# ARTIFICIAL INTELLIGENCE FOR CLIMATE CHANGE MITIGATION ROADMAP (SECOND EDITION)

CHAPTER 15:

# GREENHOUSE GAS EMISSIONS FROM AI

*David Sandalow*

November 2024

Innovation for Cool Earth Forum

**CHAPTER 15:**

# GREENHOUSE GAS EMISSIONS FROM AI

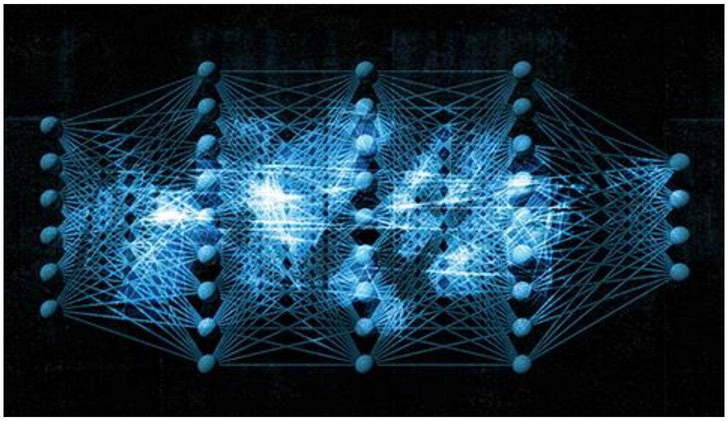
*David Sandalow*

AI systems need energy. Manufacturing silicon chips requires energy for mining minerals and operating complex machinery. Building data centers requires energy for making steel and concrete. Training and running AI models requires energy for electricity to power servers. Lighting and cooling data centers requires energy for electricity as well.

This energy use does not necessarily result in significant greenhouse gas (GHG) emissions. When the electricity for a data center comes from new solar, wind or nuclear power, for example, the GHG emissions from data-center operations are modest. Amazon, Microsoft, Google and Meta—the world's largest data center operators—are among the world's largest purchasers of renewable power.[1] However some activities essential for AI—such as making steel and concrete—use only modest amounts of low-carbon energy.

A review of the current literature suggests the following conclusions:

- Current overall impacts of AI on GHG emissions could be positive or negative. Much better data collection is needed to assess overall impacts with confidence.
- GHG emissions from generating power for AI operations at data centers and on edge devices ("AI operational emissions") are less than 1%—and perhaps much less than 1%—of global GHG emissions.
- AI operational emissions will likely increase in the years ahead. This increase could be modest or quite substantial.
- In the medium- to long-term, the overall impacts of AI on GHG emissions could be positive or negative. The GHG benefits of using AI throughout the economy could significantly outweigh GHG emissions increases due to AI. However, the opposite could occur as well. The impact of AI on GHG emissions will depend on decisions by policymakers, business leaders, researchers and others in the years ahead.



This chapter starts with background on GHG emissions from AI and data center power demand. With that foundation, the chapter examines current and future GHG emissions from AI, concluding with recommendations.

## A. Background

The phrase "GHG emissions from AI" is quite broad. It includes:

- AI operational emissions,
- GHG emissions from manufacturing equipment and building infrastructure used for AI ("AI upstream emissions") and

- The emissions impacts of applying AI in countless thousands of ways throughout the economy, some of which reduce GHG emissions (such as the many applications of AI discussed in this Roadmap) and some of which increase GHG emissions (such as when AI is used to cut the cost of some polluting activities).

Estimating GHG emissions from AI is challenging, for several reasons.

First, data collection and assessment methodologies are inadequate. The lack of standardized reporting practices and metrics across the AI industry makes it difficult to provide precise and confident emissions estimates.[2-4]

Second, the shared use of computing resources in cloud environments can make it difficult to isolate and accurately attribute emissions to AI-related activities. Data center operators do not routinely keep records distinguishing the time a server is running AI-based software from the time a server is running non-AI-based software. (Doing so would be difficult.) As a result, it can be challenging to correctly allocate overall GHG emissions from computing infrastructure to the subcategory of AI applications.

This challenge is diminished by the increasing use of specialized computing chips, such as graphics processing units (GPUs) and tensor processing units (TPUs), which are used almost exclusively for AI-based software. However allocating emissions from other AI hardware can be a challenge.

Third, data center emissions are location-specific. A data center's GHG emissions depend on the fuels used to generate electricity for that data center. Many data centers purchase electricity from local power grids, and the fuel mix in local power grids varies greatly around the world. To project future GHG emissions from data centers, one must make assumptions about not only the increase in overall data center power demand but also the locations where data centers will be built and the sources of electricity data centers will use.

Finally, AI is a transformational technology at an early stage of deployment. Forecasting how AI will impact many economic processes and societal patterns in the years and decades ahead is difficult if not impossible. As a result, forecasting the GHG impacts of AI deployment with high confidence is challenging as well.[5]

Despite these challenges, a growing body of literature seeks to estimate current and future GHG emissions. These studies are essential for understanding and managing AI's GHG impacts. After reviewing the related topic of data center power demand, we examine these studies below.

## B. Data Center Power Demand

There are roughly 11,000 data centers globally (Aljbour et al, 2024[6] at p. 11). Roughly half of global data center capacity is in the United States, 15% is in Europe and 15% is in China.[7]

Data centers are central to the AI industry. Most AI models are trained, tuned and run at data centers. Although some AI computation is beginning to move to edge devices, most AI takes place at data centers and will continue to do so for the foreseeable future.[8-10]

Data centers perform many functions other than AI—hosting websites, processing financial transactions, running email networks and much more. Only a fraction of data center workload is attributable to AI. Recent estimates of that fraction vary widely:

- KKR Insights estimates that, today, roughly 35% of the workload at Amazon, Google, Meta and Microsoft data centers is for AI and that this figure will rise to more than 50% by 2030.[11]
- A 2022 paper in *Nature Climate Change* by Lynn Kaack et al. estimates that "less than one-quarter" of the workloads and traffic of cloud and hyperscale data centers is related to machine learning (ML).[3]
- FTI Consulting estimates that roughly 10% of data center power demand globally is for AI, growing to roughly 25% by 2030.[12]
- The Electric Power Research Institute (EPRI) estimates that about 10–20% of data center electricity use comes from AI applications.[13]
- A 2024 paper in *Communications of the ACM* by David Patterson et al. estimates that, from 2019 to 2021, ML "represented between 10% and 15% of the total annual operational energy use in the Google cloud" (Patterson et al., 2024[14] at p. 88).
- Goldman Sachs estimates that the percentage of data center workload attributable to AI globally was less than 1% in 2024 but will increase to roughly 19% by 2028 (see "Data center power demand graph[15]).
- A paper published in *Nature* by Amy Luers et al. in April 2024 estimates that roughly 1% of data center power demand in 2023 came from AI processors.[5]

The wide differences in these estimates reflect different definitions of "AI" (with some studies focused on generative AI and others on ML more broadly), data gaps, the lack of standard measurement protocols and other factors.

In the past year, data center power demand has received considerable media attention, often in the context of the growth of AI.[16-18] We explore that topic below.

### i. Current data center power demand

Data centers use substantial amounts of electricity. To operate a data center, electric power is needed for servers, data storage equipment, networking equipment, cooling systems, lighting and more.

In 2023, roughly 1.5% of global electricity demand came from data centers (IEA 2024[19] at p. 19). In the United States, data centers were responsible for 3% of electricity demand.[20] The figure was 1–2% in Japan,[21] 3.5% in China[22] and 3.5% in the European Union.[23]

Although these amounts are significant, they are smaller than the electricity used in some other sectors. In 2023, for example, 4% of global electricity demand came from aluminum smelters (IEA 2024[19] at p. 19). According to IEA experts, "annual electricity consumption from data centers globally is about half of the electricity consumption from household IT appliances, like computers, phones and TVs."[24]

Data centers tend to be built in clusters. In places where data centers are concentrated, their share of power demand is much greater than the global average. In Loudon County, Virginia, USA—which has the world's largest number of data centers by far—roughly a quarter of electricity demand comes from data centers.[25] In Ireland (the largest data center hub in Europe), 21% of electricity demand came from data centers in 2023.[26] In Singapore (one of the leading data center hubs in Asia), 7% of electricity demand comes from data centers.[27]

### ii. Future data center power demand

Data center power demand is growing rapidly. Goldman Sachs Research projects 160% growth globally by 2030.[15] EPRI projects 5–15% annual growth in the United States until 2030 (EPRI 2024[6] at p. 5), several research firms project annual growth in the 7–9% range in the European Union[23,28,29] and the Open Data Center Committee projects annual growth of roughly 10% in China.[30]

The growth in data center power demand is coming from many sources, not just AI. Streaming services, 5G networks, social media and online gaming are all fueling surging data center demand.[11,31] Yet AI is an important (and perhaps the most important) factor.[30]

Although power demand from data centers is growing rapidly, it is smaller than power demand growth from several other sectors. In the IEA's Stated Policies Scenario, power demand growth for electric vehicles (EVs) and space cooling in buildings are each more than three times greater than power demand growth for data centers. According to IEA, "data centers look set to remain a relatively small driver of overall electricity demand growth at the global level in the decade to come. Nonetheless, constraints at the local level may be significant." (IEA 2024[32] at p. 188.)

Those constraints are especially significant in countries including the United States, Ireland, Singapore and Japan. In the past several years, electric utilities in these countries and other locations have received a record-breaking number of requests from data center operators for electricity interconnections. These requests are creating significant challenges. In Loudon Country, Virginia, for example, applications for electricity interconnection from data center operators are currently facing several years of delay. These applications are experiencing similar delays in many other locations as well.[12,33]

However, many of the applications for electricity interconnection submitted by data center operators do not represent actual demand. Due to delays and uncertain prospects for approvals, many data center operators have applied for more interconnections than they need, hoping that some applications will be successful. This "application frenzy" has some similarities to a run on a bank or the panic buying of essential goods at the start of the COVID epidemic.[34]

Still, data center power demand is rising rapidly.[35] In the past year, many research organizations, investment banks, consultancies and energy companies have released forecasts for increased power demand from data centers. Table 1 summarizes the results of some of these studies.

*Table 1. Power consumption projections for data centers.*

| AUTHOR | PROJECTED ANNUAL GROWTH RATE | TIMEFRAME | REMARKS |
|---|---|---|---|
| **Global** | | | |
| IEA, *Electricity 2024* (January 2024)[36] at p.31 | 21% | 2022–2026 | Electricity consumption by data centers, cryptocurrencies and AI globally increases from 460 TWh in 2022 to 620–1050 TWh by 2026 |
| IEA, *Electricity Mid-Year Report* (July 2024)[19] at p.19 | 19% | 2022–2026 | Electricity consumption of data centers increases from 1–1.3% of global demand in 2022 to 1.5–3% by 2026 |
| Goldman Sachs Research, 2024 (May 14, 2024)[15] | 14.5% | 2023–2030 | Electricity consumption by global data centers increases from 411 TWh in 2023 to 1063 TWh in 2030; AI's percent of global data center load increases from 3% in 2023 to 20% in 2030<br>Data centers increase from 1–2% of global electricity consumption now to 3–4% by end of the decade |
| SemiAnalysis, 2024[37] | 25% | 2024–2030 | Electricity consumption by data centers reaches 4.5% of global consumption by 2030 |
| Morgan Stanley, 2024[38] | 70% (GenAI only) | 2024–2027 | Global power usage from GenAI grows by 70% CAGR (compound annual growth rate) in 2024–2027 to 224 TWh |

| | | | |
|---|---|---|---|
| **United States**[33] | | | |
| EPRI, 2024 (May 28, 2024)[6] | 5–15% | 2023–2030 | Electricity consumption by US data centers increases from 150 TWh in 2023 to 196–404 TWh by 2030, taking 5–9.1% of 2030 electricity consumption |
| BCG, 2024[39] | 15–20% | 2024–2030 | Electricity consumption by US data centers increases to 800–1050 TWh (100–130 GW capacity) by 2030 |
| McKinsey, 2023[40] | 9.5% | 2022–2030 | Electricity consumption by US data centers increases from 149 TWh (17 GW capacity) in 2022 to 307 TWh (35 GW capacity) in 2030 |
| Columbia Center on Global Energy Policy, 2024[41] | | 2024–2027 | In 2027, GPUs will be roughly 4% of total US electricity sales and roughly 1.7% of total electric capacity |
| **European Union** | | | |
| Joint Research Centre EU, 2024 at pp.3,8[23] | 5–17% | 2022–2030 | Electricity consumption by EU data centers increases from 45–65 TWh in 2022 to 98.5–160 TWh in 2030 |
| Savills, 2024[29] | 8.3% | 2024–2027 | 27% increase to 13.1 GW capacity in 2027 |
| Mordor Intelligence, 2024[28] | 7.4% | 2024–2029 | Data centers reach 3.2% of EU electricity consumption in 2030, citing official EU sources |
| **China** | | | |
| China State Grid Energy Research Institute, 2021[42] | 7.1% | 2020–2030 | Electricity consumption by data centers increases from 200 TWh in 2020 (2.7% of total power demand) to 400 TWh in 2030 (3.7% of total power demand) |
| China Com-service White paper, 2023[43] | 6% | 2022–2025 | Electricity consumption by data centers in China increases from 101 TWh in 2022 to 120 TWh in 2025 |
| **Japan** | | | |
| Japan Transmission Operators, 2024[21] | 6–12% | 2022–2050 | Electricity consumption by data centers owned by three leading communications companies in Japan increases from 8.6 TWh in 2022 (slightly less than 1% of total power demand) to 43–211 TWh in 2050 |

# RESOURCE ADEQUACY

*by Mariah Frances Carter and David Sandalow*

"Resource adequacy" is the ability of an electric utility to meet the needs of its customers even during periods of peak usage or unexpected disruptions.

When a utility experiences resource adequacy problems, several issues can arise:

- First, blackouts or brownouts become more likely, especially during extreme weather events and other periods of high demand. This occurs because the utility may not have enough generation capacity or demand response resources to meet the peak electric load.
- Second, higher electricity prices are possible because the utility may need to purchase power at premium prices or rely on expensive, less efficient and more polluting peaker plants to meet demand.
- Third, the stability and resilience of the electricity system can be compromised, causing operational problems with grid management.

Surging power demand—in part due to data centers—is causing resource adequacy problems in some regions around the world. This demand surge contrasts sharply with the experience in most developed countries in recent years. For most of the past two decades, power consumption in the United States, Europe and Japan was mostly flat. However, this is changing dramatically as new factories, EVs, data centers, crypto currencies and other sources create significant new demand for electric power. The International Energy Agency (IEA) projects power demand in the United States will grow 1.5% per year in 2024–2026, with a third of that growth due to data centers (IEA, 2024[36] at p. 111). The Japanese government recently released a report forecasting an increase in long term electricity demand for the first time in twenty years, due in significant part to semiconductor plants and data centers. The report estimates that electricity demand will grow from 1 trillion kilowatt-hours (kWh) in this decade to about 1.35-1.5 trillion kWh in 2050.[44]

Power demand is growing especially fast in regions where data centers are clustered. In the United States, this includes Northern Virginia, Dallas-Ft. Worth, Chicago, Silicon Valley and Phoenix. (The Phoenix-based Arizona Public Service recently estimated average load growth in its service territory of 3.7% per year from 2023 to 2038. This is an additional 24 TWh of annual electricity consumption, with more than half of that increase coming from data centers.)[45] Globally, top areas include Frankfurt, London, Paris, Singapore, Tokyo, Hong Kong, Sydney and Querétaro (Mexico). All of these regions are facing 20–25% annual growth in data center capacity with significant related power demands.[46]

Utilities in regions with high concentrations of data centers are responding to this increased demand with new generation, demand response and other tools. In Ohio, one utility is asking permission to impose special tariffs on data center customers to help pay for expanding and strengthening the grid.[47] However the growth in power demand is outpacing the utilities' ability to respond in some places. Power connections for new data centers will need to be delayed—in some cases for years—to address resource adequacy concerns.[12]

## C. Current GHG Emissions from AI

Current overall impacts of AI on GHG emissions could be positive or negative. Assessing those impacts with confidence is difficult due to gaps in data collection, a lack of standard assessment methodologies and the rapid pace of AI deployment in recent years.

Recent studies suggest the following:

- AI operational emissions are less than 1%—and perhaps much less than 1%—of total GHG emissions.
- AI upstream emissions contribute to AI's GHG footprint. Much better data are needed to assess the magnitude of these emissions with confidence.
- The GHG impacts of applying AI in countless thousands of processes throughout the economy are difficult to assess. These impacts could be beneficial on a net basis, outweighing AI operational emissions, AI upstream emissions and other GHG increases associated with AI. However, these impacts could also be negative on a net basis, increasing global emissions.

This section discusses each of these topics in turn.

### i. AI operational emissions

Based on the existing literature, it is reasonable to conclude that GHG emissions from computing operations for AI are less than 1%—and perhaps much less than 1%—of global GHG emissions.

Relevant studies include the following.

- In a 2024 *Nature* article, Amy Luers et al. wrote that "in terms of total global greenhouse-gas emissions, we calculate that AI today is responsible for about 0.01%."[5] The estimate is based on the power consumption of AI processors in 2023.
- In a 2022 *Nature Climate Change* article, Lynn Kaack et al. estimated that cloud and hyperscale data centers are responsible for 0.1–0.2% of global GHG emissions and that roughly 25% of their workloads are related to ML.[3]
- In a 2022 study, Sasha Luccioni et al. found that GHG emissions from training several current large language models (LLMs), including GPT-3 and BLOOM, ranged from roughly 30 to 550 tonnes $CO_2$e.[48] In a 2021 paper, David Patterson et al. provided similar estimates (noting that

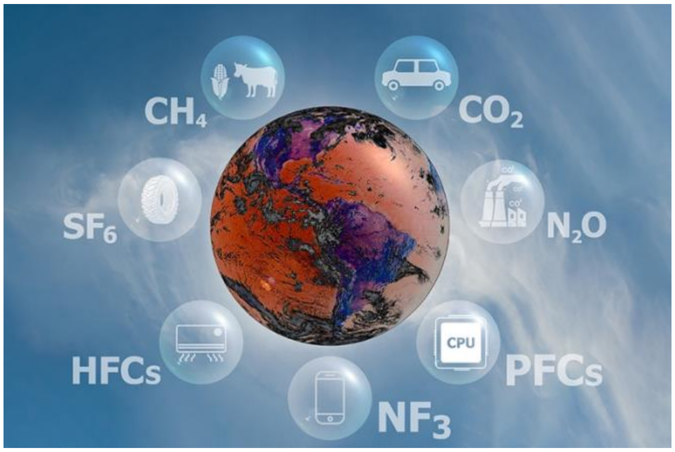
the average commercial plane emits roughly 180 tonnes $CO_2e$ flying from San Francisco to New York).[49] (550 tonnes $CO_2e$ is roughly 0.000001% ($1x10^{-8}$) of global GHG emissions, which were roughly 54 $GtCO_2e$ in 2022.)[50]

- In a 2023 report, IEA estimated that "Data centres and data transmission networks are responsible for 1% of energy-related GHG emissions." The estimate included both upstream and operational emissions.[51]
- In a 2021 paper in *Patterns*, Charlotte Freitag et al. estimated that 1.8–2.8% of global GHG emissions came from the information, communications and technology sector. This estimate included both upstream and operational emissions.[52]

These studies explore related but somewhat different topics, offering a range of results. Some of the studies are based on data that are several years old and therefore partly out of date. (The AI market is growing rapidly—at compound annual growth rates in the range of 35% according to some estimates.[53-55]) However, combined with the estimates of AI's share of data center workload (summarized in Section B of this chapter above), these studies suggest that 1% is a likely upper bound for the share of global GHG emissions from computing operations for AI and that the actual share could be much less.

### ii. AI upstream emissions

Upstream emissions from AI must be part of any complete GHG accounting for AI; however, the literature on upstream emissions from AI is sparse.[56,57] A research agenda to better assess the magnitude of AI upstream emissions should consider several factors, including the following.




First, many upstream AI activities, such as manufacturing silicon chips and making steel and cement for data centers, rely heavily on fossil fuels for energy. This contrasts with AI operations at data centers, where power use is often matched with renewable energy.

Second, major data center operators, including Google and Microsoft, report that the vast majority of their emissions are Scope 3 emissions (defined as "indirect emissions in the value chain of a company, other than emissions from the generation of purchased energy").[58] For Google, the figure is 75% (Google, 2024[59] at p. 38), and for Microsoft it is 96% (Microsoft 2024[60] at p. 15). Scope 3 is a broad category that includes many sources of emissions beyond AI upstream emissions, but still these corporate reports suggest the possibility that upstream emissions from AI could be significant and merit attention. (Again, more research is needed.)

Third, studies that have begun to explore topics related to upstream emissions from AI include:

- A 2024 paper in *Communications of the ACM* by David Patterson et al., which found that "embodied server CO2e was ~115x larger than ML operational CO2e in Google datacenters in 2021" (at p.95).[14]
- A 2024 IEEE paper by Carole-Jean Wu et al., which found that upstream GHG emissions for University LM, a multilingual language translation model, were roughly 50% of operational emissions.[61]
- A 2021 study in *HAL Open Science* by Maxime Pelcat, which found that annual emissions from semiconductor manufacturing were roughly 76.5 Mt CO2e globally (0.15% of global GHG emissions).[62] Semiconductor manufacturing is an important part of the value chain for AI, although semiconductor chips are used in countless thousands of products and only a small fraction of semiconductor chips manufactured each year are used in AI.

### iii. Impacts of AI applications on emissions

Data quantifying the current impacts of AI applications on GHG emissions are sparse.

The phrase "impacts of AI applications on GHG emissions" is potentially confusing. In this context, it means how use of AI impacts GHG emissions, not including AI operational emissions or AI upstream emissions. For example:

- When a municipality uses AI tools to help with traffic management, how much do vehicle emissions fall?
- When a commercial building uses AI tools to help with energy management, how much do emissions at that building and at the local power grid fall?
- When an industrial facility uses AI tools in its operations, how much do emissions at that facility rise or fall?

A few studies have estimated the current GHG emission benefits that come from using AI in some settings.

- In a 2021 report, BCG experts reported that their clients had achieved 5–10% emissions reductions using AI[63]
- In a 2021 report, Capgemini reported that organizations had reduced GHG emissions by 13% using AI[64]

However, the literature on this topic is sparse. Qualitative and anecdotal assessments are more common than quantitative assessments. Few if any studies have attempted to quantify the potential emissions benefits of AI-enabled breakthroughs in areas such as battery chemistry or carbon capture. Chapters 3–13 of this Roadmap contain many examples of ways in which AI is currently being used to reduce GHG emissions, including the use of AI to monitor methane emissions, optimize fertilizer application, improve low-carbon steel manufacturing and much more. Taken together, these and other AI applications may already be having a meaningful impact in reducing GHG emissions. However, much more data collection and analysis are needed to provide rigorous estimates.

The literature on the extent to which AI applications may be increasing GHG emissions is especially sparse. When AI is used in carbon-intensive industries, such as mining, manufacturing and oil-and-gas production, AI could increase GHG emissions by making carbon-emitting activities more cost-competitive. In recent years, the oil and gas industry has rapidly adopted AI tools in exploration and production activities, improving operational efficiencies and cutting costs.[65-67] Lower-cost oil and gas production seems likely to lead to higher GHG emissions, although the analysis is complicated by (1) the potential for cheap natural gas to reduce GHG emissions by displacing coal, if leakage rates for that natural gas are kept to a minimum, and (2) the partially-managed nature of global oil markets. (See text box below.)

Some AI applications are currently reducing GHG emissions. Other AI applications are probably increasing GHG emissions. Comprehensive data on the cumulative impacts of AI applications on GHG emissions are lacking.

### iv. Further study

In an interesting 2024 paper in *Scientific Reports*, Bill Tomlinson et al. compare (1) GHG emissions that come from using AI for writing and drawing tasks (both upstream and operational emissions) with (2) the GHG footprint of humans performing the same tasks. Tomlinson et al. found that "AI systems emit between 130 and 1500 times less $CO_2$e per page of text generated compared to human writers, while AI illustration systems emit between 310 and 2900 times less $CO_2$e per image than their human counterparts."[68]

The literature on GHG emissions from AI is growing.[69-71] However there are no widely used protocols or standards for measuring GHG emissions from AI systems or the GHG benefits of AI applications. Improved measurement protocols and standards—and much more research—are needed to provide precise and confident estimates of current emissions.

## AI IN THE OIL AND GAS INDUSTRY

AI is widely used in the oil and gas industry.[72-74] Some ways AI is used may increase GHG emissions; other ways may decrease emissions. On a net basis, AI appears likely to be increasing GHG emissions from the oil and gas industry, however no studies have rigorously analyzed this topic to date.

Use of AI use in the oil and gas industry has grown rapidly in recent years. AI is being used for predictive maintenance, supply chain optimization, performance improvements at refineries and much more. AI is increasing yields from reservoirs, expanding areas where drilling is economic and cutting costs in exploring for oil and gas. Many industry testimonials cite the benefits of AI for oil and gas production.[75-77]

To the extent that AI is helping oil and gas companies produce more oil and gas at lower cost, higher GHG emissions are likely to be one result. In general, lower production costs for goods put downward pressure on prices for those goods, increasing consumption. More consumption of fossil fuels, such as oil and gas, generally increases GHG emissions.

However, several factors complicate the analysis of AI's impact on GHG emissions from the oil and gas sector.

First, natural gas replaces coal in many places, with cheaper natural gas leading to less coal use. Natural gas produces roughly half the GHG emissions per unit of energy as coal when burned, so more natural gas use and less coal use can reduce GHG emissions—although only if natural gas leaks are kept to a minimum. Thus, while cheaper natural gas production due to AI creates significant risks of higher GHG emissions, there are scenarios in which it could do the opposite. The results will depend on a number of factors that vary by location.

Second, the global oil market is not a classic competitive market. Prices are determined in substantial part by the decisions of key producers (including in particular the Kingdom of Saudi Arabia), who adjust supply with the goal of keeping prices within ranges they consider desirable. In the partially managed global oil market, lower production costs enabled by AI may lead to lower prices and greater consumption but less directly and immediately than in more competitive markets.

Third, AI is also used in the oil and gas industry to help reduce GHG emissions. AI is helping to detect and control methane leaks, improve carbon capture processes and address supply chain emissions. Although these efforts appear to be smaller in scale than the use of AI to enhance oil and gas production, they have the potential to offset some of the GHG emissions increases from AI use in the industry.[78-80]

The bottom-line is that use of AI in the oil and gas sector has the potential to both increase and decrease GHG emissions. AI appears likely to be increasing GHG emissions from the oil and gas sector on a net basis, but a confident assessment requires more rigorous analysis.

## D. Future Greenhouse Gas (GHG) Emissions from AI

Future GHG emissions from AI are highly uncertain. AI has the potential to increase or decrease GHG emissions in the years ahead, in amounts that could be small or significant. The results will depend on a range of policy and investment decisions.

In the short-term, the surging demand for AI seems likely to increase GHG emissions.

- Although major data center operators would like to buy 100% low-carbon power, new data center demand exceeds the supply of low-carbon power in many locations. Growing demand for data center use, driven in part by AI, has led to deferral of some coal plant retirements in the US[17,81] and to construction of new natural gas plants in several locations, including Dublin and Phoenix.[82,83]
- Decarbonization of the processes and industries central to AI upstream emissions—including manufacturing silicon chips, steel and cement—is moving slowly.[84,85]
- Adoption of emissions-reducing applications of AI may not keep pace with increases in AI operational emissions and AI upstream emissions (although data on this topic are sparse).

In the medium- to long-term, AI could increase or decrease GHG emissions. While AI operational emissions and AI upstream emissions may both grow, AI will also be deployed in countless ways to accelerate decarbonization and reduce emissions. (See Chapters 3–13 of this Roadmap.) The net impact of AI on GHG emissions is uncertain.

A few studies have estimated future GHG emissions from AI.

- In a 2024 report, Morgan Stanley projected that $CO_2$ emissions from generative AI will reach 0.2–0.3% of global power sector $CO_2$ emissions (which is 0.1–0.15% of global $CO_2$ emissions) in 2027. Morgan Stanley said it expects the "net sustainability benefits from GenAI to be positive" (Morgan Stanley, 2024[38] at p. 4).
- In a 2021 study, BCG experts estimated that AI could reduce 5–10% of global GHG emissions by 2030, based on experiences with BCG clients.[63,86]

Several other studies have estimated future GHG emissions from data centers (including GHG emissions from data center operations unrelated to AI).

- In a 2024 report, Goldman Sachs found that "carbon dioxide emissions of data centers may more than double between 2022 and 2030."[87]
- In a 2024 blog post, International Monetary Fund (IMF) experts projected that $CO_2$ emissions from data centers could reach 0.5% of the global total by 2027.[88]

Future GHG emission from AI will be the sum of (1) AI operational emissions, (2) AI upstream emissions and (3) the GHG emissions impacts of AI applications (which could be positive or negative). The uncertainty with respect to each of these categories is significant. We consider each of them—as well as future demand for AI—below.

### i. AI operational emissions

Emissions from computing operations for AI in the years ahead will be a function of (1) improvements in the energy efficiency of AI hardware, (2) improvements in the energy efficiency of AI software, (3) rebound effects from these improvements and (4) the percentage of computing operations powered by new low-carbon sources. There is considerable uncertainty with respect to all these factors.

### a) Hardware efficiency

The energy efficiency of AI equipment has improved significantly in the past decade. This trend continues today and is likely to continue in the future. However, predicting the precise pace of improvements in the energy efficiency of AI equipment is challenging.

Some recent improvements in energy efficiency have been dramatic. Between 2015 and 2021, for example, data center workload increased by 260% while data center energy use increased by only 10%.[15,89]

Similar improvements continue today. NVIDIA's new Blackwell GPU trains large AI models with roughly 25% of the power needed for comparable tasks by older GPUs.[90,91] NVIDIA reports an astounding 45,000x improvement in the energy efficiency of their GPUs running LLMs in the past eight years.[91] In 2020, average power use effectiveness (PUE) across the industry was 1.58. (PUE is the ratio of total energy use at a data center to the energy used by its computing equipment.) Newer data centers have demonstrated PUEs of 1.1.[61,92-96]

These improvements in energy efficiency are likely to continue. Miniaturization and architectural optimization will likely drive continued energy efficiency in GPUs in the years ahead.[90,91,97] More efficient and higher-performing computational equipment, such as tensor processing units (TPUs), also offer the promise of continued improvements in energy efficiency.[92-94] More radical design concepts, such as analog-AI chips, may also result in major improvements in energy efficiency.[98] Studies of PUE at data centers suggest continued energy-efficiency improvements are possible.[61,92-96]

Yet predicting the pace at which the energy efficiency of AI equipment will improve is challenging. Hardware advances, such as new chip architectures, often follow unpredictable innovation cycles, making it difficult to forecast specific gains. Breakthroughs in quantum computing, neuromorphic

chips or AI itself could drastically improve efficiency. Supply chain disruptions or geopolitical forces could slow innovation. Significant energy efficiency gains in AI equipment are likely, but precise projections are challenging.

### b) Software efficiency

Advances in AI models have significantly improved the energy efficiency of AI in recent years. These advances include development of more efficient algorithms, such as sparse models and pruning techniques, which reduce the number of computations required to achieve the same or better results. Optimization strategies like quantization and knowledge distillation have also enabled AI models to run more efficiently on existing hardware. As a result, AI systems now require less computational power and energy to perform complex tasks, reducing their overall carbon footprint.[99,100]

Significant work is underway to further improve model architectures using these techniques and others.[92,101] Nodal and clustering optimization could have significant impacts on the overall carbon intensity of compute-heavy parts of an AI model's lifecycle. Researchers across major markets (e.g., the United States and China) have begun to investigate this potential, but more analysis is needed as new hardware becomes available.[102]

As with hardware efficiency improvements, projecting the pace of change in software development is challenging. The development of new algorithms and optimization techniques is inherently uncertain, as breakthroughs in AI often come from unexpected research directions and can be difficult to foresee.

The International Standards Organization (ISO) recently published a methodology for evaluating a software system's "software carbon intensity (SCI)." The methodology is intended to "help software practitioners make better, evidence-based decisions during system design, development, and deployment, that will ultimately minimize carbon emissions."[103] Widespread attention to the SCI methodology could help reduce emissions from AI systems.

### c) Rebound effects

In combination, the hardware and software energy advances described above offer the potential for significant—indeed extraordinary—improvement in the energy efficiency of AI in the years ahead. Whether these energy efficiency gains will have a significant impact on GHG emissions from AI is uncertain.

A core challenge in projecting GHG emissions from AI is the rebound effect (sometimes called "Jevons Paradox").[104,105] As AI tools become more energy efficient and therefore cost less, use cases for AI will expand. The power demand for AI from these new use cases could offset the energy savings from hardware and software energy efficiency improvements in part or in whole.

The rebound effect is a well-studied phenomenon in other contexts, including automotive fuel efficiency standards, where the rebound effect is estimated to offset 10–30% of a fuel efficiency standard's benefits.[106-108] A 2014 paper in the *American Economic Journal* by Lucas Davis et al. found significant rebound effects in Mexican programs to replace energy inefficient air conditioners and refrigerators.[109]

There is little research on the likely rebound effect as the energy efficiency of AI hardware and software improves in years ahead. Yet general trends in the industry suggest rebound effects may be significant. As significant energy efficiency improvements in the latest generation of GPUs were being announced in 2024, commercial orders for those GPUs skyrocketed and applications for new data center capacity continued to climb. A wide range of industry participants appear to believe that cheaper and more efficient computing power will open up new potential applications for AI, not cut back on overall power demand from the industry.[39,57]

### d) Low-carbon power

The amount of GHG emissions from AI operations in the years ahead will be determined in significant part by the amount of low-carbon power used for these operations.

Many large data center operators are deeply committed to using low-carbon power. Indeed the world's largest data center operators—Amazon, Microsoft, Google and Meta—are among the world's largest purchasers of renewable power.[1,110-112] However data center operators face significant constraints in procuring sufficient low-carbon power. Permitting delays, inadequate transmission infrastructure and land-use constraints are among the major barriers.[35]

These constraints complicate forecasting. The amount of GHG emissions from AI operations depends not just on the pace at which power demand for AI grows, but on how that power is generated. A data center or edge device powered by a grid with significant coal generation will emit far more GHGs than a data center co-located with a new low-carbon power plant.

The indirect effects of data center operators purchasing low-carbon power are also a complicating factor. If the supply of low-carbon power in a region is constrained, the purchase of low-carbon power by a data center operator may force other electricity consumers to purchase power from higher-carbon sources, indirectly increasing GHG emissions. This may currently be happening in the eastern United States.[113,114]

(Similar concerns have been raised with respect to hydrogen produced with renewable power, known as "green hydrogen." The European Union and United States have both adopted rules requiring that green hydrogen facilities use new or additional renewable power in order to receive favorable regulatory or tax treatment. There are proposals that data centers be subject to similar additionality requirements.)[115-118]

A potential solution to the problem of indirect GHG emissions increases is for data center operators to develop new low-carbon power sources for new data centers. One innovative approach is the Clean Transition Tariff developed by Google and others, in which utility regulators establish a rate structure under which data centers and other large customers pay more for new low-carbon power projects using emerging clean energy technologies.[119,120]

Another important development is the emergence of "carbon-aware computing," which schedules intensive computing tasks based on the carbon intensity of the power available to perform the computation.[121] By leveraging near-real-time data and models about renewable generation, a carbon-aware computing system can defer intensive, non-urgent AI model training tasks for time periods when renewable generation is abundant or curtailed. Intensive computing tasks could also be transferred to data centers in different locations where low-carbon electricity is available (taking into account the emissions associated with the data transfer).[122-124]

The strong commitment of leading data center operators to buying low-carbon power will help minimize the growth of GHG emissions in connection with AI in the years ahead. But there are constraints on the ability of data center operators to buy low-carbon power. Projections of low-carbon power's role in AI computing operations in the years ahead should allow for a range of possible outcomes.

### ii. AI upstream emissions

Upstream emissions from AI include emissions from manufacturing silicon chips, making steel and cement for data centers, and taking other steps necessary to build the physical infrastructure for AI operations. Many of these activities rely heavily on fossil fuel combustion and have significant GHG footprints. Future upstream emissions from AI will depend on growth in demand for AI and the pace at which these activities decarbonize.

Progress in decarbonizing some of these activities has been slow. Some forms of silicon production have a higher carbon footprint today than 20 years ago.[125] Steel and cement making are often considered "hard-to-abate" sectors, which are difficult to decarbonize.[126] (Fortunately AI could help accelerate decarbonization of some of these sectors. See Chapter 5 of this Roadmap.) The prospects for decarbonizing many of these sectors faster than AI scales may not be good, suggesting that upstream GHG emissions from AI may rise in the years ahead. However much more research is needed to make confident projections on this topic.

### iii. Emissions impacts of AI applications

In the years ahead, the impacts of AI applications on GHG emissions could be positive or negative. Indeed, these impacts could be very positive or very negative. The range of uncertainty is enormous.

As noted in Section C (iii) above, the phrase "impacts of AI applications on GHG emissions" is potentially confusing. In this context, it means how use of AI impacts GHG emissions, not including AI operational emissions or AI upstream emissions. For example, when a municipality uses AI tools to help with traffic management, how much do vehicle emissions fall? When an industrial facility uses AI tools in its operations, how much do emissions at that facility rise or fall?

A few studies have attempted to project the potential emissions benefits of AI applications in the years ahead.

- A 2023 report by BCG and Google found that "AI has the potential to unlock insights that could help mitigate 5–10% of GHG emissions by 2030"[127]
- A 2021 Capgeminii study found that executives interviewed believed AI could reduce overall GHG emissions 16% by 2024–2026[64]
- A 2019 report by PricewaterhouseCoopers (PwC)/Microsoft found that AI could reduce global GHG emissions by 1.5–4% by 2030 compared to business-as-usual pathways

However, the literature on this topic is sparse, and challenges in making projections are considerable. Data with respect to the impacts of AI applications on GHG emissions are limited. Evaluating the benefits of AI applications involves considering a counterfactual—what would happen in the same setting without AI? Such counterfactuals are often difficult to define with rigor. The potential for rebound effects from efficiencies introduced by AI creates analytic difficulties. Finally, AI is a transformational technology at early stages of development. Confidently predicting its capabilities or how it will be deployed beyond the short-term is difficult at best.

The dozens of AI applications discussed in this roadmap highlight the enormous potential for AI applications to reduce GHG emissions in the years and decades ahead. Some of these reductions are likely to be incremental—gains of perhaps 10–20% through improved operations. Other reductions could be transformational—such as dramatically reducing GHG emissions by discovering novel materials. At the same time, using AI in carbon-intensive industries could significantly increase emissions, if AI helps carbon-emitting activities become cheaper or more competitive.

### iv. Demand for AI

The pace of AI demand growth will help determine future GHG emissions in all three of the categories discussed above (AI operational emissions, AI upstream emissions and the emissions impacts of AI application). Demand for AI has been growing quickly for the past decade and is surging today. Private sector investment in AI grew 18x between 2013 and 2021,[128] and private sector demand for AI more than doubled from 2017 to 2022.[129] With the explosion of interest in AI following the release of ChatGPT in November 2022, demand for AI began to grow even faster. Many forecasters predict that AI will grow dramatically in the years ahead—at compound annual growth rates in the range of 30–35% or more.[11,53-55]

However, the pace at which demand for AI grows in the years and decades ahead is very uncertain. Some analysts question whether AI will deliver productivity benefits consistent with the enormous current investments in the technology,[130] suggesting that projections of rapid demand growth could be overstated. Regulatory frameworks, public attitudes, economic conditions, technology development and geopolitical trends will all shape demand growth. AI is a transformational, general-use technology at an early stage of adoption in most sectors. High growth rates are likely, but the range of uncertainty with respect to these rates is considerable.

## E. Conclusion

AI's impacts on GHG emissions could be positive or negative, both today and in the years ahead. Estimating with precision is challenging due to limited data and other challenges.[4]

However, there is significant potential for the overall GHG benefits of AI to exceed its costs. This could happen if (1) some of the emissions-reducing applications of AI discussed in this Roadmap deliver significant results and (2) AI operational emissions and AI upstream emissions grow slowly or fall in the years ahead. However, the opposite result is possible as well: AI applications could fail to reduce GHG emissions and AI operational emissions and AI upstream emissions could climb in the years ahead.

Supportive policies and commitment on the part of key stakeholders are needed to realize the full potential of AI to reduce GHG emissions.

## F. Recommendations

1. *AI developers, data center owners, energy experts, GHG emissions experts and standards organizations should establish robust methodologies and standards for reporting energy use and GHG emissions across the AI value chain.*

2. *AI developers and data center owners should report energy use and GHG emissions associated with their AI workloads.*

3. *Governments should adopt regulations that require AI developers and data centers owners to report their energy use and GHG emissions.*

4. *AI developers should take steps to reduce the carbon intensity of their models, using the ISO's methodology for evaluating their models' Software Carbon Intensity (SCI).*[103]

5. *Data center owners should prioritize adoption of energy-efficient hardware for AI operations and optimize AI workloads based on carbon-aware computing strategies.*

6. *Governments should promote and support policies that enable and incentivize data center owners to purchase low-carbon energy, including supporting new low-carbon power generation and grid expansion in regions with high concentrations of AI-driven data center growth.*

7. *National governments, AI developers, data center owners and philanthropies should fund researchers to develop a set of scenarios to quantify the effects that AI could have on greenhouse gas emissions under a range of assumptions. These scenarios should combine quantitative models with expert consultations, rigorously exploring a range of possible futures. The Intergovernmental Panel on Climate Change (IPCC) should include these scenarios in a special report on AI to be released within two years.*[5]

8. *All stakeholders should review and consider the dozens of other recommendations throughout this Roadmap to help reduce GHG emissions using AI tools.*

## G. References


1 Oktavia Catsaros. *Corporate Clean Power Buying Grew 12% to New Record in 2023, According to BloombergNEF;* BloombergNEF (BNEF), London, UK, https://about.bnef.com/blog/corporate-clean-power-buying-grew-12-to-new-record-in-2023-according-to-bloombergnef/ (2024).

2 Eric Masanet, Nuoa Lei & Jonathan Koomey. To better understand AI's growing energy use, analysts need a data revolution. *Joule* https://doi.org/10.1016/j.joule.2024.07.018.

3 Lynn H. Kaack *et al.* Aligning artificial intelligence with climate change mitigation. *Nature Climate Change* 12, 518-527 (2022). https://doi.org/10.1038/s41558-022-01377-7.

4 Peter Henderson *et al.* Towards the systematic reporting of the energy and carbon footprints of machine learning. *J. Mach. Learn. Res.* 21, Article 248 (2020). https://doi.org/10.48550/arXiv.2002.05651.

5 Amy Luers *et al.* Will AI accelerate or delay the race to net-zero emissions? *Nature* 628(8009), 718–720 (2024). https://doi.org/10.1038/d41586-024-01137-x.

6 Jordan Aljbour, Tom Wilson & Poorvi Patel. *Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption* Electric Power Research Institute (EPRI), Palo Alto, California, https://www.epri.com/research/products/000000003002028905 (2024).

7 Statista. *Share of worldwide hyperscale data center capacity in 2nd quarter 2022, by region* New York, New York, https://www.statista.com/statistics/1350992/global-hyperscale-data-center-capacity/ (Accessed August 2024).

8 Shandra Earney. *The Advantages and Disadvantages of Edge AI Simply Explained;* Xailient, Dover, Delaware, https://xailient.com/blog/the-advantages-and-disadvantages-of-edge-ai-simply-explained/ (2022).

9 Wevolver. Edge AI Market Analysis and Trends in *2024 State of Edge AI Report* Ch. 1, (Wevolver, Amsterdam, North Holland, 2024, https://www.wevolver.com/article/2024-state-of-edge-ai-report/edge-ai-market-analysis-and-trends).

10 Wevolver. Edge AI Challenges and Real-World Mitigations in *2024 State of Edge AI Report* Ch. 10, (Wevolver, Amsterdam, North Holland, 2024, https://www.wevolver.com/article/2024-state-of-edge-ai-report/edge-ai-challenges-and-real-world-mitigations).

11 Waldemar Szlezak & Andrew Peisch. *Data Centers: The Hubs of Digital Infrastructure;* Kohlberg Kravis Roberts & Co. L.P. (KKR), New York, New York, https://www.kkr.com/insights/hubs-digital-infrastructure#:~:text=Two%20megatrends%20are%20driving%20the,of%20artificial%20intelligence%20(AI) (2024).

12 Fengrong Li *et al. Current Power Trends and Implications for the Data Center Industry;* FTI Consulting, Washington, D.C., https://www.fticonsulting.com/insights/articles/current-power-trends-implications-data-center-industry (2024).

13 Jordan Aljbour, Tom Wilson & Poorvi Patel. *Executive Summary – Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption;* Electric Power Research Institute (EPRI), Palo Alto, California, https://www.epri.com/research/products/000000003002028905 (2024).

14 David Patterson *et al.* Energy and Emissions of Machine Learning on Smartphones vs. the Cloud. *Commun. ACM* 67, 86–97 (2024). https://doi.org/10.1145/3624719.

15 Goldman Sachs. *AI is poised to drive 160% increase in data center power demand;* New York, New York, https://www.goldmansachs.com/insights/articles/AI-poised-to-drive-160-increase-in-power-demand (2024).

16 Sayumi Take. *Boom in data centers challenges clean power goals in Asia;* Nikkei Asia, Tokyo, Japan, https://asia.nikkei.com/Business/Energy/Boom-in-data-centers-challenges-clean-power-goals-in-Asia (2024).

17 Antonio Olivo. *Internet data centers are fueling drive to old power source: Coal;* The Washington Post, Washington, D.C., https://www.washingtonpost.com/business/interactive/2024/data-centers-internet-power-source-coal/ (2024).

18 Melody Petersen. *Power-hungry AI data centers are raising electric bills and blackout risk;* Los Angeles Times, Los Angeles, California, https://www.latimes.com/environment/story/2024-08-12/california-data-centers-could-derail-clean-energy-goals (2024).

19 International Energy Agency (IEA). *Electricity Mid-Year Update – July 2024;* Paris, France, https://www.iea.org/reports/electricity-mid-year-update-july-2024Electricity Mid-Year Update - July 2024 (2024).

20 Carly Davenport *et al. Generational Growth: AI, data centers and the coming US power demand surge;* Goldman Sachs Group, Inc., New York, New York, https://www.goldmansachs.com/insights/goldman-sachs-research/generational-growth-ai-data-centers-and-the-coming-us-power-demand-surge (2024).

21 Nanako Takehara. *OCCTO Panel: Communications-Related Electricity Demand Expected to Increase Sharply to 2050;* Japan Atomic Industrial Forum, Inc (JAIF), Tokyo, Japan, https://www.jaif.or.jp/en/news/7022 (2024).

22 Jefferies' Equity Research Team. *How Data Centers Are Shaping the Future of Energy Consumption;* Jeffries Global Research & Strategy, New York, New York, https://insights.jefferies.com/the-big-picture/how-data-centers-are-shaping-the-future-of-energy-consumption (2024).

23 European Commission: Joint Research Centre, George Kamiya & Paolo Bertoldi. *Energy consumption in data centres and broadband communication networks in the EU;* Publications Office of the European Union, Luxembourg, Luxembourg, https://data.europa.eu/doi/10.2760/706491 (2024).

24 International Energy Agency (IEA). *What the data centre and AI boom could mean for the energy sector;* Paris, France, https://www.iea.org/commentaries/what-the-data-centre-and-ai-boom-could-mean-for-the-energy-sector (2024).

25 Dominion Energy. *Powering Your Every Day – 2023 Annual Report;* Richmond, Virginia, https://s2.q4cdn.com/510812146/files/doc_downloads/2024/2024/03/21/Dominion-Energy-2023-Annual-Report-and-Annual-Report-on-Form-10-K-1.pdf (2023).

26 Eoin Burke-Kennedy. *Data centres now account for 21% of all electricity consumption;* The Irish Times, Dublin, Ireland, https://www.irishtimes.com/business/2024/07/23/electricity-consumption-by-data-centres-rises-to-21-eclipsing-urban-households/ (2024).

27 National University of Singapore (NUS) College of Design and Engineering. *Taking a byte out of data centre emissions;* Singapore, https://cde.nus.edu.sg/taking-a-byte-out-of-data-centre-emissions/Taking a byte out of data centre emissions (2024).

28 Mordor Intelligence. *Europe Data Center Power Market Size & Share Analysis - Growth Trends & Forecasts (2024 - 2029);* Hyderabad, India, https://www.mordorintelligence.com/industry-reports/europe-data-center-power-market (2024).

29 Lydia Brissy *et al. Spotlight: European Data Centre – 2024;* Savills, London, UK, https://www.savills.com/research_articles/255800/362604-0 (2024).

30 Xuemei Xie, Yuhang Han & Hao Tan. Greening China's digital economy: exploring the contribution of the East–West Computing Resources Transmission Project to CO2 reduction. *Humanities and Social Sciences Communications* 11, 466 (2024). https://doi.org/10.1057/s41599-024-02963-0.

31 Aaron Tilley & Yang Jie. *Apple Is Developing AI Chips for Data Centers, Seeking Edge in Arms Race;* The Wall Street Journal, New York, New York, https://www.wsj.com/tech/ai/apple-is-developing-ai-chips-for-data-centers-seeking-edge-in-arms-race-0bedd2b2 (2024).

32 International Energy Agency (IEA). *World Energy Outlook 2024;* Paris, France, https://www.iea.org/reports/world-energy-outlook-2024 (2024).

33 John D. Wilson & Zach Zimmerman. *The Era of Flat Power Demand is Over (at slide 10);* Grid Strategies, LLC (sponsored by Clean Grid Initiative), Bethesda, Maryland, https://gridstrategiesllc.com/wp-content/uploads/2023/12/National-Load-Growth-Report-2023.pdf (2023).

34 Shayle Kann. Under the hood of data center power demand in *Catalyst. (Latitute Media, 2024) https://podcasts.apple.com/lu/podcast/under-the-hood-of-data-center-power-demand/id1593204897?i=1000658881876.*

35 Vijay V. Vaitheeswaran. *Big tech's great AI power grab;* The Economist, London, England, https://www.economist.com/business/2024/05/05/big-techs-great-ai-power-grab (2024).

36 International Energy Agency (IEA). *Electricity;* Paris, France, https://www.iea.org/reports/electricity-2024 (2024).

37 Dylan Patel, Daniel Nishball & Jeremie Eliahou Ontiveros. *AI Datacenter Energy Dilemma - Race for AI Datacenter Space;* SemiAnalysis, https://www.semianalysis.com/p/ai-datacenter-energy-dilemma-race (2024).

38 Morgan Stanley. *Powering the AI Revolution;* New York, New York, https://www.morganstanley.com/ideas/ai-energy-demand-infrastructure (2024).

39 Vivian Lee *et al. U.S. Data Center Power Outlook: Balancing competing power consumption needs;* Boston Consulting Group (BCG) on LinkedIn, Sunnyvale, CA, https://www.linkedin.com/pulse/us-data-center-power-outlook-balancing-competing-consumption-lee-iz4pe/ (2024).

40 Srini Bangalore *et al. Investing in the rising data center economy;* McKinsey & Company – Technology, Media & Telecommunications Practice, New York, New York, https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/investing-in-the-rising-data-center-economy (2023).

41 Taiba Jafari *et al. Projecting the Electricity Demand Growth of Generative AI Large Language Models in the US;* Columbia University School of International and Public Affairs (SIPA) Center on Global Energy Policy, New York, New York, https://www.energypolicy.columbia.edu/projecting-the-electricity-demand-growth-of-generative-ai-large-language-models-in-the-us/ (2024).

42 China Daily. *Green data centers in focus;* The State Council of the People's Republic of China, Beijing, China, https://english.www.gov.cn/statecouncil/ministries/202112/09/content_WS61b13edac6d09c94e48a1f81.html#:%7E:text=According%20to%20a%20report%20from%20China%27s%20State%20Grid,for%202.7%20percent%20of%20the%20country%27s%20electricity%20consumption. (2021).

43 China Com-Service. *White Paper;* https://aimg8.dlssyht.cn/u/551001/ueditor/file/276/551001/1684888884683143.pdf (2023).

44 Katya Golubkova, Chang-Ran Kim & Sonali Paul. *Japan sees need for sharp hike in power output by 2050 to meet demand from AI, chip plants;* Reuters, Toronto, Canada, https://www.reuters.com/business/energy/japan-sees-need-sharp-hike-power-output-by-2050-meet-demand-ai-chip-plants-2024-05-14/ (2024).

45 Arizona Public Service Company (APS). *2023 Integrated Resource Plan;* Phoenix, Arizona, https://www.aps.com/-/media/APS/APSCOM-PDFs/About/Our-Company/Doing-business-with-us/Resource-Planning-and-Management/APS_IRP_2023_PUBLIC.pdf?la=en&sc_lang=en&hash=DF34B49033ED43FF0217FC2F93A0BBE6 (2023).

46 CBRE (Coldwell Banker Richard Ellis). *Global Data Center Trends 2024;* Dallas, Texas, https://www.cbre.com/insights/reports/global-data-center-trends-2024 (2024).

47 Matthew Connatser. *Ohio power plants want special tariffs on datacenters to protect regional grid;* The Register, London, England, https://www.theregister.com/2024/05/21/ohio_datacenter_power/ (2024).

48 Alexandra Sasha Luccioni, Sylvain Viguier & Anne-Laure Ligozat. Estimating the carbon footprint of bloom, a 176b parameter language model. *arXiv preprint arXiv:2211.02001* (2022). https://doi.org/10.48550/arXiv.2211.02001.

49 David Patterson *et al.* Carbon emissions and large neural network training. *arXiv preprint arXiv:2104.10350* (2021). https://doi.org/10.48550/arXiv.2104.10350.

50 Hannah Ritchie, Pablo Rosado & Max Roser. *Greenhouse gas emissions;* Our World in Data, Oxford, England, https://ourworldindata.org/greenhouse-gas-emissions (2020 (updated 2024)).

51 International Energy Agency (IEA). *Data Centres and Data Transmission Networks;* https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks (Accessed September 2023).

52 Charlotte Freitag *et al.* The real climate and transformative impact of ICT: A critique of estimates, trends, and regulations. *Patterns* 2, 100340 (2021). https://doi.org/10.1016/j.patter.2021.100340.

53 Grand View Research. *Artificial Intelligence Market Size & Trends;* San Francisco, California, https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-market (2024).

54 Markets and Markets Research Private Ltd. *Artificial Intelligence (AI) Market;* Hadapsar, India, https://www.marketsandmarkets.com/Market-Reports/artificial-intelligence-market-74851580.html (Accessed August 2024).

55 Bergur Thormundsson. *Artificial intelligence (AI) market size worldwide from 2020 to 2030;* Statista, New York, New York, https://www.statista.com/forecasts/1474143/global-ai-market-size#:~:text=AI%20market%20size%20worldwide%20from,2030%20(in%20billion%20U.S.%20dollars)&text=The%20market%20for%20artificial%20intelligence,billion%20U.S.%20dollars%20in%202030 (2024).

56 THe Salata Institute for Climate and Sustainability. *Cutting the carbon footprint of future computer chips;* Harvard University, Boston, Massachusetts, https://salatainstitute.harvard.edu/cutting-the-carbon-footprint-of-future-computer-chips/ (2024).

57 Gauthier Roussilhe, Anne-Laure Ligozat & Sophie Quinton. A long road ahead: a review of the state of knowledge of the environmental effects of digitization. *Current Opinion in Environmental Sustainability* 62, 101296 (2023). https://doi.org/10.1016/j.cosust.2023.101296.

58 Greenhouse Gas Protocol. *FAQ;* World Resources Institute, Washington, D.C., https://ghgprotocol.org/sites/default/files/2022-12/FAQ.pdf (Accessed August 2024).

59 Google. *2024 Environmental Report;* Mountain View, California, https://sustainability.google/reports/google-2024-environmental-report/ (2024).

60 Microsoft. *2024 Environmental Sustainability Report;* Redmond, Washington, https://www.microsoft.com/en-us/corporate-responsibility/sustainability/report (2024).

61 Carole-Jean Wu *et al.* Beyond Efficiency: Scaling AI Sustainably. arXiv:2406.05303 (2024). https://ui.adsabs.harvard.edu/abs/2024arXiv240605303W.

62 Maxime Pelcat. GHG emissions of semiconductor manufacturing in 2021. (Univ Rennes, INSA Rennes, CNRS, IETR – UMR 6164, F-35000 Rennes (hal-04112708v1), 2023; https://hal.science/hal-04112708v1/file/semiconductor_ghg.pdf).

63 Charlotte Degot *et al. Reduce Carbon and Costs with the Power of AI;* Boston Consulting Group (BCG), Boston, Massachusetts, https://www.bcg.com/publications/2021/ai-to-reduce-carbon-emissions (2021).

64 Anne-Laure Thieullent *et al. Climate AI: How artificial intelligence can power your climate action strategy;* Capgemini Research Institute, Paris, France, https://www.capgemini.com/insights/research-library/climate-ai/ (2020).

65 Gaurav Sharma. *How Multibillion Dollar Investments In AI Are Driving Oil And Gas Sector Innovation;* Forbes Media, Jersey City, New Jersey, https://www.forbes.com/sites/gauravsharma/2023/08/14/how-multibillion-dollar-investments-in-ai-are-driving-oil-and-gas-sector-innovation/ (2023).

66 Sudeep Srivastava. *Unleashing the Potential of Artificial Intelligence in the Oil and Gas Industry – 10 Use Cases, Benefits, Examples;* Appinventiv, Noida, Uttar Pradesh, India, https://appinventiv.com/blog/artificial-intelligence-in-oil-and-gas-industry/ (2024).

67 Karen Hao. *Microsoft's Hypocrisy on AI;* The Atlantic, Washington, D.C., https://www.theatlantic.com/technology/archive/2024/09/microsoft-ai-oil-contracts/679804/ (2024).

68 Bill Tomlinson *et al.* The carbon emissions of writing and illustrating are lower for AI than for humans. *Scientific Reports* 14, 3732 (2024). https://doi.org/10.1038/s41598-024-54271-x.

69 Renee Cho. *AI's Growing Carbon Footprint;* Columbia Climate School: State of the Planet, https://news.climate.columbia.edu/2023/06/09/ais-growing-carbon-footprint/ (2023).

70 U. Gupta *et al.* "Chasing Carbon: The Elusive Environmental Footprint of Computing" in *2021 IEEE International Symposium on High-Performance Computer Architecture (HPCA)*, 854-867, https://www.doi.org/10.1109/HPCA51647.2021.00076,(2021).

71 Eric Masanet *et al.* Recalibrating global data center energy-use estimates. *Science* 367, 984-986 (2020). https://www.doi.org/10.1126/science.aba3758.

72 Dmitry Koroteev & Zeljko Tekic. Artificial intelligence in oil and gas upstream: Trends, challenges, and scenarios for the future. *Energy and AI* 3, 100041 (2021). https://doi.org/10.1016/j.egyai.2020.100041.

73 Adogioye Monigha. AI-Driven Cost Optimization in Oil and Gas Projects. *International Journal of Petroleum and Gas Engineering Research* 7(2), 17-32 (2024). https://eajournals.org/ijpger/vol-7-issue-2-2024/ai-driven-cost-optimization-in-oil-and-gas-projects/.

74 Abdulhamid Musa. Review Paper Revolutionizing Oil and Gas Industries with Artificial Intelligence Technology. *International Journal of Computer Sciences and Engineering* 11(5), 20–30 (2023). https://www.researchgate.net/publication/374380602_Review_Paper_Revolutionizing_Oil_and_Gas_Industries_with_Artificial_Intelligence_Technology.

75 C3.ai. *C3 AI Addresses Key Issues Across the Oil and Gas Value Chain;* Redwood City, California, https://c3.ai/industries/enterprise-ai-for-oil-and-gas/ (Accessed October 2024).

76 Lance Mortlock *et al. AI in oil and gas: Extracting collective value by humanizing data;* Ernst & Young LLP, London, UK, https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/oil-and-gas/ey-ai-in-oil-and-gas.pdf (2019).

77 Wassim Ghadban. *The journey of AI in the oil & gas industry;* Oil & Gas Middle East, Dubai, United Arab Emirates, https://www.oilandgasmiddleeast.com/business/insights/the-journey-of-ai-in-the-oil-gas-industry (2024).

78 Mike Lyons *et al. The AI Angle in Solving the Oil and Gas Emissions Challenge;* Boston Consulting Group (BCG), Boston, Massachusetts, https://www.bcg.com/publications/2021/ai-in-oil-and-gas-emissions-challenge (2021).

79 Jim Magill. *Oil, Gas Companies Deploy AI In The Fight To Reduce Carbon Emissions;* Forbes, Jersey City, New Jersey, https://www.forbes.com/sites/jimmagill/2021/03/28/oil-gas-companies-deploy-ai-in-the-fight-to-reduce-carbon-emissions/ (2021).

80 Charlie Fazzino *et al. Building a Lower-Carbon Future With HPC and AI in Energy;* NVIDIA, Santa Clara, California, https://resources.nvidia.com/en-us-upstream-energies/gtc24-s62121 (

81 Matthew Connatser. *Surging datacenter power demand slows the demise of US coal plants;* The Register, London, England, https://www.theregister.com/2024/05/31/datacenter_power_crunch/ (2024).

82 Peter Judge. *Microsoft granted permission to run its Dublin data center on gas;* Data Center Dynamics (DCD), London, UK, https://www.datacenterdynamics.com/en/news/microsoft-granted-permission-to-run-its-dublin-data-center-on-gas/ (2023).

83 Catherine Boudreau. *Phoenix plans to build gas plants to help power new data centers;* Business Insider, New York, New York, https://www.businessinsider.com/phoenix-expanding-its-natural-gas-plants-to-power-data-centers-2023-10 (2023).

84 Jan-Hinnerk Mohr *et al. For Chip Makers, the Decarbonization Challenge Lies Upstream;* Boston Consulting Group (BCG), Boston, Massachusetts, https://www.bcg.com/publications/2023/why-chip-makers-need-to-focus-on-the-upcoming-decarbonization-challenges (2023).

85 Canary Media. *The Tough Stuff: Decarbonizing steel, cement and chemicals;* New York, New York, https://www.canarymedia.com/the-tough-stuff-decarbonizing-steel-cement-and-chemicals (Updated 2023; Accessed August 2024).

86 Amane Dannouni *et al. Accelerating Climate Action with AI;* Boston Consulting Group (commissioned by Google), Boston, Massachusetts, https://web-assets.bcg.com/72/cf/b609ac3d4ac6829bae6fa88b8329/bcg-accelerating-climate-action-with-ai-nov-2023-rev.pdf (2023).

87 . (!!! INVALID CITATION !!! 15,75).

88 Shafik Hebous & Nate Vernon-Lin. *Carbon Emissions from AI and Crypto Are Surging and Tax Policy Can Help;* International Monetary Fund (IMF), Washington, D.C., 2024).

89 Ciaran Flanagan. *How Data Centers Are Driving The Renewable Energy Transition;* Forbes, https://www.forbes.com/sites/siemens-smart-infrastructure/2023/03/13/how-data-centers-are-driving-the-renewable-energy-transition/?sh=1396b0fb4214 (2023).

90 Shar Narasimhan. *Beyond 'Data-Driven': How Energy-Efficient Computing for AI Is Propelling Innovation and Savings Across Industries;* Nvidia – Blogs, Santa Clara, California, https://blogs.nvidia.com/blog/energy-efficient-ai-industries/ (2024).

91 Dion Harris. *Sustainable Strides: How AI and Accelerated Computing Are Driving Energy Efficiency;* Nvidia – Blogs, Santa Clara, California, https://blogs.nvidia.com/blog/accelerated-ai-energy-efficiency/ (2024).

92 Meredith Fowlie. *Data Centers Are Booming;* Energy Institute Blog by the Energy Institute at Haas (University of California (UC) Berkeley), https://energyathaas.wordpress.com/2023/10/09/data-centers-are-booming/ (2023).

93 D. Patterson *et al.* The Carbon Footprint of Machine Learning Training Will Plateau, Then Shrink. *Computer* 55, 18-28 (2022). https://www.doi.org/10.1109/MC.2022.3148714.

94 Ashish Dandekar. *8 power usage effectiveness (PUE) best practices for your data center;* Computer Weekly, https://www.computerweekly.com/tip/8-power-usage-effectiveness-PUE-best-practices-for-your-data-center (2011).

95 Colleen Josephson *et al.* The Sky is Not the Limit: Untapped Opportunities for Green Computing. *SIGENERGY Energy Inform. Rev.* 3, 33–39 (2023). https://doi.org/10.1145/3630614.3630621.

96 T. Hernandez *et al.* "Energy Efficiency in Data Center Management" in *2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE)*, 1214-1218, https://doi.org/10.1109/CSCE60160.2023.00203,(2023).

97 Nvidia. *Glossary: Power Efficiency;* Santa Clara, California, https://www.nvidia.com/en-us/glossary/power-efficiency/ (Accessed August 2024).

98 Antonella Gina Fleitas. *What Is Data Center PUE? Defining Power Usage Effectiveness;* Data Center Knowledge, https://www.datacenterknowledge.com/sustainability/what-data-center-pue-defining-power-usage-effectiveness (2023).

99 Torsten Hoefler *et al.* Sparsity in Deep Learning: Pruning and growth for efficient inference and training in neural networks. arXiv:2102.00554 (2021). https://doi.org/10.48550/arXiv.2102.00554.

100 Everton Gomede. *Sparse Neural Networks and Pruning: Trimming the Fat for Efficient Machine Learning;* Medium, San Fancisco, California, https://medium.com/aimonks/sparse-neural-networks-and-pruning-trimming-the-fat-for-efficient-machine-learning-5b9d2920c526 (2023).

101 S. Ambrogio *et al.* An analog-AI chip for energy-efficient speech recognition and transcription. *Nature* 620, 768-775 (2023). https://doi.org/10.1038/s41586-023-06337-5.

102 Wenli Ni *et al.* CO2 emission-mitigation pathways for China's data centers. *Resources, Conservation and Recycling* 202, 107383 (2024). https://doi.org/10.1016/j.resconrec.2023.107383.

103 International Organization for Standardization (ISO). *ISO/IEC 21031:2024 Information technology — Software Carbon Intensity (SCI) specification;* Geneva, Switzerland, https://www.iso.org/standard/86612.html (2024).

104 David Owen. *The Efficiency Dilemma (in Annals of Environmentalism);* The New Yorker, New York, New York, https://www.newyorker.com/magazine/2010/12/20/the-efficiency-dilemma?_sp=c1f8b808-3111-41ce-a832-08891ee99fa5.1723848270771 (2010).

105 Mario Giampietro & Kozo Mayumi. Unraveling the Complexity of the Jevons Paradox: The Link Between Innovation, Efficiency, and Sustainability. *Frontiers in Energy Research* 6 (2018). https://doi.org/10.3389/fenrg.2018.00026.

106 Ruiwen Lee, Gernot Wagner & Environmental Defense Fund. *The Rebound Effect in a More Fuel Efficient Transportation Sector (Working Paper No. 2012/3);* New York University School of Law – Institute for Policy Integrity, New York, New York, https://policyintegrity.org/files/publications/The_Rebound_Effect.pdf (2013).

107 Kenneth Gillingham, David Rapson & Gernot Wagner. *The Rebound Effect and Energy Efficiency Policy;* Yale School of the Environment, New Haven, Connecticut, https://resources.environment.yale.edu/gillingham/GillinghamRapsonWagner_Rebound.pdf (2015).

108 David Berreb. *As Use of A.I. Soars, So Does the Energy and Water It Requires;* Yale School of the Environment – Yale Environment 360, New Haven, Connecticut, https://e360.yale.edu/features/artificial-intelligence-climate-energy-emissions (2024).

109 Lucas W. Davis, Alan Fuchs & Paul Gertler. Cash for Coolers: Evaluating a Large-Scale Appliance Replacement Program in Mexico. *American Economic Journal: Economic Policy* 6, 207–238 (2014). http://dx.doi.org/10.1257/pol.6.4.207.

110 Amazon – Sustainability. *Carbon-Free Energy;* Seattle, Washington, https://sustainability.aboutamazon.com/climate-solutions/carbon-free-energy?energyType=true (Accessed August 2024).

111 Google Data Centers. *24/7 Carbon-Free Energy by 2030;* Council Bluff, Iowa, https://www.google.com/about/datacenters/cleanenergy/#:~:text=A%20sustainability%20moonshot,with%20purchases%20of%20renewable%20energy (Accessed August 2024).

112 Michelle Patron. *Advocating for decarbonization of the power sector;* Microsoft, Redmond, Washington, https://blogs.microsoft.com/on-the-issues/2023/08/16/apec-sustainability-decarbonization-fusion-energy/ (2023).

113 Jackson Morris. *Data Centers Gobbling Up Existing Nukes Threatens Grid Decarb Goals;* Natural Resources Defense Council (NRDC), New York, New York, https://www.nrdc.org/bio/jackson-morris/data-centers-gobbling-existing-nukes-threatens-grid-decarb-goals-0 (2024).

114 Jeff St. John. *Data centers want to tap existing nuclear power. Is that good or bad?;* Canary Media, New York, New York, https://www.canarymedia.com/articles/nuclear/data-centers-want-to-tap-existing-nuclear-power-is-that-good-or-bad (2024).

115 Directorate-General for Energy. *Renewable hydrogen production: new rules formally adopted;* European Commission, Brussels, Belgium, https://energy.ec.europa.eu/news/renewable-hydrogen-production-new-rules-formally-adopted-2023-06-20_en (2023).

116 US Department of Energy (DOE). *Clean Hydrogen Production Tax Credit (45V) Resources;* Washington, D.C., https://www.energy.gov/articles/clean-hydrogen-production-tax-credit-45v-resources (2023).

117 Brian Deese & Lisa Hansmann. *America Needs an Energy Policy for AI;* Heatmap News Inc., Larchmont, New York, https://heatmap.news/technology/ai-additionality-framework (2024).

118 Cy McGeady. *AI Additionality Is the Wrong Solution to a Real Problem;* Center for Strategic & International Studies, Washington, D.C., https://www.csis.org/analysis/ai-additionality-wrong-solution-real-problem (2024).

119 Lisa Martine Jenkins. *Can Google's clean transition tariff remake utility incentives?;* Latitude Media, Mount Pleasant, Michigan, https://www.latitudemedia.com/news/the-clean-transition-tariff-is-googles-play-to-remake-utility-incentives#:~:text=Dubbed%20the%20%E2%80%9Cclean%20transition%20tariff,passing%20costs%20on%20to%20ratepayers (2024).

120 Duke Energy. *Responding to growing demand, Duke Energy, Amazon, Google, Microsoft and Nucor execute agreements to accelerate clean energy option;* PR Newswire, New York, New York, https://www.prnewswire.com/news-releases/responding-to-growing-demand-duke-energy-amazon-google-microsoft-and-nucor-execute-agreements-to-accelerate-clean-energy-options-302158155.html (2024).

121 Philipp Wiesner *et al.* Vessim: A Testbed for Carbon-Aware Applications and Systems. arXiv:2306.09774 (2023). https://doi.org/10.48550/arXiv.2306.09774.

122 Green Software Foundation. *Carbon-Aware Computing Whitepaper - How UBS succeeded in measuring and reducing carbon emissions of their core risk platform;* San Francisco, California, https://greensoftware.foundation/articles/carbon-aware-computing-whitepaper-how-ubs-succeeded-in-measuring-and-reducing-car (2023).

123 A. Radovanović *et al.* Carbon-Aware Computing for Datacenters. *IEEE Transactions on Power Systems* 38, 1270-1280 (2023). https://doi.org/10.1109/TPWRS.2022.3173250.

124 Dan Benitah. *Green-Software-Foundation / carbon-aware-sdk* GitHub, San Francisco, California, https://github.com/Green-Software-Foundation/carbon-aware-sdk (Accessed August 2024).

125 Gudrun Sævarsdottir, Halvor Kvande & Thordur Magnusson. Greenhouse Gas Emissions from Silicon Production -Development of Carbon Footprint with Changing Energy Systems. *Proceedings of the 16th International Ferro-Alloys Congress (INFACON XVI)* (2021). http://dx.doi.org/10.2139/ssrn.3926088

126 Chris Bataille. Decarbonizing Steel and Cement in *Columbia Energy Exchange. (Columbia University School of International and Public Affairs – Center on Global Energy Policy, New York, New York, 2023) https://www.energypolicy.columbia.edu/decarbonizing-steel-and-cement/.*

127 Amane Dannouni *et al. How AI Can Speed Climate Action;* Boston Consulting Group (BCG), Boston, Massachusetts, https://www.bcg.com/publications/2023/how-ai-can-speedup-climate-action (2023).

128 Emma Chervek. *Hyperscaler sustainability scorecard: Rating Google's, Microsoft's and AWS' environmental strategies;* SDxCentral, https://www.sdxcentral.com/articles/analysis/hyperscaler-sustainability-scorecard-rating-google-microsoft-and-aws/2023/07/ (2023).

129 Quantum Black AI by McKinsey. *The state of AI in 2022—and a half decade in review;* https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review (2022).

130 Goldman Sachs. *Top of Mind – Gen AI: too much spend, too little benefit? (Report);* New York, New York, Report available at https://www.goldmansachs.com/insights/top-of-mind/gen-ai-too-much-spend-too-little-benefit (2024).