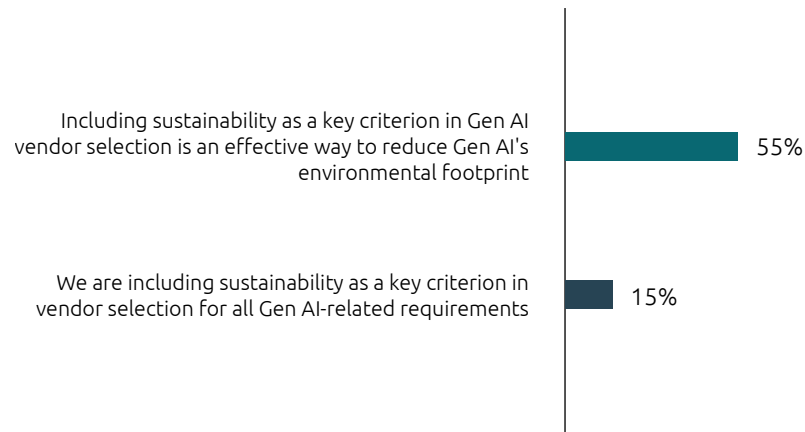
Boris Gamazaychikov, Senior Manager of Emissions Reduction at Salesforce, agrees: *“For companies not directly training models, the key is demanding transparency from AI providers on energy efficiency and carbon footprints – treating it like fuel efficiency for a new car purchase.”*⁴¹

You should also evaluate the energy sources of your providers. Google offers a “Carbon Sense Suite” to help organizations reduce energy consumption in their cloud workloads. Users of cloud providers can monitor public reports on sustainability progress.⁴²

Figure 11.

Gen AI’s environmental footprint is not a vendor or model selection parameter today

Percentage of executives that agree with the statements below:



Source: Capgemini Research Institute, Gen AI and Sustainability survey, August-September 2024, N = 2,000 executives from organizations that are working on Gen AI initiatives.

Choose the right model for your use cases

Selecting the right model is critical to reducing the environmental footprint and resource usage. In each case, organizations should consider a few key points:

- **Small versus large models:** Finding the efficiency ‘sweet spot’ that offers maximum performance gains per unit of resource input is critical to sustainable Gen AI deployment.⁴³ Recent studies demonstrate that SLMs can be fine-tuned to achieve comparable or even superior performance to that of their larger counterparts in specific tasks.⁴⁴

SLMs bring many benefits, notably their capacity for rapid training and faster inference speed, contributing to a more sustainable footprint. In addition, SLMs strike a harmonious balance between performance and resource efficiency. Training SLMs is much more cost-effective due to the reduced number of parameters. Offloading the processing workload to edge devices further cuts infrastructure and operating costs. Andrew Ng, Founder of LandingAI and Managing General Partner at AI Fund, told us: *“For the most complex reasoning tasks, an LLM is much more effective. But if you want a grammar checker, then you do not need a trillion-parameter model trained in science, philosophy, and ancient history. A SLM works just fine for this sort*

*of task. Also, there are a lot of use cases for SLMs that can run locally on device, for reasons of privacy and security.”*⁴⁵

- **Prebuilt versus tailored models:** Although tailored models can be customized to meet your specific or deep-domain expertise requirements, training them, not to mention maintenance and updates, demands considerable computational power and energy. For use cases with less stringent requirements, a general prebuilt model is preferable.

Monitor and report your Gen AI footprint

Conducting a lifecycle analysis to estimate the environmental impact of Gen AI solutions will be the key to mitigating its impact. The factors to consider are:

- Location of model training
- Type and quantum of hardware used – e.g., GPU, TPU, or CPU
- Training type: pre-training or fine-tuning
- Estimated carbon footprint of the model, calculated in real time
- Token consumption,⁴⁶ etc.

Organizations can use tools such as [CodeCarbon](#), [ML CO2 Impact](#), [Cloud Carbon Footprint](#) and [Green Algorithms](#),

to measure, monitor, and track Gen AI footprints and conduct an impact analysis before scaling. Capgemini’s [Sustainable AI](#) framework also focuses on measuring the environmental impact of AI based on life cycle assessment (LCA) and Green AI framework. Our tools enable organizations to estimate the multi-criteria environmental footprint of their AI systems, allowing them to assess their AI impact and align with current trends in AI and global sustainability.

Organizations should also include AI carbon footprint in their disclosures and reporting and set reduction targets. Currently, only 28% of those measuring Gen AI footprint disclose it, and only 24% have set targets (see figure 12).

The lifecycle assessment should also consider the impact of activities that Gen AI is complementing or replacing in order to evaluate the overall net impact.

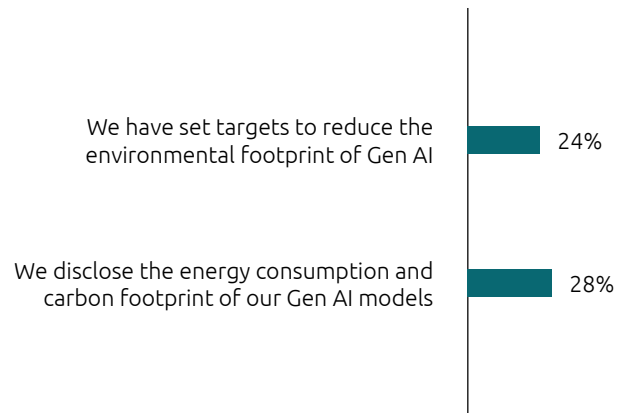
24%

of organizations have set targets to reduce the environmental footprint of Gen AI.

Figure 12.

Of the organizations measuring Gen AI's environmental footprint only one-quarter disclose it

Percentage of executives that are measuring Gen AI's environmental footprint that agree with the statements below:



Source: Capgemini Research Institute, Gen AI and Sustainability survey, August-September 2024, N = 233 executives from organizations that are measuring the footprint of Gen AI initiatives.



“We are looking at different deployment models, even switching programming languages, which can reduce the energy consumption of an application by up to 50%.”

Folake Fabunmi

Managing Executive,
Strategic Partnerships at IBM

Implement sustainable practices throughout Gen AI's lifecycle

Organizations can adopt a range of practices at development, deployment, and operational phases to cut down Gen AI footprints. Ricardo Guerra, CIO of Latin America's biggest bank, Itaú Unibanco Group, says: *"We continually optimize our AI models by reducing energy consumption, often using smaller models. We're learning when to use different solutions but with an emphasis on sustainable data centers and green technologies. We're also monitoring the market closely and prioritizing providers that offer green solutions. Lastly, we're trying to collaborate with other players and stakeholders for industry-wide sustainability initiatives."*⁴⁷ Below, we suggest a range of measures:

- Prioritize **smaller, targeted, domain-specific, pre-trained** Gen AI models over general-purpose models. Two in three (66%) technology executives in our research believe task-specific models will reduce Gen AI footprints.
- If managing hardware in-house, use more **energy-efficient and recyclable hardware** specifically designed for AI/Gen AI applications to reduce the operational carbon footprint, as well as embodied carbon footprint. When working with partners, prioritize those with advanced, energy-efficient hardware. A majority (60%) of tech executives also believe

that reducing power consumption of hardware during idle periods can contribute significantly.

- Explore upcoming **innovative and alternative sustainable hardware** such as neuromorphic or analog computing. Dr. Zina Jarrahi Cinker, a globally renowned exponential technology and deep science strategist, adds: *"Our brain can perform the equivalent of an exaflop – a billion-billion mathematical operations per second – and runs on 10–15 watts. Do you know what kind of power/energy our supercomputers run on? This is what the field of neuromorphic computing addresses. In the age of Gen AI, and the looming problem with energy consumption of LLMs, I believe [we could find] some of the answers in the world of biology."*⁴⁸ Liquid foundational models also have a reduced memory footprint compared to transformer architectures.

Some startups are also emerging in this space:

- Dutch startup Innatera uses biologically inspired spiking neural networks (SNNs). These networks use precisely timed events ("spikes"), to communicate as neurons in our brains do, yielding extreme energy efficiency.⁴⁹
- Gen AI startup Mythic is developing analog chips. The company promises exceptional performance on specific types of AI calculations, such as matrix

multiplications and signal processing, by going from digital chips (0/1 based) to continuous analog voltage signals. This higher degree of parallelism and efficiency leads to faster calculating speeds and a smaller energy footprint.⁵⁰

- Consider **model-optimization techniques** (such as quantization, distillation, and pruning) for deployment in resource-constrained devices or systems, without compromising performance. Folake Fabunmi, Managing Executive, Strategic Partnerships at IBM, adds: *"We are looking at different deployment models, even switching programming languages, which can reduce the energy consumption of an application by up to 50%."* A recent academic research paper proposed linear-complexity multiplication algorithm that approximates floating point multiplications with integer additions, significantly reducing energy consumption by up to 95% for element-wise tensor multiplications and 80% for dot products.⁵¹ More than half (53%) of organizations in our research say that model switching (embedding several Gen AI models in a single overarching model, where depending on the task, system itself decides which model to use) can help to reduce the environmental footprint of Gen AI.

Also, adding checkpoints to Gen AI code helps minimize the energy consumption and computation resources by allowing training jobs to resume from where they left off, rather than starting from scratch.



"We have designed new algorithms that reduce time and memory complexity of foundation models, resulting in considerable energy and carbon-footprint savings."

Ramin Hasani

Co-Founder and CEO of Liquid AI



- Implement **prompt and query optimization** techniques to cut down the computational resources required for each task, in turn reducing energy use and operational costs. Using techniques such as prompt caching⁵² (a technique that allows developers to temporarily store frequently used context between API [application programming interface] calls to LLMs), and concise chain of thought⁵³ (CCoT, a prompt engineering technique that instructs the LLM to think step by step, while encouraging brevity) can help to reduce energy consumption.
- Additionally, consider utilizing **edge computing** devices to reduce energy usage associated with data transfer and distribution.
- Select cloud providers with **energy-efficient data centers**. A study by Google found that using a more efficient AI model architecture, processor, and a greener data center can reduce the tech's carbon footprint by between 100 and 1,000 times.⁵⁴ Martin Elwin, Senior Engineering Director at Swedish fintech company Klarna, says: *“One of the things we recommend is to use AI, especially inference, through providers that have shared on-demand AI inference environments. The more people using a public cloud service, the higher the utilization rates. The improvement in the use of resources in running power-hungry AI applications could make a difference to your organization's overall carbon footprint.”*⁵⁵
- **Choose a region for your server that ensures smaller environmental impact.** The environmental consequences and the GHG footprint of Gen AI projects can vary significantly due to the varying energy mix of the locations of the servers and data centers hosting AI applications. Jesse Dodge, Senior Research Scientist at the Allen Institute for AI, says: *“The carbon footprint of AI in places where the power grid is relatively clean, such as France, will be much lower than it is in places with a grid that is heavily reliant on fossil fuels, such as some parts of the US. While the electricity consumed by running AI models is fixed, we might be able to reduce the overall carbon footprint of these models by running them in areas where a larger proportion of the power grid is made up of renewable sources.”*⁵⁶ Scheduling AI/Gen AI workloads in regions and times where renewable energy sources are more available by leveraging real-time data on carbon intensity (the amount of CO2 emitted per unit of electricity consumed) can help to reduce the environmental impact of Gen AI.
- Finally, continuously **monitoring and reporting** the usage of Gen AI systems to identify and reduce unnecessary or excessive use is critical; 64% of executives across functions agree.





“It is urgent for businesses to embed sustainability into their AI strategies. By leveraging smaller models, renewable energy, and transparent practices from AI and GenAI vendors, we can mitigate environmental impacts while harnessing AI to drive both innovation and sustainability.”

Vincent Charpiot

Head of Group Sustainability
Business Accelerator at Capgemini

Harness Gen AI's potential to accelerate sustainable business by investing and scaling the right use cases

Responsible Gen AI has the potential to accelerate sustainable business value. When it comes to using Gen AI to accelerate progress toward business and sustainability goals, the first step is to identify suitable use cases. Only four out of ten (39%) executives who participated in our survey have a clear understanding of where Gen AI can be used for accelerating the sustainability transformation. As mentioned previously, taking a holistic approach by combining the benefits of AI and Gen AI can help organizations to drive growth, efficiency and sustainable business value. Organizations should:

- Identify the sustainability use cases where Gen AI can help garner **significant benefits** in terms of reduction of GHG emissions, energy efficiency, water conservation, waste reduction, regulatory compliance, etc. Also, these

use cases should be **easily scalable** across the entire organization.

- Evaluate the use cases on associated business benefits (financial, environmental, operational) versus the associated cost (financial, environmental, expected disruption, time to implement) to **shortlist use cases that can bring maximum value with minimal associated cost**. A senior executive for a German energy company elaborates: *“Challenge number one is finding a business case for your use case. For Gen AI in sustainability, it might be more difficult to find the right business case. How does the investment pay off? What is my actual return on investment? What is reducing my current footprint by X% worth?”*
- Substitute alternative technologies for Gen AI only for those use cases where **Gen AI brings significant additional value**.
- Define the expected value or sustainability outcomes, as **quantified by KPI metrics**, to be measured and tracked after deploying Gen AI for each use case.

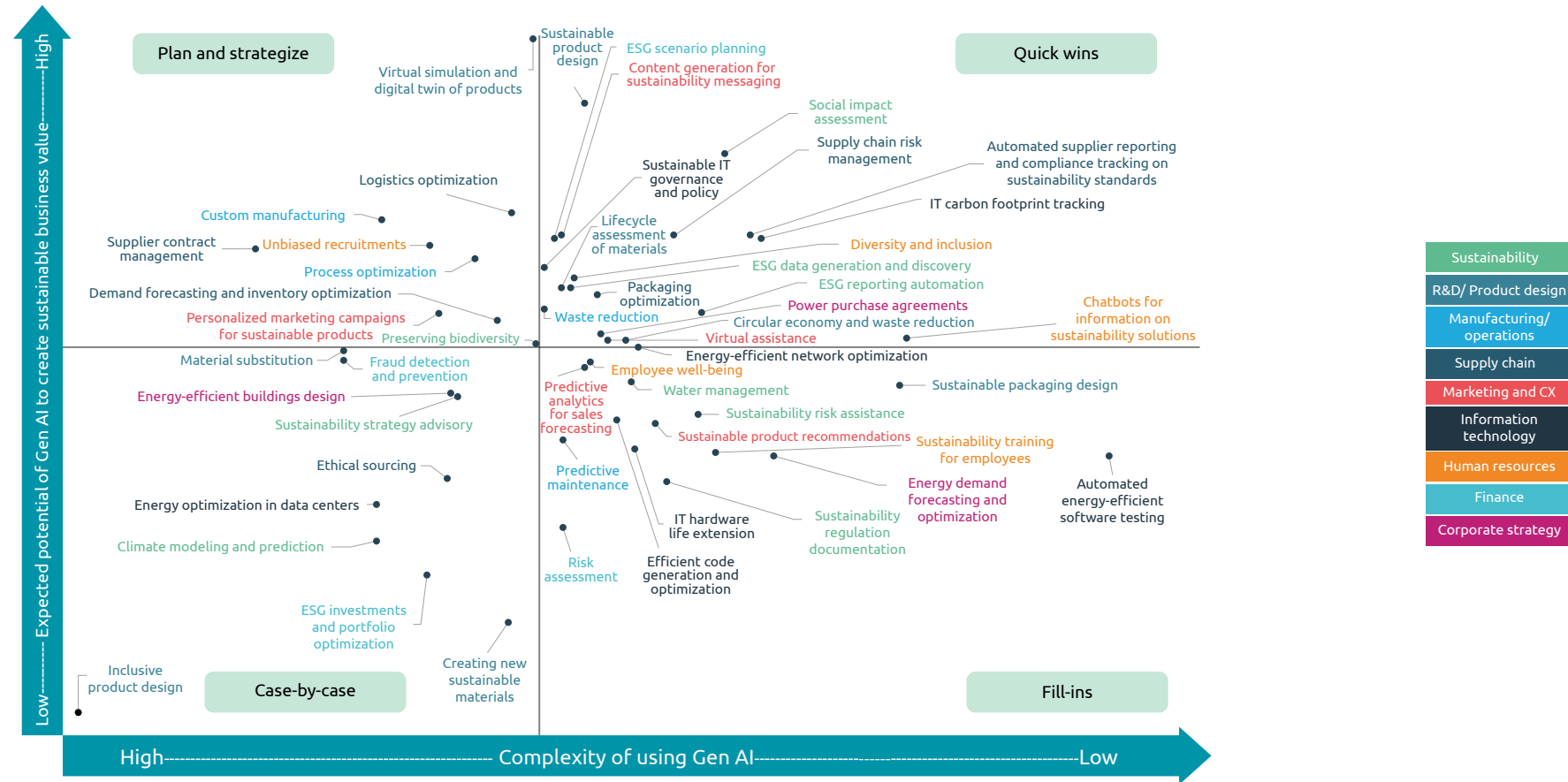
In our research, we looked at more than 100 use cases across functions and sectors and assessed them across two dimensions: complexity of implementation and the potential to create a sustainable business impact, as perceived by the survey respondents. Each use case falls into one of four categories:

- **Quick wins:** These offer the dual advantage of high expected sustainability-related benefit and low complexity of Gen AI implementation
- **Plan and strategize:** These use cases offer high benefits, but are accompanied by high complexity of implementation, so organizations will need to position themselves carefully for effective use.
- **Fill-ins** can be classified as “can do,” as they offer low complexity and low benefit
- **Case-by-case** offer high complexity and low benefit

It should be noted that in many cases organizations plan to deploy a combination of technologies to achieve a more comprehensive approach to problem solving and innovation.

Figure 13.

Across functions, a few quick-win use cases emerge



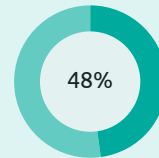
Note: The list of use cases is indicative and not comprehensive

Source: Capgemini Research Institute, Gen AI and Sustainability survey, August-September 2024, N = 2,000 executives from organizations that are working on Gen AI initiatives

A deep dive into “quick wins” by function: Those use cases where AI/Gen AI has high potential to create sustainable business value and has lower complexity of implementation

Sustainability

- ESG reporting automation
- ESG data generation and discovery
- Social impact assessment



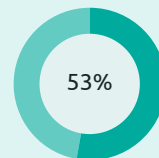
of sustainability executives say they are already piloting use of Gen AI to automate ESG reporting.

US-based battery manufacturing company EnerSys is using Gen AI to analyze and **generate insights from large datasets related to sustainability metrics**, including Scope 1 and 2 emissions, travel data, and waste data.⁵⁷

Capgemini's Sustainability Data Hub provides a robust foundation for ESG intelligence.⁵⁸ The **ESG Lens** solution provides organizations with an AI-powered solution incorporating NLP, LLMs, and Gen AI that delivers ESG insights from unstructured data with complete transparency.⁵⁹

R&D/ Product development

- Sustainable product design
- Lifecycle assessment of materials
- Circular economy and waste reduction



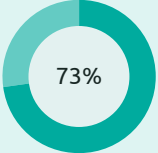
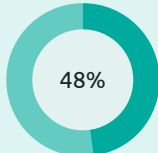
of executives from the R&D function rank using Gen AI for **sustainable product design** among the top three priorities in the next 12 months.

Automotive player Toyota is investing heavily in Gen AI techniques. It has developed a text-to-image tool that applies engineering constraints to early design sketches. It also accepts parameters such as aerodynamic drag as input to assist designers in **creating fuel-efficient vehicles** with fewer iterative cycles.⁶⁰

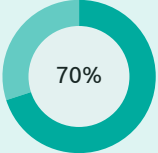
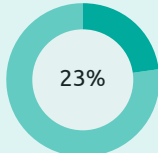
Bio-tech organizations including Cradle, Basecamp Research, Arzeda, Biomatter Designs, Cambrium, and Absci are **designing new proteins from scratch** using Gen AI.⁶¹

Siemens' Gen-AI powered HEEDS platform aids organizations to accelerate **sustainable product design** while reducing cost.⁶²

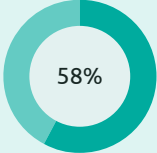
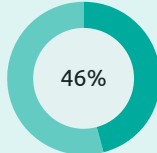
"continued to the next page"

<p>Manufacturing/ operations</p>	<ul style="list-style-type: none"> Waste reduction 	 <p>73%</p> <p>of executives from operations rank using Gen AI for waste reduction in manufacturing/operations among the top three priorities in the next 12 months.</p>	<p>Walmart is piloting a Gen AI-powered waste-management solution that helps predict and prevent wastage, recommending price adjustments or donation of items nearing expiration.⁶³</p>
<p>Supply Chain</p>	<ul style="list-style-type: none"> Automated supplier reporting and compliance tracking on sustainability standards Supply chain risk management Packaging optimization 	 <p>48%</p> <p>of supply chain executives say they are already piloting use of Gen AI for automated supplier reporting and compliance tracking on sustainability standards.</p>	<p>British-Dutch multinational CPG company Reckitt is using AI and Gen AI to slash its Scope 3 emissions, which represent 96% of its total emissions. Partnering with CO2 AI and Quantis, Reckitt gathered over 300,000 data points for emissions per product, and used AI and Gen AI to transform these into a faster and more accurate carbon footprint. Gen AI is helping the company to reveal new ways to reduce its emissions by 2030.⁶⁴</p>

"continued to the next page"

<p>Marketing and CX</p>	<ul style="list-style-type: none"> • Content generation for sustainability messaging • Virtual assistance 	 <p>70%</p> <p>of marketing/sales/CX executives rank using Gen AI for sustainability messaging among the top three priorities in the next 12 months.</p>	<p>French grocery firm Carrefour has deployed Hopla, a chatbot based on ChatGPT to assist customers with their daily shopping. Hopla also provides anti-waste solutions to customers, suggesting practical ways of reusing ingredients and crafting associated recipes.⁶⁵</p>
<p>Information technology</p>	<ul style="list-style-type: none"> • Energy-efficient network optimization • IT carbon footprint tracking and reporting • Sustainable IT governance and policy 	 <p>23%</p> <p>of technology executives say they are already piloting use of Gen AI for sustainable IT governance and policy.</p>	<p>Gen AI can help to translate code into more energy-efficient languages. Nicole Onuta, Lead AI risk management at a large financial services company in Netherlands says, <i>"I have seen a huge benefit in transitioning the code from a legacy programming language to a new one using generative AI."</i>⁶⁶</p>

"continued to the next page"

<p>Human resources</p>	<ul style="list-style-type: none"> Promoting diversity and inclusion Chatbots for information on sustainability solutions 	 <p>58%</p> <p>of HR executives rank the use of Gen AI for diversity and inclusion among their top three priorities in the next 12 months.</p>	<p>Canada-based startup Diversio offers a Gen AI-led solution to identify patterns in HR data and suggest actionable changes to help businesses evaluate and improve their inclusivity efforts.⁶⁷</p>
<p>Finance and corporate Strategy</p>	<ul style="list-style-type: none"> ESG scenario planning Power purchase 	 <p>46%</p> <p>of finance executives rank use of Gen AI for ESG scenario planning among their top three priorities in the next 12 months.</p>	<p>FactSet, a US-based company that provides software and financial data solutions to investment professionals, offers a Gen AI-powered platform that uses ML algorithms to assess financials, market trends, and ESG metrics, offering insights into how ESG factors could affect companies' financial performance in various scenarios.⁶⁸</p>

To complement the functional use cases, we also analyzed some key sector-specific use cases where AI/Gen AI can create a sustainable business value. Based on analysis of our survey results, as well as our internal expertise, we highlight the following use cases:

<h3>Aerospace and defense</h3>	<ul style="list-style-type: none"> • Optimization of flight paths to reduce fuel consumption • Design lightweight materials to contribute to fuel efficiency 	<p>In 2023, NASA's Goddard Space Flight Center used Gen AI to design specialized parts. These parts save up to two-thirds of the weight compared to traditional components.⁶⁹</p> <p>Emirates Airlines is using Gen AI for efficient flight-route planning.⁷⁰</p>
<h3>Automotive</h3>	<ul style="list-style-type: none"> • Energy efficient vehicle design • Material optimization • Virtual vehicle testing and simulation • Sustainable manufacturing process • Sustainable battery ecosystem • Sustainable supply chain 	<p>General Motors utilizes generative design to create lighter vehicle parts.⁷¹</p> <p>US-based startup Aionics is using Gen AI models trained on existing battery materials data to create or design new molecules targeted at a certain application. The startup is already working with several companies, including Porsche's battery manufacturing subsidiary Cellforce.⁷²</p>
<h3>Consumer products and retail</h3>	<ul style="list-style-type: none"> • Waste reduction • Sustainable product design • Material substitution • Sustainable supply chain 	<p>Unilever uses Gen AI to optimize product formulations, reducing the use of harmful chemicals and minimizing waste.⁷³</p> <p>Adidas has utilized Gen AI to design sustainable products. The algorithm analyzes data on materials and consumer preferences to generate designs that meet sustainability standards, while also appealing to customers.⁷⁴</p> <p>US-based Citrine Informatics leverages Gen AI-based platform to help organizations to accelerate sustainable materials and chemicals development – enhancing product performance and reducing research time.⁷⁵</p>

"continued to the next page"

Energy and utilities

- Renewable energy optimization
- Climate tech development and adoption
- Dynamic grid management
- Optimizing energy storage
- Sustainable battery ecosystem

Sergio Merchán, Global CIO at Iberdrola, states: *“We see [Gen] AI playing a central role in delivering the smart grid, renewables, and storage technology that will drive the energy transition.”* Iberdrola plans to harness Gen AI to **optimize renewable energy production** by integrating data from its 400 renewable energy sites.⁷⁶

Danish energy company Ørsted has implemented a Gen AI system that analyzes data from wind turbines and **predicts potential failures**. This not only helps reduce maintenance costs but also ensures more efficient use of resources.⁷⁷

More than eight in ten (82%) executives in our previous research say that Gen AI will have a high impact on accelerating the **adoption of climate tech**.⁷⁸ For instance, Microsoft has trained a Gen AI model with US nuclear regulations and licensing documents to assist with generating paperwork for SMR projects.⁷⁹

Capgemini leveraged a spatial network graph and neural network to improve the quality of asset data for Western Power Distribution (WPD) – UK’s electricity distribution network operator (DNO). This model predicts and corrects asset attributes like voltage, conductor size, and material, **ensuring accurate network data** and aiding in energy transition.⁸⁰

Financial services

- Identification of sustainable investment opportunities and assessment of ESG risks
- Fraud detection
- Automated claims assessment

JPMorgan Chase employs AI/Gen AI to **analyze ESG data sources, identifying pertinent ESG themes** for specific sectors, stocks, and management practices, enabling clients to screen for and compare organizations. This expedites pattern analysis and report generation, highlighting viable investment opportunities and showcasing how technology enhances strategic capabilities for long-term sustainability.⁸¹

"continued to the next page"

<h2>Life sciences</h2>	<ul style="list-style-type: none"> • Drug discovery with Gen AI • Design of more efficient clinical trials for reduced environmental footprint • Development of environmentally friendly chemical processes for drug synthesis and manufacturing 	<p>Merck has launched a Gen AI-based solution to integrate drug discovery and synthesis. It virtually screens compounds from over 60 billion chemical targets and proposes the optimum method for synthesizing these drugs. Karen Madden, CTO, Life Science business sector at Merck, comments: <i>“Our platform enables any laboratory to count on generative AI to identify the most suitable drug-like candidates in a vast chemical space. This helps ensure the optimal chemical synthesis route for development of a target molecule in the most sustainable way possible.”</i>⁸²</p>
<h2>Public sector</h2>	<ul style="list-style-type: none"> • Public health monitoring • Disaster response planning • Urban planning • Climate risk mitigation and adaptation 	<p>Boston in the north-eastern US is using Gen AI to increase the bike-friendliness of its city infrastructure. Jinhua Zhao, Professor of Cities and Transportation at MIT, says: <i>“With the help of generative AI, we can overlay a Copenhagen-style bicycle infrastructure on top of the existing Boston infrastructure and building environment, so that Bostonians can get a real sense of what a bicycle [network] for our city would look like.”</i>⁸³</p>
<h2>Telecom</h2>	<ul style="list-style-type: none"> • Optimize network traffic and reduce energy consumption 	<p>An executive from a large European telecom comments: <i>“Currently, we use generative AI to optimize network capacity and reduce energy consumption. Moving forward, we aim to use it to address Scope 3 emissions and enhance ESG reporting and calculation processes.”</i>⁸⁴</p>

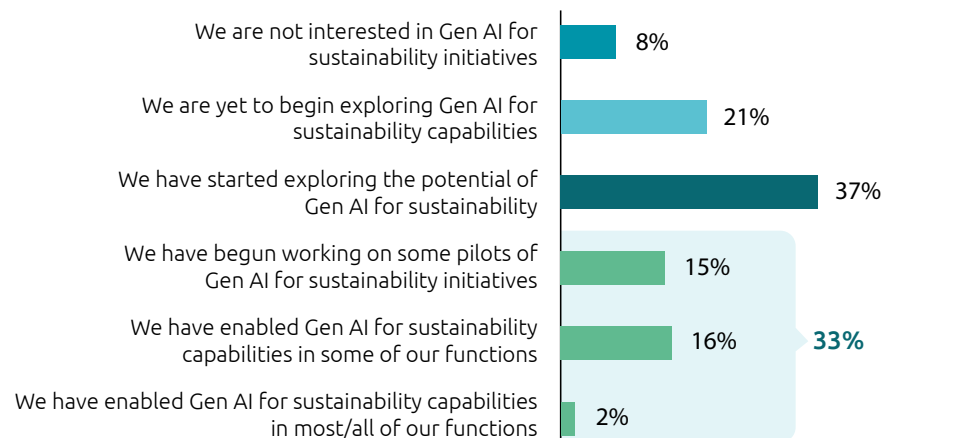


In fact, **one-third (33%) of executives mentioned that they have already started using Gen AI for sustainability initiatives**, and a further 37% have started to explore its potential for sustainability (see figure 14). Moreover, a majority (84%) also expect an average increase of 16% in Gen AI investments for accelerating sustainability in the next 3–5 years.

Figure 14.

One in three organizations is already using Gen AI for sustainability initiatives

Percentage of executives that agree with the statements below:



Source: Capgemini Research Institute, Gen AI and Sustainability survey, August-September 2024, N = 2,000 executives from organizations that are working on Gen AI initiatives.

Although most organizations still fall short on measuring the environmental footprint of their Gen AI use, they believe that Gen AI-driven use cases have the potential to reduce GHG emissions. As figure 15 shows, 66% of executives say they expect a reduction of more than 10% in GHG emissions in the next 3–5 years because of deploying Gen AI for sustainability initiatives. Canadian startup BrainBox AI uses Gen AI to reduce the carbon footprint of commercial buildings. The company uses historical data to predict interior building temperatures, cutting heat, ventilation, and air-conditioning (HVAC) costs by up to 25% and GHG emissions by 40%.⁸⁵

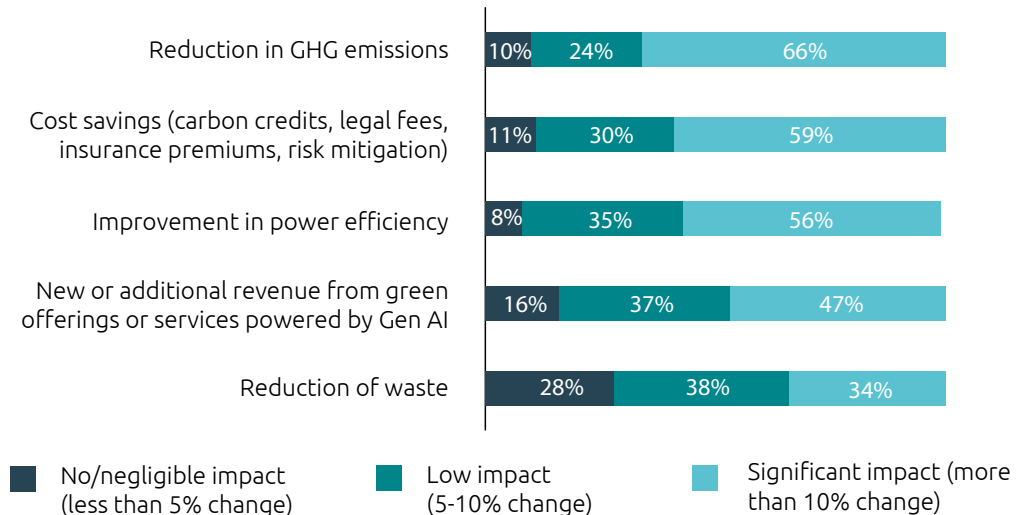
66%

of executives say they expect a reduction of more than 10% in GHG emissions in the next 3–5 years because of deploying Gen AI for sustainability initiatives.

Figure 15.

Organizations expect benefits of harnessing Gen AI for sustainability initiatives

Percentage of executives saying they expect to achieve the below benefits by using Gen AI for sustainability initiatives in the next 3-5 years



Source: Capgemini Research Institute, Gen AI and Sustainability survey, August-September 2024, N = 2,000 executives from organizations that are working on Gen AI initiatives.

The rise of AI agents for creating a sustainable business value

Unlike traditional AI systems, AI 'agents' are adept at decision-making, interpretation of instructions and planning, and are capable of adaptation and acting independently. Our previous research reveals that, while only 10% of organizations currently employ AI agents, a large majority (82%) intend to integrate them within 1–3 years.⁸⁶

Given their ability to streamline and automate complex tasks or processes, AI agents can play a pivotal role in a sustainability drive. Below, we highlight a few examples of their rising trajectory:

- Germany-based Compliance Solutions has launched an ESG AI-Agent that fully automates research, evaluation, and compliance violation reports on ESG topics.⁸⁷

- Australia-based startup, ESG&I, has launched a suite of AI agent, ChatESG™, to help with ESG consultation.⁸⁸
- Startup GreenMetrica has also launched AI agents to help organizations with compliance and reporting, such as CSRD, optimize operations and supply chains to reduce carbon impact, and enhance environmental data quality.⁸⁹

Moreover, agentic AI allows problems to be broken down into multiple steps and sent to the optimal Gen AI model. Prompts that do not require a general purpose LLM go to smaller models for quicker processing and lower energy usage.⁹⁰



“Generative AI will be key to unlocking a more sustainable future and accelerating the attainment of net zero ambitions.”

Christopher Scheefer
Global Sustainability Lead for
Data and AI at Capgemini

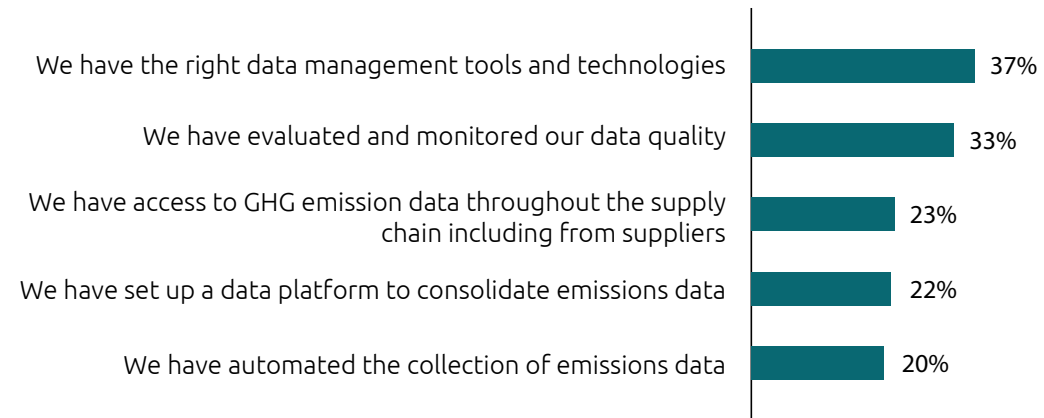
Develop the right data and technology foundations

Data is the backbone of Gen AI. To harness the potential of Gen AI, data foundations require greater focus. Our previous research highlights that nearly three out of five data executives agree that integrating Gen AI will require significant changes to their data collection, storage, retrieval, and governance.⁹¹ Only 37% of the executives claim their organization has the right data management tools and technologies for Gen AI, and only one-third (33%) evaluate and monitor data quality. Only two out of ten executives agree that their organization has automated the collection of emissions data (see figure 16).

Figure 16.

Organizations lag on various dimensions of data maturity to embrace Gen AI for sustainability

*Percentage of executives that agree with the following statements:
"To utilize Gen AI for accelerating our sustainability transformation ..."*



Source: Capgemini Research Institute, Gen AI and Sustainability survey, August-September 2024, N = 2,000 executives from organizations that are working on Gen AI initiatives.

Organizations should:

Define the data strategy for a holistic approach towards AI/Gen AI: Key components of a data strategy include:

- A vision of how data will be used to support sustainability ambitions
- An assessment of gaps in the availability and quality of data
- A view of internal and external data sources with adequate levels of coverage, granularity, and accuracy
- Technology architecture and solutions required to collect, process, and compute data
- Governance mechanisms to manage issues such as data auditability and traceability

Have a sound process for obtaining trustworthy data: For Gen AI models to be effective in collating information to derive meaningful insights, ensuring the reliability and correctness of input and output data is critical. Jaydeep Ghosh, a data management and analytics professional in the healthcare sector, adds: *“Organizations want to merge their internal data with the power of generative AI to create more specific insights. We must define the good data sources.”*⁹²

Set standards and prequalify sources to ensure the data your organization feeds into the AI/Gen AI models follows specific protocols. Erik Brynjolfsson, Professor at the Stanford Institute for Human-Centered AI and Director of the Stanford Digital Economy Lab, talks about how data quality can become a competitive advantage: *“Not all companies have done a good job of capturing and curating data, but the ones who have will find that they are sitting on a goldmine. Companies with better data are going to be the biggest winners in this paradigm.”*⁹³

Develop data platform capabilities to industrialize the ingestion, storage, and processing of data: To enhance data coverage, granularity and accuracy, organizations should automate the ingestion of data from multiple external and internal sources. They will then need to feed ingested data into a single data platform to provide a ‘single source of truth.’

Ensure democratization of data through data mesh: Democratization of data is essential to disseminate data to business users at speed, in the right form and degree of granularity. This enables business users to explore data and derive actionable insights without waiting for the technical teams. Data mesh enables this by making data products and their metadata available to different business units in a single platform, as well as domain-oriented decentralized data ownership, self-serve data infrastructure as a platform, and federated computational governance.

“continued to the next page”

Implement strong data governance with clear accountability for data standards, data quality, audits, data collection touchpoints, processes, and schedules. Given the interdependencies between AI, data, and technology, collaboration between the chief AI officers (CAIO), chief information officers (CIO), and chief data officers (CDO) is essential.

Strive to ensure sustainable data operations: Organizations should also consider that collecting, storing, accessing, and utilizing data comes with its own sustainability price, too. It involves being more selective about what data is necessary, addressing the issue of data waste, and understanding that 'big' data is not always synonymous with 'better' data.⁹⁴

Invest in the right talent pool: As well as investing in infrastructure, organizations must look in the right places for the talent to design, develop, and run these systems. In our survey, only 24% of executives say they have the skills to implement sustainable AI/Gen AI.



Set governance for a sustainable Gen AI

Organizations need to implement **governance models** to ensure safe, transparent, ethical, and sustainable usage and design of Gen AI. A majority (62%) of executives believe that setting the right guardrails and governance for sustainable development/deployment of Gen AI would reduce the environmental impact of Gen AI. Nearly half (49%) also rank lack of clear governance models among the top five challenges of implementing Gen AI for sustainability.

To adopt a responsible approach, it is key to **set up a governance body** to implement accountability measures. Our research suggests that a centralized governance body can take decisions faster, including resource allocation based on requirements. It will have a clear mandate to reduce the Gen AI footprint and will have responsibility for drafting a clear action plan that can be enforced collaboratively with all stakeholders. At the same time, this will ensure efficient Gen AI implementation for the most promising sustainability use cases.

Effective governance structures are multidisciplinary, as well as close collaboration between the teams managing sustainability (such as sustainability, human rights, ethics, legal compliance), teams that develop or purchase Gen AI technologies (such as engineering, product, IT, research and development, procurement), and teams that use Gen AI technologies (such as operations, supply chain, human resources) is required. Our research shows that 57% of executives rank “lack of collaboration between Gen AI implementation teams and sustainability teams” as a key challenge today.

Also, it is important for organizations to **create a framework to approve or reject Gen AI/AI pilots**, based on parameters such as cost, feasibility, timelines, environmental footprint, and sustainability benefits.

As with any initiative, to manage and mitigate Gen AI footprint and harness Gen AI for sustainability requires all stakeholders to have access to the relevant information and work cohesively towards a common goal. Karthik S, General Manager, AI at LogicMonitor, says: **“Co-operating with governments ensures alignment with regulatory standards and policy development, while partnerships with universities drive research and innovation in**

sustainable AI technologies.”⁹⁵ Stakeholders across the Gen AI ecosystem – organizations, Gen AI model developers, hyperscalers, technology integrators, startups, research institutions, sustainability experts, and governments – must collaborate to share best practices, develop new solutions and sustainable AI standards, and promote collective action.

62%

of executives believe that setting the right guardrails and governance for sustainable development/deployment of Gen AI would reduce the environmental impact of Gen AI.

Conclusion

Gen AI has significant environmental costs associated with its development and usage. Organizations should not ignore the high energy consumption, considerable carbon footprint, increased water usage, and mounting e-waste across its lifecycle. However, Gen AI can also play a pivotal role in accelerating sustainable business transformation. As we stand on the brink of an AI-driven future, it is crucial for organizations to understand how to make informed, responsible choices regarding the opportunities that Gen AI offers while managing its environmental footprint.

It is imperative that organizations build sustainability considerations into Gen AI business cases. By choosing the appropriate technology; optimizing model sizes, data and algorithms; using energy-efficient and alternative hardware; using green data centers powered by low-carbon energy sources; and incorporating sustainability measures across the Gen AI lifecycle, organizations can shrink their environmental footprints as well as costs.

At the same time, it is important for organizations to identify and prioritize the right sustainability use cases to tap into Gen AI, based on financial and environmental costs, benefits, associated risks, and alignment with the organization's sustainability goals. Organizations must build the required data and technology foundations and develop the skillsets to derive maximum benefit from Gen AI in sustainability.

Building an effective multidisciplinary governance structure and collaborating with various players within the ecosystem can go a long way toward implementing Gen AI effectively within the guardrails of sustainability.

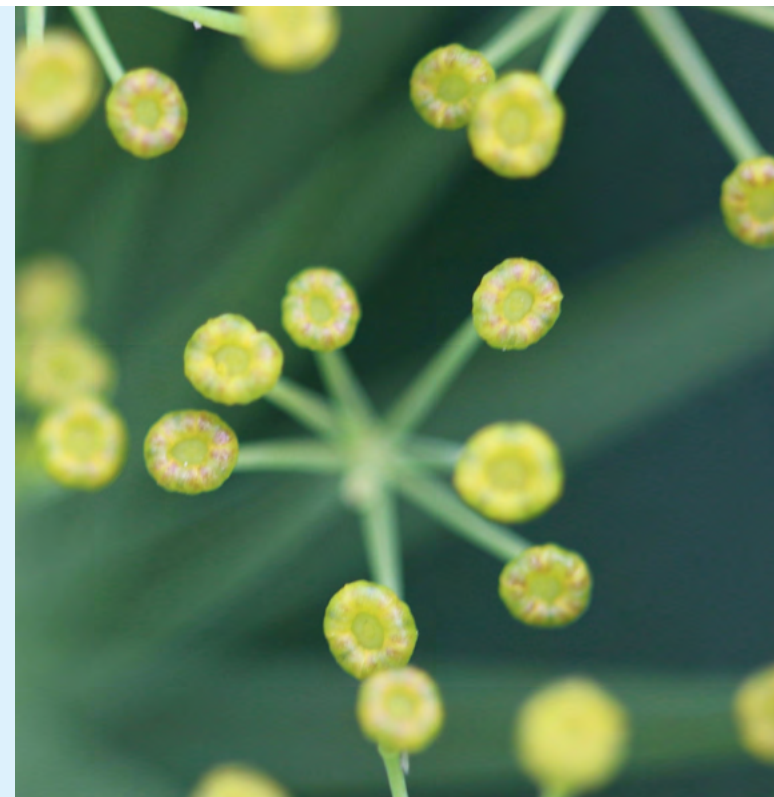
The journey towards sustainable Gen AI is a collective responsibility that requires commitment, innovation, and a willingness to prioritize the conservation of our environment and well-being alongside technological advancement and commercial gain. Gen AI's potential to progress organizational sustainability and UN SDGs shows that within the technology itself lies the potential to offset those dangers.

Appendix

Definitions

For the purposes of this research, we use the following definitions:

- **Artificial intelligence (AI):** Development of computer systems capable of performing tasks that historically required human intelligence, such as recognizing speech, making decisions, and identifying patterns. AI is an umbrella term that encompasses a wide variety of technologies, including machine learning, deep learning, and natural language processing (NLP).⁹⁶
- **Machine learning (ML):** This is a subfield of AI that uses algorithms trained on data sets to create self-learning models that are capable of predicting outcomes and classifying information without human intervention. It is used today for a wide range of commercial purposes, including suggesting products to consumers based on their past purchases, predicting stock market fluctuations, translating text from one language to another, and much more.⁹⁷
- **Generative AI (Gen AI):** It is a type of AI that has the capability to learn and reapply the properties and patterns of data for a wide range of applications, from creating text, images, and videos in different styles to generating tailored content. It enables machines to perform creative tasks previously thought exclusive to humans.⁹⁸



How Gen AI can support the UN's Sustainable Development Goals

Gen AI, if developed and used responsibly, can accelerate business impact on sustainable development. Gen AI can track progress on the UN's Sustainable Development Goals (SDGs) and identify patterns that can inform evidence based decision-making and targeted actions to sustain and accelerate that progress.⁹⁹ In the figure below, we delve into case examples of the transformative role of Gen AI in some SDGs.



Precision agriculture to increase crop yield

Syngenta Group, a Switzerland-based agri-tech organization, launched Cropwise AI, a multi-lingual Gen AI-powered conversational LLM that provides farmers with tailored recommendations for precise input applications, product placement, and strategies for pest and disease control, with the potential to increase yield by an estimated 5%.¹⁰⁰



Accelerating drug discovery

Insilico Medicine, a Hong Kong and New York-headquartered biotech company, is using Gen AI to develop a drug for idiopathic pulmonary fibrosis, a rare disease that causes progressive lung decline. Historically, such a drug-discovery process would have taken an estimated six years and cost over \$400 million. Using Gen AI, Insilico reduced the cost to around \$40 million over two and a half years.¹⁰¹

Improving healthcare

Using Google Cloud's latest Vertex AI and Gen AI technologies, Apollo Hospitals has developed a new clinical intelligence engine (CIE) to provide secure and reliable healthcare services across India. Utilizing Google Cloud's LLMs and the vast dataset collected by Apollo Hospitals, CIE offers proprietary solutions to assist doctors in quickly determining the best course of action for patients. Additionally, Apollo Hospitals intends to utilize the more advanced Med-Pathways Language Model (MedPaLM) 2, a Google-developed LLM that can answer medical enquiries and generate summaries of clinical text.¹⁰²

"continued to the next page"



Gen AI for education

India's Khan Academy has developed Khanmigo, a Gen AI tutor assisting young students in various subjects.¹⁰³

Dastudy, a Cameroon-based Gen AI-driven startup, offers a platform that supports users in learning new skills, allows them to share content, find up-to-date information on the Cameroonian socio-cultural environment, and get help with everyday administrative tasks, including writing and translation.¹⁰⁴

Capgemini is partnering with UNESCO and AWS for the seventh edition of Global Data Science Challenge. The challenge focuses on developing agentic AI systems that empower educators through powerful data tools.¹⁰⁵



Gen AI for climate action

Andrew Ng, Founder of LandingAI and Managing General Partner, AI Fund, says: *“One area we must study seriously is sunlight reduction, which is also called climate geoengineering via solar-radiation management. High-altitude stratospheric aerosol injection could effectively put a parasol around the planet to reflect sunlight and cool us down. We are now able to train very large foundation models to predict more accurately the effects of stratospheric aerosol injection. Given the world's collective inability to reduce CO2 emissions climate engineering [could be our best option]. I think AI, especially large AI foundation models of climate, have a large role to play in that.”*¹⁰⁶

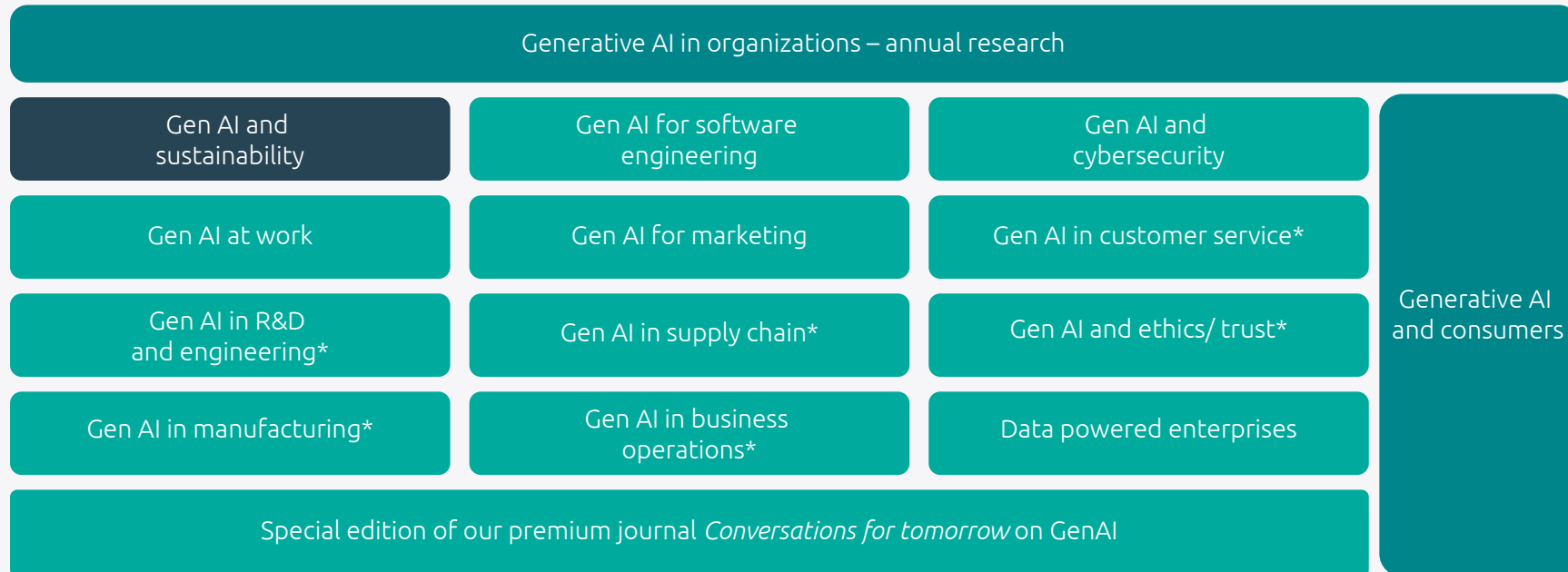
Nvidia utilizes Gen AI in its NVIDIA Earth-2 platform to create high-resolution weather forecasts and simulate climate scenarios, helping organizations anticipate extreme weather events and take proactive measures.¹⁰⁷



Protecting biodiversity

Researchers from the University of Valencia, the Instituto de Física Corpuscular in Spain, and the University of Sussex experimented with Gen AI to analyze the coexistence patterns of different species in vegetation patches and the implications for biodiversity. The researchers trained two Gen AI systems to simulate likely compositions of patches and evaluate them against increasingly complex ecological metrics.¹⁰⁸

This report is a part of Capgemini Research Institute's series on Generative AI



*Upcoming reports

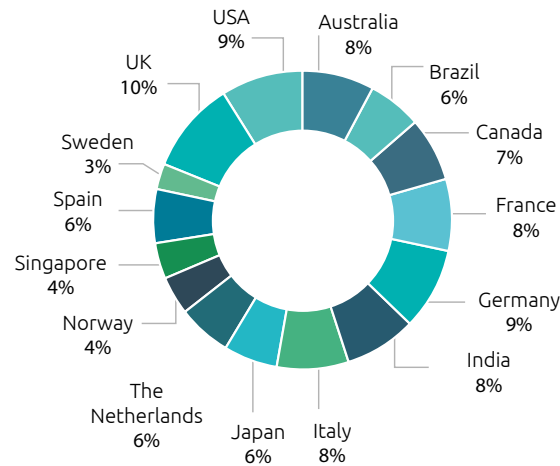
To find out more, please go to <https://www.capgemini.com/insights/research-institute/>

Research methodology

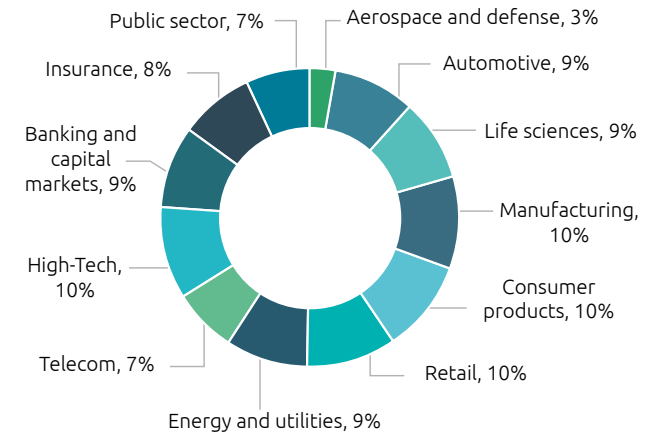
We surveyed 2,000 executives from organizations with more than \$1 billion in annual revenue, across 15 countries in North America, Europe, and Asia-Pacific and in 12 sectors, and that were reasonably/well/highly informed about their organizations' strategies and initiatives around Gen AI, as well as environmental and social sustainability initiatives. We selected the executives from organizations already working with Gen AI, either at the pilot stage or across a few/most functions/locations. The executives we surveyed were at director level and above, with one-third from technology functions, one-third from corporate innovation/strategy and sustainability functions, and the remaining third from business functions such as finance, legal, HR, marketing and communications, sales, R&D, procurement, manufacturing, operations, and logistics. The global survey took place in August 2024. In the following charts, we provide the distribution of executives and their organizations. Lastly, we conducted interviews with 15 executives and academics at leading global organizations.

Findings from the study reflect the views of the respondents to our online questionnaire for this research and are aimed at providing directional guidance. Please contact one of the Capgemini experts listed at the end of the report to discuss specific implications.

Percentage of executives by location of headquarters of current organization

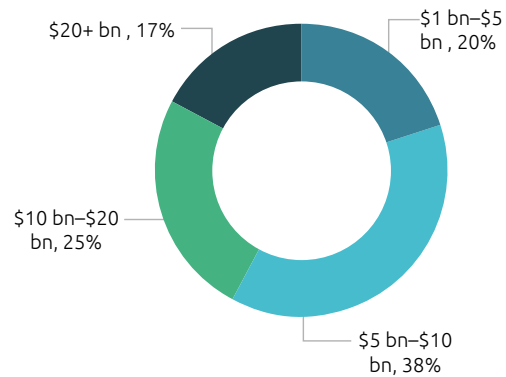


Percentage of executives by sector of current organization

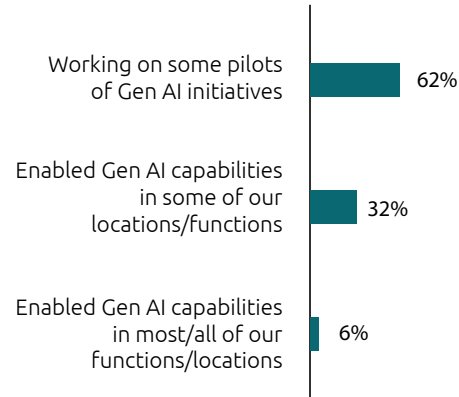


Source: Capgemini Research Institute, Gen AI and Sustainability survey, August-September 2024, N = 2,000 executives from organizations that are working on Gen AI initiatives.

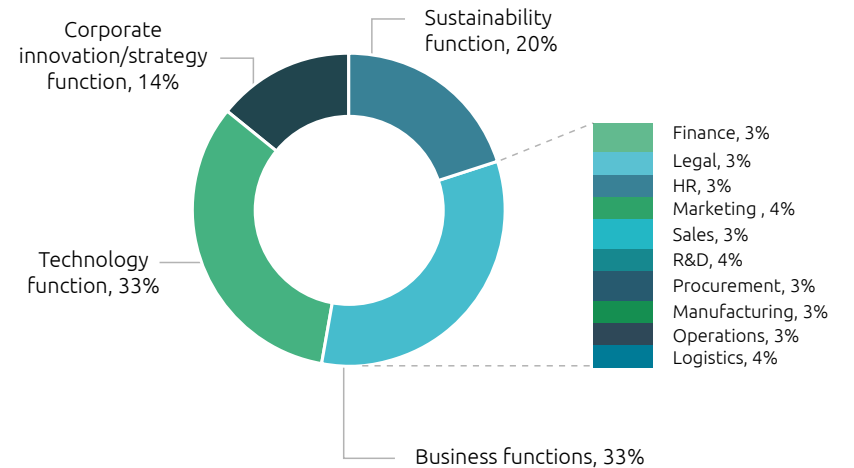
Percentage of executives by organization enterprise-level revenue, in USD



Percentage of executives by adoption stage of Gen AI within organization



Percentage of executives by business function



Source: Capgemini Research Institute, Gen AI and Sustainability survey, August-September 2024, N = 2,000 executives from organizations that are working on Gen AI initiatives.

References

1. Frontline Magazine, "E-waste from AI computers could 'escalate beyond control': study," October 2024.
2. The Verge, "How much electricity does AI consume?" February 2024.
3. Towards Data Science, "Environmental implications of the AI boom," May 2024.
4. Capgemini Research Institute, What matters to today's consumer, January 2025.
5. Capgemini Research Institute, Harnessing the value of generative AI - Vol II, July 2024.
6. Harvard Business Review, "How to make generative AI greener," July 2023.
7. Nvidia, "Sustainable strides: How AI and accelerated computing are driving energy efficiency," July 2024.
8. Design Life-Cycle, "Nvidia FE GPU," accessed in October 2024.
9. TechTarget, "Generative AI's sustainability problems explained," October 2023.
10. bid.
11. CIO.com, "3 things CIOs can do to make gen AI synch with sustainability," May 2024.
12. The Verge, "How much electricity does AI consume?" February 2024.
13. Towards Data Science, "Environmental implications of the AI boom," May 2024.
14. David Patterson, Joseph Gonzalez, Urs Hölzle, Quoc Le, Chen Liang, Lluís-Miquel Munguia, Daniel Rothchild, David So, Maud Texier, and Jeff Dean. "The carbon footprint of machine learning training will plateau, then shrink," April 2022. <https://doi.org/10.48550/ARXIV.2204.05149>.
15. IEA, Electricity 2024: Analysis and forecast to 2026, January 2024.
16. Vox, "AI already uses as much energy as a small country. It's only the beginning," March 2024.
17. IEA, Electricity 2024.
18. Goldman Sachs, "AI/data centers' global power surge and the sustainability impact," April 2024.
19. Microsoft, How can we advance sustainability? 2024 Environmental Sustainability Report, 2024.
20. Google, Environmental Report 2024, July 2024.
21. CNBC, "Amazon goes nuclear, to invest more than \$500 million to develop small modular reactors," October 2024.
22. Constellation Research, "Google plans to use Kairos Power small modular reactors for data centers," October 2024.
23. Amazon, "Amazon signs agreements for innovative nuclear energy projects to address growing energy demands," October 2024.
24. Forbes, "Microsoft can take the lead in small modular reactors for powering AI," August 2024.
25. Financial Times, "US tech groups' water consumption soars in 'data centre alley'," August 2024.
26. OECD, "How much water does AI consume? The public deserves to know," November 2023.
27. Capgemini Research Institute, Creative and generative AI, July 2023.
28. IEA, Electricity 2024.
29. UNICEF, "Water scarcity: Addressing the growing lack of available water to meet children's needs," accessed in October 2024.
30. OECD, "How much water does AI consume? The public deserves to know," November 2023.

31. World Economic Forum, "The water challenge for semiconductor manufacturing: what needs to be done?," July 2024.
32. Data Center Dynamics, "Energy and e-waste: The AI tsunamis," October 2023.
33. Frontline Magazine, "E-waste from AI computers could 'escalate beyond control': study," October 2024.
34. AP News, "Google falling short of important climate target, cites electricity needs of AI," July 2024.
35. CIO.com, "3 things CIOs can do to make gen AI synch with sustainability," May 2024.
36. Latitude media, "Nvidia data center exec on near-term efficiency for AI computing," October 2024; TensorFlow, "Quantization aware training," accessed in October 2024; Amazon, "AWS unveils next generation AWS-designed chips," November 2023; Meta, "Introducing the AI Research SuperCluster – Meta's cutting-edge AI supercomputer for AI research," January 2022; Microsoft, "Sustainable by design: innovating for energy efficiency in AI, part 1," September 2024; Microsoft, "Splitwise: efficient generative LLM inference using phase splitting," July 2024; Salesforce, "Sustainable AI policy principles," accessed in October 2024; AWS, "Optimize generative AI workloads for environmental sustainability," September 2023; CNBC, "Amazon goes nuclear, to invest more than \$500 million to develop small modular reactors," October 2024; Reuters, "Meta Platforms strikes geothermal energy deal to power US data centers," August 2024; Meta, 2024 Sustainability Report, accessed in October 2024; Salesforce, "Sustainable AI policy principles," accessed in October 2024; Red Hat, "What is InstructLab?" May 2024.
37. TechTarget, "Generative AI energy consumption grows, but ROI role unclear," September 2024.
38. Technology Magazine, "Green AI: Building sustainability into AI initiatives," September 2024.
39. Capgemini Research Institute, Conversations for Tomorrow Edition 9.
40. Technology, "Green AI: Building sustainability into AI initiatives," September 2024.
41. Techopedia, "Exclusive interview with Salesforce: Can AI's energy demands destroy sustainability efforts?" May 2024.
42. Harvard Business Review, "How to make generative AI greener," July 2023.
43. Microsoft, "Reducing the environmental impact of generative AI: a guide for practitioners," April 2024.
44. Microsoft, "Tiny but mighty," April 2024.
45. Capgemini Research Institute, Conversations for Tomorrow Edition 9.
46. Token consumption is the process of breaking down text into smaller units or tokens, which the model processes to generate or understand language. The number of tokens consumed impacts the computational resources required and, consequently, the environmental footprint of the AI model.
47. Capgemini Research Institute, Conversations for Tomorrow Edition 9.

48. The Innovator, "Interview of the week: Zina Jarrahi Cinker, exponential technologies strategist," July 2024.
49. Exoswan, "These 6 neuromorphic computing startups could change AI forever," April 2024.
50. Mythic.ai, accessed in October 2024.
51. Arxiv, "Addition is all you need for energy-efficient language models," October 2024.
52. Element 84, "Sustainable AI: Reduce your carbon footprint, cut costs, and keep your users happy," October 2024.
53. Matthew Renze, Erhan Guven, "The benefits of a concise chain of thought on problem-solving in large language models," January 2024. <https://arxiv.org/html/2401.05618v1/#S1>
54. Carbon Credits, "How big is the CO2 footprint of AI models? ChatGPT's emissions," May 2023.
55. CIO.com, "3 things CIOs can do to make gen AI synch with sustainability," May 2024.
56. MIT Technology Review, "AI's carbon footprint is bigger than you think," December 2023.
57. Thomson Reuters, "ESG Case study: How EnerSys uses GenAI to drive efficiency, ensure accuracy, and safeguard sustainability & ESG data," August 2024.
58. Capgemini, "Sustainability data hub," accessed in November 2024.
59. Capgemini, "ESG lens," accessed in November 2024.
60. Automotive Dive, "GenAI's impact on the automotive industry and tips to capitalize on it," January 2024.
61. Forbes, "This company is using generative AI to design new antibodies," January 2023.
62. Siemens, "Heeds," accessed in December 2024.
63. Walmart Global Tech, "How Walmart plans to use AI to reduce waste," April 2024.
64. Reckitt, "Reckitt's first GenAI results demonstrate changing face of marketing during Cannes Lions international festival of creativity," July 2024.
65. Carrefour, "Carrefour integrates OpenAI technologies and launches a generative AI-powered shopping experience," June 2023.
66. Capgemini Research Institute, Turbocharging software with Gen AI: How organizations can realize the full potential of generative AI for software engineering, July 2024.
67. AIHR, "Generative AI in HR: Examples & how to successfully start using AI," accessed in November 2024.
68. Capgemini, "The synergy of generative AI and sustainable investing," November 2023.
69. NASA, "NASA turns to AI to design mission hardware," May 2024.
70. LEK, "Deploying AI and Generative AI in the airline industry," April 2024.
71. Autodesk, "General Motors: Driving a lighter, more efficient future of automotive part design," accessed in November 2024.
72. TechCrunch, "How generative AI is creeping into EV battery development," October 2023.
73. The Code Work, "The role of generative AI in ESG," September 2024.
74. Palo IT, "The role of generative AI In ESG," June 2024.
75. Citrine Informatics website, accessed in December 2024.
76. Amazon, "Iberdrola harnesses AWS generative AI technologies to drive innovation and sustainability in the energy sector," August 2024.
77. Palo IT, "The role of generative AI In ESG," June 2024.
78. Capgemini Research Institute, Climate tech: Harnessing the power of technology for a sustainable future, December 2023.

79. Data Center Dynamics, "Microsoft trains generative AI to handle SMR nuclear regulatory process," December 2023.
80. Capgemini, Technovision 2022, May 2022.
81. Capgemini, "The synergy of generative AI and sustainable investing," November 2023.
82. Merck, "Merck launches first ever AI solution to integrate drug discovery and synthesis," December 2023.
83. World Economic Forum, "These 5 cities are making innovative use of generative AI," July 2024.
84. Capgemini Research Institute, A world in balance 2024: Accelerating sustainability amidst geopolitical challenges, September 2024.
85. AWS Startup Blogs, "How climate tech startups use generative AI to address the climate crisis," February 2024.
86. Capgemini Research Institute, Harnessing the value of generative AI: 2nd edition Top uses cases across sectors, July 2024.
87. Compliance Solutions, accessed in November 2024.
88. ESG&I, "The disruption of traditional consulting by AI agents: a case study on ESG&I," accessed in November 2024.
89. Greenmetrica, "AI agents for sustainable enterprise: a leader's guide," accessed in November 2024.
90. Tech Target, "Generative AI energy consumption grows, but ROI role unclear," September 2024.
91. Capgemini Research Institute, Data-powered enterprises: The path to data mastery, August 2024.
92. Ibid.
93. Capgemini Research Institute, Conversations for Tomorrow Edition 9.
94. Capgemini, Technovision: Prompt the future, January, 2024.
95. Technology Magazine, "Green AI: Building sustainability into AI initiatives," September 2024.
96. Coursera, accessed in October 2024.
97. Ibid
98. Capgemini Research Institute, Harnessing the value of generative AI: Top use cases across industries, July 2023.
99. Roundtable on "Generative AI for the SDGs: Friend or foe, hope or hype?" hosted by the Broadband Commission, ITU, Global Partnerships Forum, and the Permanent Mission of Romania, May 2023.
100. Syngenta Group, "Syngenta Group adds cutting-edge generative artificial intelligence (GenAI) to Cropwise," September 2024.
101. Forbes, "How generative AI is accelerating drug discovery," June 2024.
102. Deccan Herald, "Apollo Hospitals, Google Cloud to deliver gen AI-powered healthcare service," September 2023.
103. World Economic Forum, "With Generative AI we can reimagine education — and the sky is the limit," February 2024.
104. UNDP, "How can generative AI contribute to improving quality education and research to impact the development of Cameroon?" October 2023.
105. Capgemini, The Capgemini Global Data Science Challenge: Harnessing AI for a sustainable future, accessed in November 2024.
106. Capgemini Research Institute, Conversations for Tomorrow Edition 9.
107. Nvidia, "How generative AI is empowering climate tech with NVIDIA Earth-2," April 2024.
108. British Ecological Society, "A deep generative Artificial Intelligence system to predict species coexistence patterns," February 2022.

Authors

Meet the experts



Cyril Garcia

Head of Global Sustainability Services and Corporate Responsibility and Member of the Group Executive Board, Capgemini cyril.garcia@capgemini.com

With more than 30 years of cross-industry consulting experience, Cyril has been at the forefront of the development of new services for Capgemini Group and its clients. Since 2023, Cyril Garcia has been acting as Head of Global Sustainability Services and Corporate Responsibility. In this role, he is responsible for driving positive impact across Capgemini's own operations, Corporate Social Responsibility activities, and through services for our clients, supporting sustainable enterprise transformation initiatives by deploying data, AI and climate technologies at scale.



Etienne Grass

Managing Director, Capgemini Invent France etienne.grass@capgemini.com

Having joined the Capgemini Group in 2017, Etienne dedicated his initial four years to the Public Services sector in France and later expanded his scope globally within Capgemini Invent. He is an expert in healthcare and administration transformations and an essayist (Génération réenchantée, Calmann-Lévy, 2016) and columnist for Les Echos.



Anne-Laure Thibaud

Head of AI/GenAI & Analytics Global Practice, Capgemini annelaure.thibaud@capgemini.com

Anne-Laure Thibaud has over two decades of experience in large-scale big data analytics and AI, and has observed the pivotal role that data and insights have come to play in shaping competitive organizations. Since joining the group in 2000, Anne-Laure has risen to a position of global leadership heading Capgemini's worldwide portfolio in data, analytics, intelligent automation, AI and GenAI. Her role is crucial in guiding Capgemini's clients to future-proof their businesses and stand out in today's data economy.



Vincent Charpiot

Executive Vice President, Head of Group Sustainability Business Accelerator, Capgemini vincent.charpiot@capgemini.com

Based in New York, Vincent has over 25 years of experience supporting global organizations across industries on their digital transformations and more recently, on their sustainability journeys as Capgemini Group Sustainability Business Accelerator head since 2023. Vincent graduated from EM Lyon in France and started his career in the technology industry in the UK and France, before joining Capgemini in 2008. His passion is to help companies leverage innovative digital and engineering solutions to deliver sustainable and business outcomes.



Steve Jones

EVP of Data Driven Business and Generative AI, Capgemini steve.g.jones@capgemini.com

Steve brings over 21 years of innovation experience. He was a pioneer in mobile applications within the Capgemini group and played a key role in advancing Capgemini's early initiatives in SaaS and Cloud. Since then, Steve has established the Master Data Management and Big Data businesses at Capgemini. Currently, he leads the Group's efforts in Trusted AI and Collaborative Data Ecosystems and is a prominent writer and conference presenter on these topics.

Authors

Meet the experts



Christopher Scheefer

Vice President, Global Data & AI Sustainability Lead, Intelligent Industry, GenAI Ambassador, Capgemini
christopher.scheefer@capgemini.com

Christopher Scheefer is the Global Sustainability Leader for Data and AI at Capgemini, based in North America, with over two decades of experience in sustainability advisory and data & analytics leadership. A recognized thought leader, speaker, and author, Chris specializes in driving sustainable business transformation through artificial intelligence and automation at scale. As a Global Generative AI Ambassador, he has played a pivotal role in integrating Artificial Intelligence, Climate Tech, and Energy Transition technologies into corporate value chains, fostering resilience and purpose-led growth.



Elizabeth Lugnier

Director, Head of Sustainability Portfolio Offer Management, Capgemini
elizabeth.lugnier@capgemini.com

Elizabeth oversees the sustainability offer portfolio at Capgemini, collaborating with both internal and external ecosystems to design and deliver services that create both business and planetary value. Prior to joining Capgemini, Liz spent ten years in data-driven strategy roles where she built actionable, integrated, and automated reporting across disparate systems, departments, and geographies at Hewlett Packard.



Dorothea Pohlmann

CTO Sustainability, Capgemini Engineering Global
dorothea.pohlmann@capgemini.com

Dorothea Pohlmann holds a Ph.D. in Physics and has worked for more than 17 years at Capgemini. In her current role as global CTO Sustainability at Capgemini Engineering, she is responsible for advising clients on business and engineering transformation projects focusing on sustainable products, circular economy, climate tech, enabled by the use of digital technologies, such as generative AI and digital twins.



Vincent De Montalivet

Director, Head of Sustainability Insights Data Practices, North America, Capgemini
vincent.de-montalivet@capgemini.com

Vincent is an architect based in Chicago with 15+ years of experience, and sponsor of the Capgemini Sustainable AI Program. He helps clients leverage digital technologies to transition to a low-carbon future and ensure ESG compliance. A graduate in engineering and sustainability business management from Stanford, Vincent is an expert in CSR, ESG reporting, and carbon accounting. He empowers organizations to use data for sustainability transformations and to monitor their ESG commitments. Vincent has also delivered a TEDx talk on how AI can help save the planet and developed the GreenAI framework.

Authors

Meet the Capgemini Research Institute



Jerome Buvat

Head of
Capgemini Research Institute
jerome.buvat@capgemini.com

Jerome is the head of the Capgemini Research Institute. He works closely with industry leaders and academics to help organizations understand the business impact of emerging technologies.



Nancy Manchanda

Director,
Capgemini Research Institute
nancy.manchanda@capgemini.com

Nancy has over 16 years of experience in market research, consulting and digital transformation. She collaborates with business leaders to understand challenges and opportunities relating to technology, people and sustainability transformation.



Subrahmanyam KVJ

Senior Director,
Capgemini Research Institute
subrahmanyam.kvj@capgemini.com

Subrahmanyam is a Senior Director at the Capgemini Research Institute. He loves exploring the impact of technology on business and consumer behavior across industries in a world being eaten by software.



Amrita

Senior Manager,
Capgemini Research Institute
amrita.a.amrita@capgemini.com

Amrita has 13 years of experience in market research and consumer insights, with extensive exposure to CPG. Her doctoral work focused on understanding consumer behavior for mobile wallets usage.

The key contributors would like to thank Amol Khadikar and Anjali Roy for their contribution to this research.

The key contributors would also like to thank Ajay Mohan, Alex Marandon, Andy Vickers, Ane-Marte Weng, Aniruddha Guha Biswas, Anne-Sophie Colomberon, Antoine Tillette De Mautort, Aurélie Lustenberger, Ayan Bhattacharya, Bala Natarajan, Catherine Joce, Chris Hodgson, Conor MCGovern, Danica Vukadinovic Greetham, Daniela Rittmeier, Edmond Segalen, Emmanuelle Bischoffe-Cluzel, Eric Benoits, Florent Andrillon, Franco Amalfi, Greg Bentham, Hinrich Thölken, James Robey, Jean-Baptiste Perrin, Joyce Chew, Karen Balbi, Kary Bheemaiah, Alexandre Lapene, Lee Beardmore, Liza Garay De Vaubernier, Lukas Birn, Marie Uhlirova, Marie-Fleur Revel, Mark Oost, Mark Roberts, Maxime Wyka, Mona, Naresh Khanduri, Natalia Sanchez Bella, Neelam Gupta, Neha Bhatt, Noémie Lauer, Ozlem Bozyurt, Pascal Brier, N Periyasamy, Peter King, Philipp Fürst, Phillippe Cordier, Pierre-Adrien Hanania, Pierre-yves Glever, Robert Engels, Ron Tolido, Salomon Salinas, Sergey Patsko, Stephane Girard, Thordur Arnason, Valérie Perhirin, Victoire Grux, Vincent Godiner, WeiWei Feng, Yashowardhan Sowale, Shashank Kaurav, Punam Chavan, Amitabha Dutta, Jaydeep Neogi, Suparna Banerjee, Aparajita Paul for their contribution to the report.

About the Capgemini Research Institute

The Capgemini Research Institute is Capgemini's in-house think tank on all things digital. The Institute publishes research on the impact of digital technologies on large traditional businesses. The team draws on the worldwide network of Capgemini experts and works closely with academic and technology partners. The Institute has dedicated research centers in India, Singapore, the UK, and the US. We are proud to have been ranked #1 in the world for the quality of our research by independent analysts for six consecutive times - an industry first.

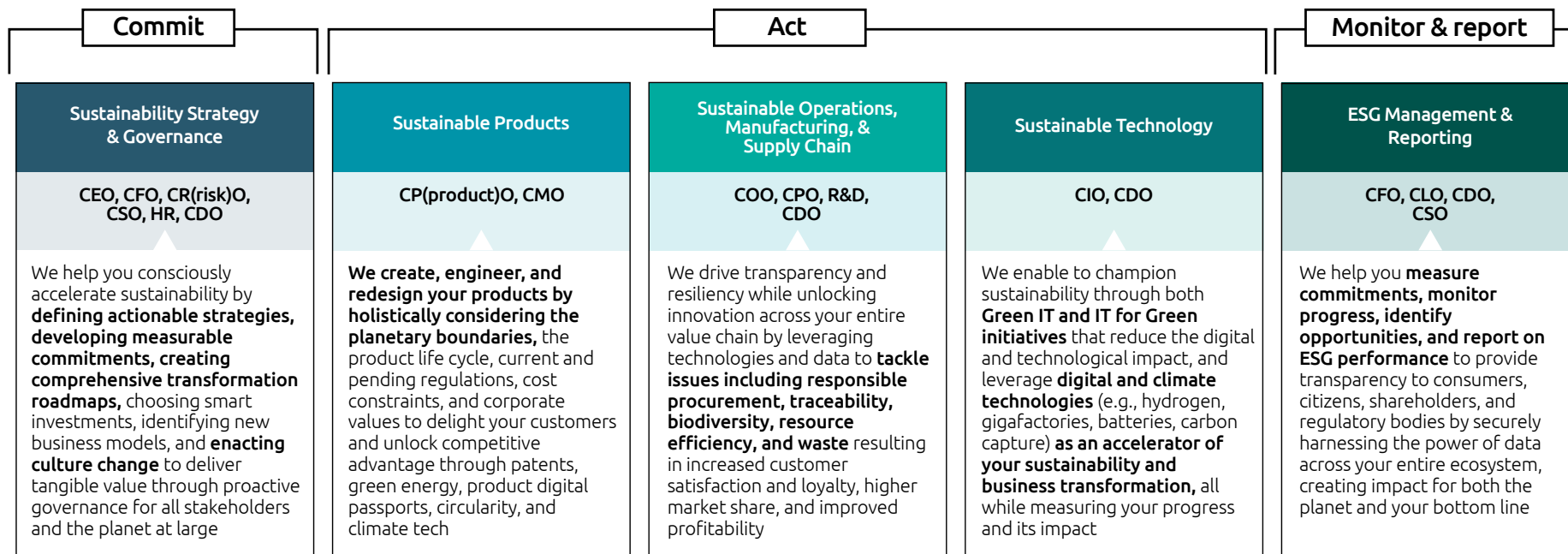
Visit us at www.capgemini.com/researchinstitute/

Why Partner with Capgemini?

Gen AI has the potential to accelerate sustainable business initiatives, optimize costs, drive efficiencies and propel growth. Nevertheless, this potential must be weighed against the technology's environmental cost. As the leader of the Eco-Digital Revolution, we are committed to helping our clients harness the immense potential of Generative AI while considering its cost to the planet and society.

We help clients to identify the right technology for each use case, assess and mitigate GenAI's environmental impact, harness Gen AI's potential to accelerate sustainable business value by scaling the right use cases, develop the right data and technology foundations, and set a governance for sustainable generative AI. We do this through our comprehensive set of generative AI offerings built on our market leading Data and AI experience in addition to our portfolio of sustainability offers which helps organizations consciously accelerate their sustainability journey, from commitments to sustainable achievements, creating value for business and the planet.

Why Partner with Capgemini?



"continued to the next page"

Generative AI strategy		
Define and prioritize the most relevant Gen AI use cases for business, and lay the right foundations in terms of people, process, and technology as well as a compelling business case to deliver business outcomes at the right cost.		
Custom generative AI for enterprise		
Move beyond the generic models and solutions with custom models created specifically to address your use case and industries. Industrialize and accelerate custom Gen AI deployments in a cost-efficient manner, with the relevant guardrails to ensure trusted outcomes.		
Generative AI for customer experience	Generative AI for software engineering	Generative AI for software product engineering
Personalize every customer experience and empower agent efficiency across sales, service, commerce, and marketing.	Improve efficiency and quality across the software lifecycle, accelerate the time to market for new software and improve security.	Enhance software products with hyper-personalized experiences, improve operational efficiency, and enable scalable growth.

Learn more about our sustainability portfolio: <https://www.capgemini.com/services/sustainability>

Learn more about our Generative AI offerings and capabilities: <https://www.capgemini.com/services/data-and-ai/generative-ai>

As part of Capgemini's [Code of Ethics for AI](#), we pledge for Sustainable AI, which includes call for increased transparency from hyperscalers, Gen AI model providers, on their AI usage and environmental footprint, through a cradle to grave methodology, and deployment of automated Gen AI systems only for use cases where significantly higher business values are created as compared to alternate, more energy efficient solutions. We also support [energy scoring of AI models](#).

For more information, please contact:

Cyril Garcia

Head of Global Sustainability Services and Corporate Responsibility and Member of the Group Executive Board, Capgemini
cyril.garcia@capgemini.com

Vincent Charpiot

Executive Vice President, Head of Group Sustainability Business Accelerator, Capgemini
vincent.charpiot@capgemini.com

Florent Andrillon

Head of Group Climate Tech, Vice President Capgemini
florent.andrillon@capgemini.com

Mark Oost

Global Offer Leader AI, Analytics, and Data Science, Insights and Data, Capgemini
mark.oost@capgemini.com

Ayan Bhattacharya

Vice President, Generative AI Insights & Data, North America, Capgemini
ayan.bhattacharya@capgemini.com

Vincent De Montalivet

Director, Head of Sustainability Insights Data Practices, North America, Capgemini
vincent.de-montalivet@capgemini.com

Anne-Laure Thibaud (Thieullent)

Head of AI/GenAI & Analytics Global Practice, Capgemini
annelaure.thibaud@capgemini.com

Christopher Scheefer

Vice President, Global Data & AI Sustainability Lead, Intelligent Industry, GenAI Ambassador, Capgemini
christopher.scheefer@capgemini.com

Ozlem Bozyurt

Principal, Sustainability, Capgemini
ozlem.bozyurt@capgemini.com

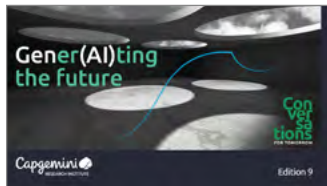
Étienne Grass

Managing Director, Capgemini Invent France
etienne.grass@capgemini.com

Alex Marandon

Vice President, Global Head of Generative AI Accelerator, Capgemini Invent
alex.marandon@capgemini.com

More Capgemini Research Institute Publications



Gener(AI)ting the future



Harnessing the value of generative AI 2nd edition: Top use cases across sectors



Turbocharging software with Gen AI: How organizations can realize the full potential of generative AI for software engineering



A World in balance 2024: Accelerating sustainability amidst geopolitical challenges



The Eco-Digital Era™: The dual transition to a sustainable and digital economy



Data for net zero: Why data is bridging the gap between net zero ambition and action



Data-powered enterprises: The path to data mastery



Climate tech: Harnessing the power of technology for a sustainable future



Sustainable IT: Why it's time for a Green revolution for your organization's IT



Climate AI: How artificial intelligence can power your climate action strategy

Subscribe to latest research from the Capgemini Research Institute



Receive copies of our reports by scanning the QR code or visiting

<https://www.capgemini.com/capgemini-research-institute-subscription/>

Capgemini Research Institute

Fields marked with an * are required

First Name *

Last Name *

Email *

By submitting this form, I understand that my data will be processed by Capgemini as indicated above and described in the [Terms of use](#).

Submit





About Capgemini

Capgemini is a global business and technology transformation partner, helping organizations to accelerate their dual transition to a digital and sustainable world, while creating tangible impact for enterprises and society. It is a responsible and diverse group of 340,000 team members in more than 50 countries. With its strong over 55-year heritage, Capgemini is trusted by its clients to unlock the value of technology to address the entire breadth of their business needs. It delivers end-to-end services and solutions leveraging strengths from strategy and design to engineering, all fueled by its market leading capabilities in AI, cloud and data, combined with its deep industry expertise and partner ecosystem. The Group reported 2023 global revenues of €22.5 billion.

Get the Future You Want | www.capgemini.com