

How Al may help or harm the climate: A historical and sectoral review Emily Han

Abstract

As AI capabilities have accelerated over the past decade, so have the questions surrounding their environmental impact. This paper traces sector-specific perceptions of AI's climate effects throughout the past decade, focusing on three key eras of development from 2014. Drawing on peer-reviewed literature, corporate sustainability reports, and public publishing outlets, we developed a custom Sentiment Concern Index (SCI) to quantify shifts in optimism and concern across academia, industry, and publishing houses.

The findings suggest that while early academic and industrial discourse framed AI as a promising but untested tool, more recent years have seen both increased deployment and growing criticism, especially regarding the energy demands of large-scale models. Despite these concerns, the landscape is shifting toward "green AI," carbon-aware infrastructure, and environmentally responsible development practices. The paper concludes with a forward-looking discussion of integrated strategies, emphasizing the need for coordinated policy, technical innovation, public transparency, and cross-sector collaboration. As AI becomes further embedded in society, ensuring that it functions as a climate asset and not liability, will be one of the defining sustainability challenges of the coming decade.

Background

Over the past decade, artificial intelligence (AI) has evolved from a niche field into a central force shaping everyday life. Whether powering virtual assistants, automating vehicles, or enabling scientific research, AI systems have become deeply embedded in how we live, work, and solve problems. Simultaneously, the world has grappled with mounting climate pressures: global temperatures continue to rise, and extreme weather events grow more frequent, displacing families and leading to countless deaths.

This study investigates the double-edged role of AI in the context of climate change. On one hand, AI's increasingly sophisticated abilities hold promise as a powerful tool for climate modeling, energy grid optimization, and disaster prediction. On the other, its development, especially the training and deployment of large-scale models, demands significant amounts of energy. To assess whether AI is helping to mitigate or inadvertently worsening climate problems, we must understand not only how it functions, but also the infrastructure and systems that support it.

Climate change itself, as described by the United Nations, refers to "long-term shifts in weather patterns" caused primarily by human activities, especially the emission of greenhouse gases like



carbon dioxide and methane ("What Is"). These gases trap heat in the atmosphere, leading to a cascade of grim effects: hotter summers, prolonged droughts, more intense storms, and even disruptions to geological systems, such as increased volcanic activity. As AI becomes more energy-intensive, understanding its contribution to such emissions becomes critical.

This brings us to the concept of energy consumption. Every AI model, from a simple classifier to a complex language model, requires electricity to operate. Yet, quantifying that energy use is no simple task. Estimating the energy consumption of a model involves a detailed understanding of the hardware it runs on, how long it's trained, and the efficiency and energy sources of the facilities that house its servers (O'Donnell). This becomes even more complex once cloud-based models are taken into consideration, as everything from the time of day they are run to the geographical region can change the outcome, making it difficult to predict (Dodge et al., 2022). Thus, these estimates are often opaque, leaving a significant gap in our understanding of Al's true environmental impact.

A key part of this puzzle of energy consumption lies in the specialized hardware that powers AI. Modern models depend heavily on AI accelerators such as GPUs (Graphics Processing Units) and TPUs (Tensor Processing Units), which are designed to process massive volumes of data and perform high-level computations with speed and efficiency (Brodtkorb, et. al). While these accelerators enable groundbreaking advances—particularly in fields like Natural Language Processing (NLP), which involves teaching machines to understand and generate human language—they also draw substantial amounts of electricity.

To gauge how efficiently that electricity is used, many companies turn to a metric called Power Usage Effectiveness, or PUE (Yuventi). Widely used in the tech industry, including by firms like Google or Samsung ("Growing the Internet", n.d.), PUE measures how well a data center converts energy into useful computational work. It's calculated by dividing the total energy used by the facility by the energy used solely by its IT equipment. A perfect PUE of 1.0 means that every watt of electricity powers computing directly, with none lost to cooling or other overhead. In practice, PUE serves as a benchmark for energy efficiency and a target for sustainability improvements in AI infrastructure.

The development of AI can be understood in three key phases, each marking a distinct stage in its evolution. From 2014 to 2017 was the "pre-transformer era", which saw progress in deep learning, but lacked the architectural breakthroughs that would define the next wave. That wave came in 2017 with the introduction of transformers, ushering in what will be henceforth called the "conception of the transformer era," spanning from 2017 to 2022. During this time, AI models became more powerful, more specialized, and vastly more energy-intensive, as new models were generated from the original transformers model. Then, from 2022 onward, the field entered its current phase: the "post-conception era," where transformer-based models such as large



language models (LLMs) began to scale rapidly and reach mainstream adoption and widespread use through tools like AI chatbots and productivity assistants.

Looking at Al's evolution through the lens of these three periods is essential as it provides a clear structure for tracking the technological milestones that have shaped Al, such as the introduction of transformer models, which marked an immense turning point in both computational capacity and energy demands. This segmentation of the timeline also allows precise analysis of how energy consumption patterns, hardware advancements, and environmental impacts have shifted with each new wave of innovation. Furthermore, it allows for more detailed review of the literature, as the studies, both academic and industry, can be situated within their relevant context, and enables clearer research into how the public perceptions of Al have shifted as well. Ultimately, organizing Al's evolution into these distinct eras enables a clearer understanding of its dual role as both a driver of energy-intensive progress and a potential tool for climate mitigation.

Together, these concepts of climate change, energy systems, computational infrastructure, and an outline of Al's technical evolution frame the central question of this review: as Al becomes more developed and more widespread, is it exacerbating the very climate problems it might be capable of helping to solve?

Era #1 - 2014-2017 - Pre-transformer era

While this era did not see any of the shockingly rapid growth that subsequent eras did in the development or use of AI models, it includes the essential precursors that enable this advancement in the future.

One of the most essential turning points during this era was the introduction of GPUs for use in AI model training. Without them, using only the CPU, it would have been completely impossible to train any AI model to a useful level of functionality due to the gargantuan amount of computations that were required to train a model. This is because GPUs were much more specialized than any other existing processing unit, which is essential since specialization allows for incredible amounts of time reduction and energy efficiency in large-scale tasks like AI model training (Baji). This is why it is known as an AI accelerator.

Originally created for use in gaming graphics and 3D graphics rendering in the 1990's, used to output 2D images from a 3D world, GPUs were originally built to handle huge amounts of calculations, especially in contexts such as matrix multiplication (McClanahan, 2011). Due to this extraordinary ability in comparison to the powers of the CPU in that same area, they were slowly introduced to outside applications starting in 1993, and finally fully broke free of the graphics industry in the 2000's. By the end of 2017, nearly 28% of the global population had a



device that used a GPU, which demonstrates how useful it was and how it radicalized the industry standard, setting the stage for further Al accelerator development in future years.

This era also saw the monumental historical moment of the introduction of the transformer model. But first, it was preceded by the sequence-to-sequence model, or the Seq2Seq model, introduced in 2014, which was the first ever transformer-like model to exist at the time with the use of an encoder-decoder architecture, and providing the basis for the introduction of the transformer model (Yin & Wan, 2022; Sutskever et al., 2014). This model utilizes LSTM (long short-term memory), working from end-to-end, to read inputs and give outputs that do not have a fixed dimensionality (Sutskever). The capabilities of this model are surprisingly similar to those of current-day models, given that it can generate text, albeit to a lesser degree of human-likeness than today's models, and paraphrase or summarize long documents. In other words, it is especially capable of data-to-text generation (Yin), providing generated text based on a structured input. It was even proposed by some scientists in the field at the time to be used on social-media app algorithms to predict and recommend content to users (Torres et al., 2020).

Then, in 2017, the Transformer model was first introduced to the world through the paper "Attention is All You Need", mainly written by Google Brain or Research members. This was the first paper that revolutionized the idea of transformers (Vaswani). Transformers work according to an encoder-decoder structure. One of its distinguishing features is the self-attention mechanism, which essentially highlights the importance of certain words so that it can only pay attention to the most important ones. It considered not only the individual words as previous models had, but more so their relationship to others — for example, the word "the" held less meaning than the subject of the sentence, rather than the same importance. Thus, it was much more efficient, optimizing the number of operations necessary, and also that it was able to parallelize work, computing multiple inputs at once.

Era #2 - 2018-2022 - Conception of the transformer

After the introduction of the transformer in 2017, numerous different researchers and companies were sparked to develop their own AI model creations. There was also the introduction of pre-trained models in 2018, which made models all the more efficient and well-suited to perform specific tasks.

The overarching name for these models are foundation models, and they began their rise in development in 2021 (Bommasani. These foundation models were mainly used for Natural Language Processing, or NLP. They could be easily scaled at this point because of the well-developed hardware, the depth of research into transformers already existing, and increased databases that could be used to train large models more accurately. Even to this day, Google Search relies on foundation models to function.



The way foundation models work is that they are first trained on large amounts of data across hundreds of GPUs, which can sometimes take months. This can create large energy costs and thus carbon emissions, but it is only one time during the initial training process. These models were able to be more specialized, which improved their productivity. However, they also had some negative impacts, such as bias and inequality occurring, which can lead to legal concerns.

Among the most famous of these foundation models is the GPT model. The first instance, GPT-1, was built in 2018, created specifically in order to overcome the challenges faced by past transformer models on the account of a shortage of labeled data (Zong & Krishnamachari, 2022). This model by OpenAI was also able to perfectly recreate human speech patterns, as well as capture information from large, dense texts unsupervised, and complete tasks like text summarization, code generation, and story creation (Zhang). Its main strength was its ability to multitask, and its ability to learn to do tasks proficiently with very few data points. OpenAI continued on to create GPT models 2 and 3, which both developed and surpassed the previous models in its ability to both comprehend and produce language (McGuffie). The applications of GPT-3 stretched to include customer service chatbots and report summaries, and with the creation of the GPT-3 playground, enabled even inexperienced users to be able to train and create their own downstream models (Zong & Krishnamachari, 2022).

Another prevalent model is BERT, which was also created using the Transformer model as a basis, using attention layers to utilize the self-attention method (Han et al., 2021). It too was built in 2018, developed by Google Al Language, and also operates as a solution to multiple different problems, serving as a sort of "swiss army knife solution" to multiple commonly found tasks language models perform, including sentiment analysis, summarization, and text generation and prediction (Muller, 2022). BERT was made publicly available through open source code, meaning that the pre-trained model was able to be used for an even larger number of purposes as many developers could use BERT easily and quickly in their own specific models (Muller, 2022).

Models like BERT and GPT had a strength in that they were able to act as a backbone for more specific models with applications in more specialized sectors, due to the fact that they came pre-trained (Han et al., 2021).

The development of models like BERT, which were larger and more compute-intense models, requiring the training of billions of texts online, was made possible through the introduction of the TPU, a further specialized AI accelerator from the GPU (Muller, 2022).



Era #3 - 2022-2025 - Integration into everyday life

In this most modern era, it seems that AI is developing and integrating into our lives faster than we can think twice about its possible dangers. Clearly, its benefits are numerous, and only growing, and more than ever, this era focuses on the applications of AI in various sectors of life.

Projects that could take months without AI take less than half the time with the use of AI (Slimani). Furthermore, applications like AutoML are allowing AI to become even further accessible to even non-experts outside of the field, increasing efficiency not only within the AI field but other sectors as well (Salehin). With its clear benefits and abilities, making ML techniques more friendly for a non-expert is beneficial, if not essential in this day and age (Schmitt).

One of the most prevalent uses of AI is its use in multimodal models: namely, ChatGPT and the like. This means everything from image generation to text generation to summaries to feedback. These uses are often harmless, especially in the context of simple work or school-related uses they are usually used for, or the recently rising "Ghibli-style image" trends (Di Placido, 2025).

However, there have also been serious problems, specifically when using Al-generated research in professional instances like court-cases. Starting in 2023 and ranging all the way to current-day 2025, courts have been struggling with lawyers as they use Al-generated evidence that is eventually revealed to be untrue, created by Al's hallucinations. The most well-known is the Mata v. Avianca court case, where lawyers submitted fake, Al-generated extracts and citations from a court case to the court (Mata v. Avianca, 2022). They failed to check whether this case actually existed, and the case turned out to be one that Al fabricated. Lawyers' rationale to using Al in these cases range from misplaced belief that their associates would fact check the evidence to feeling tired and being pressed for time.

Another, lesser-seen application is in retail and ecommerce. Used by large-scale companies like Amazon, it aids companies in analyzing customer interests to help them improve ad campaigns and gain more customers, as well as in going through job applications. "Amazon Go" is another famous application, in which Amazon utilized AI in order to establish a chain of cashierless convenience stores in the US (Wankhede, 2018). This allowed them to streamline the process for customers, making visits even more time efficient and encouraging a larger customer base of regular users (Polacco & Backes, 2018). Another usage was by Macy's, a large department store based in the US, who created *Macy's On Call* with AI, utilizing it to function as a chatbot to answer customers' questions real time on the company website (Huang et al., 2021).

It is also used in healthcare, specifically in medical image processing and in monitoring patients. This includes analyzing the hospital scans such as CTs, MRIs, and PETs (Eklund). Al's ability for pattern recognition allows it to excel in this area. Another healthcare application is known as



human activity recognition (HAR) (Poulose), which uses AI to analyze the usual activities and motions of a person in order to predict what an abnormal act for that person may be. The use of AI rather than humans improves the accuracy of this monitoring, especially due to AI's excellent prediction abilities, and is thus incredibly helpful in a healthcare context.

Another application is one that nearly everyone has likely heard of by now: autonomous vehicles (AVs). This application requires AI to map out environments using sensors, determine the fastest routes using GPS data, and recognize important elements of the road, such as lanes, pedestrians, and other vehicles. Most importantly, it must be capable of making decisions, especially relating to when to stop, or which direction to go when it faces an obstacle. These AVs are supposed to minimize fuel usage, lower accident rates, and improve traffic flow (Bathla), all while allowing drivers to be more productive as they may engage in other activities during driving times (Shi).

Last but not least, this is the first era that has seen machine learning actively used to aid climate change mitigation. Its applications fall mostly under the jurisdiction of predictions. According to a paper written in 2024 by Ukoba et al., one of its purposes is helping to optimize renewable energy systems by analyzing complex environmental data and predicting patterns. Fluctuations in energy produced by green energy sources, such as solar or wind power, often makes energy systems efficient, preventing them from being integrated into large-scale power grids; predicting weather patterns via machine learning combats the unpredictability, allowing them to become more effective, efficient, and reliable. Another example is helping to predict the energy consumption of cars, buildings, or factories to gather data on their impact and reduce emissions as a whole. However, it is still true that machine learning still faces problems like ethical concerns, interpretability of data, and data quality in this application as well.

Analysis

Part I: Impact of AI on Climate Change

As AI has evolved, so has our understanding of its connection to climate change, both as a source of greenhouse gas emissions and as a tool for solutions to combat environmental issues. Early discussions were largely speculative, centered on AI's potential energy demands and ethical concerns. Over time, assessment methods became more precise, focusing on measurable factors such as the carbon dioxide equivalent emissions from training and deploying large-scale models, the energy consumption of data centers, and the environmental cost of manufacturing and maintaining hardware. At the same time, AI's contributions to climate action through applications like energy optimization and climate forecasting grew more apparent. As the scope of knowledge on the possible impacts of AI widened, so did the conclusions. This



section examines the quantifiable environmental impact of AI over time, detailing how metrics such as carbon output and energy usage have evolved alongside advances in both AI technology and climate measurement practices.

Era 1 (2014-2017):

In 2014, the global average temperature was the highest recorded temperature up to that year, with an average temperature of 0.69°C, or 1.24°F (NOAA, 2015). This was a further rise from the average temperature of 2010, which was 0.04°C. Six out of the twelve months ended up hitting record highs in history. Furthermore, the ocean temperature was 0.57°C, a dismal number that beat the previous record highs from 1998 and 2003, 16.1°C, by 0.05°C. Considering the size of the ocean, this is a vast increase.

Even at the end of the era, the situation did not improve. In 2017, the average temperature was 0.84°C, with a consistent increase rate since 2014 of around 0.06°C (NOAA, 2018). However, in the month of March, it hit even higher temperatures, hitting a record at 1.03°C. While this year did fall behind the heat of 2015 and 2016, the four years between 2014 and 2017 took up the top four hottest years in history, since 1998. Previously, new record highs had only been set after 13 years; now, each year there was a new record.

There were multiple possible contributors to this worsening climate situation. In general, the UN considers the leading causes to be fossil fuels, generated from man-made causes such as manufacturing, transportation, and generating power ("Causes and Effects", n.d.).

Most importantly, AI came to fruition as a possible rising contributor, taking the climate to an even more unfavorable position. In general, these four years between 2014 and 2017 constituted the initial boom of AI and ML. This meant that models became increasingly complex, involving more compute time, requiring more energy and hardware like the development of the GPU. They thus become more resource-intensive.

Most papers and academics interested in AI at this time viewed the AI and climate change relationship as a positive development, focusing most on its areas of possibility (Faghmous & Kumar, 2014). Al's high sensitivity meant that it could be used to factor for many unknowns or highly variable factors, making it especially apt at investigating mysterious, unknown topics like the environment (Bzdok, 2017).

Still, there were some cons of AI that had already arisen at this stage. For one, GPU energy efficiency. A paper written by Mittal and Vetter (2014) already expresses concern about the GPU energy usage of high-computation models, showing early awareness that the application of GPUs and rapid growth of AI capabilities has also affected its energy consumption. Another point of concern is the dataset size and training energy (Mittal & Vetter, 2014). In early 2018, just



after the tail end of this era, concern was also shown for Deep Neural Networks (DNNs) due to their high use in AI models despite their high computational requirements, and the challenges and complications faced in the process of attempting to create more energy-efficient alternatives (Chen et al., 2018).

These papers also eventually did reach out of the tighter circle of AI experts into the field of policymakers and more general academics.

Era 2 (2018-2022):

Following the trend of increasingly warm weather patterns, 2018 likewise fell within the rankings as the fourth warmest year in history (NOAA, 2019). It was similar to previous years, which had always had temperatures just above 1.0°C, with a global average temperature of 0.97°C.

There was more attention at this time to the negative impacts of AI on climate change, as it slowly became more researched and well-integrated into various other fields, following the publishing of the initial transformer paper that allowed others to utilize AI and machine learning.

Notably, Strubell et al. (2019) published the paper titled "Energy and Policy Considerations for Deep Learning in NLP", which speaks about the energy consumption in pounds of CO2 for a multitude of different well-known NLP models compared to the consumption of air travel, or an entire human life. This was the first ever paper that seriously considered this topic in detail. According to this paper, while a model like a basic transformer only produces 26 CO2-equivalent emissions, this number increases exponentially to reach 626,115 for BERT models.

After hearing of this, there was a rebuttal published by Google, with the argument that these numbers were inaccurate considering the efficiency of the materials used, such as highly specialized TPUs, and the geographic location, since the proportion of renewable energies used can also vary (Patterson et al., 2021). It is possible that some of these numbers may have been cherry-picked or not necessarily accurate, given that Google has Gemini within its products and thus wishes to improve its sales.

This was significant because it first brought light to one of the largest problems faced by NLP: the struggle to accurately measure or even estimate the exact energy consumption of these models. There was also the event of a social media post made by Jeff Dean (Dean, 2024), the lead of Google AI, focusing on the rebuttal of Google emissions. This worked to spread the news of AI and its impact on climate change past the confines of scientific papers, despite the fact that the information itself is likely difficult to trust.

Era 3 (2020-2025):



In the most recent era from 2022 to 2025, there was also a similar trend of increase in temperature (NOAA, 2025). As of May 29, 2025, 2024 was the hottest recorded year in terms of average surface temperature in history since 1850, when temperature recordings first began. The trend is shown clearly in the graph below, where temperatures began rising steadily in the 1970s, and seems to be increasingly exponential in form.

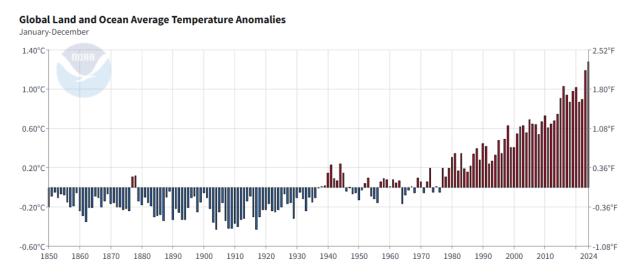


Figure from: NOAA, 2024

There has also been a rapid increase in attention on finding accurate methods to calculate the carbon outputs and impacts of Machine Learning systems (Luccioni). This came from the spark that started in 2019 between Strubell and Google. Papers currently being written, unlike those which focused primarily on the development of Al's computational skills without regard to energy efficiency, now focus much more on how it can help on climate mitigation (Amiri).

Now, even articles that are targeted to be read by the general public, like one written by MIT, writes about this topic as very nuanced, with the possibility of being both positive and negative (Stackpole). The conclusion is that more research and implementation is required to see exactly whether Al's climate impact can be mitigated without sacrificing its computational power.

General Impacts

In some ways, the impact of AI usage on the climate may be viewed as positive. One of the most common uses is in modelling, especially in situations that require processing large amounts of data. AI has been used to model for situations like weather patterns or natural disasters, to help mitigate their consequences. Datasite collections like the Climate Data User Guide ("Expert Insights", n.d.) have also been established due to this effort, which aided the development of this community. An example of this is research done in 2024, where machine



learning was used to predict the effect of climate change on US heat waves (Trok et al., 2024). The results of this trial usage of Al agreed with other human-processed data, demonstrating how it can successfully be used in this context. Other than natural phenomena, modelling is also useful when used for predicting emissions from engines or buildings, which are huge contributors to global emissions. Outside of simply predicting emissions produced, models can also predict how new biofuels or eco-friendly fuels may perform, leading to faster solutions to fuel-related emission problems (Khurana et al., 2021).

However, these advantages come at a cost: according to OpenAI themselves, AI models require a huge amount of energy in order to function, and this amount has only been growing as systems grow more complex and powerful in order to process more and more data (Amodei & Hernandez, 2018). The increase in function quality required the use of AI accelerators such as GPUs, which also contributed to the general increase in energy usage over time, as the more advanced a model became, the more components and computational resources it required (Luccioni & Hernandez-Garcia, 2023). Furthermore, not only do they require actual energy to power all of the necessary components, but also, as the servers begin to run they generate a large amount of heat, leading to the need of cooling systems that use up even more energy (Landram, 2025).

Part II: Discourse on AI and Climate

Following the quick development of AI, discourse surrounding its environmental impact has also grown. This section specifically examines how AI's relationship with climate change has been perceived in each of the three eras. This evaluation will focus on three main publication fields: academia, industry, and online publishing houses. By analyzing sentiment trends and categorizing proposed solutions, this section shows the evolution of views surrounding AI's environmental footprint.

Media Sampling

Academic analysis was based on articles accessed through platforms such as Google Scholar, IEEE Xplore, and arXiv. Papers were selected based on relevance to Al's environmental impact, using search terms like "Al and energy consumption," "Al and climate change," and "green Al." For each era, at least five papers were chosen to ensure there were a wide range of views represented.

Industry sentiment was assessed using over 25 publicly available sources, including whitepapers, annual ESG reports, and technical blogs from companies such as Google,



Microsoft, IBM, DeepMind, and smaller climate AI startups. Only documents that explicitly addressed AI's energy use or climate applications were included.

The view of public publishing houses was determined by choosing 10+ articles for each era from some of the most visited English-language news outlets when measured by web traffic as of March 2025 (Press Gazette), as well as more genre-specific sites. The most visited sites were *The New York Times, BBC, CNN, The Guardian,* and *News18*. The countries represented here are the US, the UK, and India. However, the NYT, BBC, and CNN are often read outside of their "host" countries, seen more as global networks. For example, BBC reaches around 318 million people according to their official website, updated recently, which far surpasses the number of people just in the UK ("Global news services"). More genre-specific sites include *Scientific American* and *MIT Technology Review*.

Sentiment Scale

To measure how each sector perceived Al's role in climate change, a numerical scale was developed, which reflected the level of concern or optimism toward Al's environmental impact, calculated using analysis of sources.

Score	Interpretation
+3	Strongly optimistic (Al significantly helps climate)
+2	Moderately optimistic (Al helps climate significantly more than harms)
+1	Slightly optimistic (Al helps climate slightly more than harms)
0	Neutral/mixed (Al helps and harms climate equally)
-1	Slightly concerned (Al harms climate slightly more than helps)
-2	Moderately concerned (AI harms significantly more than helps)
-3	Strongly alarmed (Al significantly harms climate)

Era 1: 2014-2017, Discovery and Novelty

Academia (+1)

In the field of academia, AI was explored as a tool for climate modeling and energy optimization, although this mainly took place towards the end of this era, in 2017 or the early months of 2018.



Research was cautiously optimistic, highlighting theoretical benefits. This includes papers such as a paper written by Yang et. al in 2017, which used AI techniques such as the artificial neural network (ANN) to predict the inflow levels of reservoirs based in the US and China (Yang et al., 2017). A paper written by Amasyali and EI-Gohary in early 2018 also developed a machine learning-based method to optimize HVAC systems, displaying this early academic interest in applying AI to energy efficiency.

However, the energy use of training models was already being flagged in parallel research, laying early groundwork for "Green AI" discourse. This occurred as early as 2016, where a paper by Li et al. examined the energy consumption of training then-popular neural network models on computer vision tasks, containing one of the first deep-dives and breakdowns of the energy used during training and inference for models utilizing GPUs.

Still, with the lack of concrete evidence or research at the time, there was not much that could be said on the topic.

Industry (0)

On the other hand, few common real-world applications that had actually been implemented existed at the time, due to Al being used mostly for research purposes within academia.

From environment reports from Google, Samsung, and IBM, it can be seen that corporate interest was nascent. In the Environmental Report by Google (2016), Al was discussed primarily in terms of data center automation, not environmental applications. IBM's 2017 Corporate Responsibility Report mentioned Al-driven enterprise solutions but lacked reference to climate or environmental targets (IBM, 2017). As can be seen, most conversations within these companies were centered on automation, not sustainability, leaving environmental effects largely absent from discourse.

Similar effects happened within startups in the AI sector, which focused on innovation narratives with no significant sustainability viewpoint. In a list of 25 rising startups in this era written for Forbes magazine, on which the detailed missions and purposes of each startup focus most on solving problems such as improving datasets and engines, and aiding manufacturing companies, without a hint of any startup working towards solving a climate-based problem (Columbus, 2018).

News Media (-1)

In more scientific articles, such as those from Scientific American, AI was being viewed positively throughout 2015-2017. This includes articles like "Springtime for AI: The Rise of Deep Learning", which highlights the new applications of deep learning and neural networks that had



arisen (Bengio). In 2015, it even published an article titled "Deep Learning is the AI Breakthrough We've Been Waiting For", stating clearly that the possibilities of the newly developed deep learning, allowing computers to be autonomous, could suggest positive consequences (Stix). With the introduction of AI being so recent at the time, these articles focus primarily on the small developments made at the time like photo recognition and the ability for computers to autonomously play video games.

But for some, rather than actual concrete information or data points about AI, skepticism about "AI hype" was common. The New York Times published such articles in 2014, for example, such as "Artificial Intelligence as a Threat", which speaks about how it may not take long for AI to "spiral out of control" (Bilton). An article from The New York Times in 2016 spoke on the topic of AI's great awakening in regards to its use in Google Translate, and showed clearly how the integration of AI into multiple Google products had changed the entire landscape of the market, also did not mention anything about its possible environmental impacts (Lewis-Kraus). At the time, celebrities like Elon Musk and books had been published about the grim outlook of AI usage. However, this article, and many others in common, mostly speak to general AI ethics rather than concerns with its environmental impact.

Era 2: 2018-2021, Deployment and Dissonance

Academia (+2)

Academic consensus grew around Al's potential to enhance energy forecasting, emission tracking, and climate modeling. For instance, Rolnick (2019) published "Tackling Climate Change with Machine Learning," a landmark survey outlining 13 domains where ML could accelerate sustainability. A paper by Campos et al. (2019) explored the uses of neural networks in making simulations of the Gulf of Mexico, specifically for the long-term predictions of weather and wave patterns. Wang et al. (2019) proposed hybrid models combining deep learning with Earth system science to better simulate climate. These signaled growing consensus that Al could support decarbonization, despite simultaneously ongoing ethical debates.

It is true that ethical debates on AI fairness and environmental equity also intensified. Most notably, the publishing of the paper analyzing carbon outputs of different AI models, written by Strubell et al. (2019). Overall, these were overshadowed by the mounting potential being discovered.

Industry (+1)

Al adoption in climate tech expanded greatly. Some applications were smart grids and precision agriculture. Microsoft's "Al for Earth" program, launched in 2017 and matured by 2019, provided open-access tools for land cover mapping and climate modeling (Spencer). IBM launched Al-powered energy management platforms, called Watson IoT for smart buildings ("IBM"), as



well as deployed a site through which companies without coding experience could easily design and deploy their own AI models ("Agent Builder", n.d.).

However, OpenAI's GPT-2 and similar models triggered controversy due to their high energy demands, signaling tension between utility and emissions, as noted in an article in a business article by Wired, titled "AI Can Do Great Things—if It Doesn't Burn the Planet" (Knight, 2020). Growth in data centers and compute-heavy AI models led to further mixed perceptions. Because the growth in uses and overall need for AI drastically increased in this time period of initial deployment, there was a larger focal point on distribution and development for speed and effectiveness over a focus on environmental impacts or energy efficiency.

In 2020, for instance, Microsoft wrote a study on the carbon benefits of cloud computing, showing how large-scale companies were interested in promoting a healthy environment and were attempting to balance efficiency of models alongside the growing need for carbon efficiency, making ambitious zero-carbon or zero-emission goals by 2025 and 2030 (Microsoft)

Public Publishing Houses (0)

In comparison to the fearful or overtly negative tones of previous years, coming from the concern about "Al hype", a more balanced tone emerged in news articles and media in this era. Overall, the tone balanced critique with cautious endorsement, showing rising literacy and awareness from writers and from the public about Al's environmental tradeoffs.

Highlighting this change were stories which praised AI for its profound positive impacts, such as tracking deforestation or improving disaster prediction. Scientific American, while it largely focused on AI's other applications in this era, contributed to this growing concept of climate change and AI with the article titled "What AI Can Do for Climate Change, and What Climate Change Can Do for AI" (Dimock, 2022). The New York Times first wrote on this concept in 2019, writing an article titled "How A.I. Can Help Handle Severe Weather" (Tugend, 2019).

Still there existed some particles that critiqued Al's environmental tradeoffs, particularly surrounding NLP model energy use. In 2019, MIT Technology Review's article titled "Training a single Al model can emit as much carbon as five cars in their lifetimes" (Hao, 2019) spotlighted the famous GPT-2 study by Strubell et al. (2019), which had also shaken the academic world in the very same year.

Era 3: 2022-2025 Reflection & Regulation

Academia (+1)

Recent academic work emphasizes responsible AI, defined by IBM as a set of principles used to guide AI's design, development, and uses, ensuring that AI solutions do not cause damage but



are able to aid organizations and stakeholders involved (Stryker, 2024). This includes papers written on green AI design, promotion of carbon-aware training, and AI for adaption strategies to the ever-changing climate. These were especially prevalent starting in the year 2023. In a paper by Liu and Yin (2023), green AI is discussed in detail, not only including facts on their carbon footprints but specifically emphasizing mitigation strategies, with a specific section dedicated to experimental setup on determining the best ways to mitigate the environmental impact of LLMs.

More and more studies have been committed to accurately measuring the energy costs of Al models. This first began with a paper by Samsi written in 2023, which was one of the first to analyze the computational and resources required by LLMs (Samsi et al., 2023). Although not focused on the environmental impacts that energy-intensive LLMs had, this was still a step forward in uncovering accurate data, past the simple estimations proposed in Strubell's 2019 paper.

Positive framing continues, but is now paired with accountability frameworks. Filho et al. (2022) discuss in their paper how Al can help enhance and enforce government and policy coherence, showing a shift in the fact that Al has expanded its boundaries and can now be used to aid outside of academic or research uses, in the society and community. This reflects its more widespread use by the wider community through the creation of multiple Al chatbots, as well as its continual integration into widely used products such as Siri (Paul & Robins-Early, 2024).

Industry (+2)

Recently big tech companies like Google or Microsoft have made climate pledges (e.g., "net-zero AI," "sustainable ML"). This applies especially to Google, who began creating a specific environmental report just for its AI usages starting from July of 2023 ("Reports", n.d.). Within, they feature multiple environmental sustainability strategies and targets, specifying both past and present metrics to show how they are working towards these goals (Google Sustainability, 2023). In the 2024 environmental review by Google, they boasted a low PUE score of 1.10, which measures the Power Usage Effectiveness of their facilities in terms of total facility energy consumption over total IT equipment energy consumption—a low score of 1.10 implies that most of the energy used is used directly by necessary equipment, meaning high energy efficiency. They also developed the TPU, and showed their use of AI prediction in projects like cool roofs, implementing reflective roofs to save energy from heating and AC, and to reduce emissions (Google Sustainability, 2024). Google's website also provides supporting information, writing multiple articles on their research on the uses and effects of Al's sustainability ("Driving Global Progress", n.d.). For instance, a recent article on Google's implementation of AI in farming, greatly boosting crop productivity ("AI's role", n.d.). They also specify that they are working together with the government and NGOs by providing them with climate data, and have even written an article on how AI can help advance progress on UN



SDG goal number 7, clean energy, in the future ("Al powered", 2024). This shows their faith and dedication to this new climate movement arising in the industry.

Startups have also chosen to focus more on finding solutions to mitigate the climate impact of AI, deploying AI in emissions tracking and carbon market validation. Some famous companies include Mistral AI and Anthropic, known for Llama and Cloud respectively. Google has aided this shift by choosing to fund AI startups which have a focus on climate change aid or mitigation purposes, such as tackling deforestation or ecosystem conservation (ESG News, 2025).

News Media (+1)

Unlike covering Al's effects or possible damaging outcomes, coverage shifted more towards solutions, including carbon-efficient algorithms, open-source climate datasets, and climate Al competitions. While some criticism persists, it is increasingly constructive.

For instance, in the Scientific American, the article "AI Needs to Be More Energy Efficient" discusses the solution of urging the public to be more involved in promoting the development of energy-efficient methodologies and models in the industry, ending the article with a clear call to action that clearly displays the shift towards constructive criticism (Scientific American, 2025). In the New York Times, articles from 2022 onwards have consistently been published on the topic of AI's usages in climate prediction and control, as well as its energy efficiency, such as "Will A.I. Ruin the Planet or Save the Planet?" (Lohr, 2024), Dealbook newsletter, which mainly focuses on finance, showing how AI and climate issues have managed to catch the attention of other sectors (Sorkin et al., 2024), and most recently, the article "Can You Choose an A.I. Model That Harms the Planet Less?" by Mulkey in 2025 reflects this same sentiment.

Proposed Solutions to Lessen Al's Climate Impact

Across all sectors, proposed solutions generally fell into five categories.

Firstly, the most tech-centric solution: working towards computational efficiency. Specifically, the focus on Green AI, such as optimizing training efficiency, pruning models, and developing low-power architectures, with suggested motivations like cash prizes for the most effective (as outlined in the Google paper). In media (specify what kind of media), more instances of these cases are being highlighted, such as Meta's low-carbon infrastructure and DeepMind's use of AI to reduce data center cooling costs.

Secondly, using AI to aid environmental intelligence, to counter its negative effects with positive ones. This translates to being used as a tool in academia and research for climate modelling,



satellite image analysis, biodiversity mapping, and disaster response optimization (find examples of all). In the industry, it is aiding in insurance risk assessments, precision agriculture, and supply chain monitoring, especially for companies like Amazon. Case stories about real-time wildfire detection or glacier monitoring have also been shared across media.

The third solution involves the cooperation of the government, proposing more governmental regulations and governance in this sector. For example, the call for standards to be raised upon responsible AI model development, and for academic models to carbon-label AI models and disclose their full environmental impact to the public or to other researchers, also mentioned in Google's rebuttal paper.

Fourth, market-based and social incentives are suggested. For example, giving carbon credits for efficient code or models, having Al-enabled ESG investing platforms, and public awards for climate-conscious innovation like green-buildings do could all incentivize researchers and industries alike to target energy efficiency.

Lastly, multiple sources suggest a solution related to education, aimed at the public, in order to garner more public engagement from the wider community. This includes pushing for transparency in revealing Al's environmental costs from researchers in academia, and hosting more climate-related Al hackathons or open-data challenges to crowdsource solutions and ignite further interest or passion into this topic, accelerating developments.

Discussion

The evolving relationship between artificial intelligence and climate change mitigation suggests a nuanced trajectory ahead. As transformer-based architectures become more deeply woven into the fabric of everyday life, from personal assistants and automated logistics to real-time energy optimization systems, their aggregate impact on emissions and climate outcomes will grow. On the one hand, transformers offer substantial efficiency gains in forecasting, supply chain coordination, and environmental monitoring. On the other hand, their continued scale-up and deployment could exacerbate concerns around computational emissions, particularly if training and inference continue to rely on carbon-intensive energy sources.

Looking forward, there is reason for cautious optimism. Recent shifts in both academia and industry demonstrate increasing recognition of Al's carbon footprint and a parallel push toward "green Al," low-power model design, and carbon labeling. As transformers become ubiquitous, their environmental impact could be mitigated by integrating carbon-aware scheduling, leveraging renewable-powered data centers, and prioritizing efficiency in model design.



Secondarily, the findings of this investigation suggest that future strategies should move beyond single-sector solutions. Effective climate-aligned AI development may rely on combining regulatory and market incentives and policies requiring carbon reporting paired with credits or tax benefits for low-carbon AI development. Technical innovation is also required, as in advancing in the creation of sparse models and efficient training protocols, learning to reduce the amount of data movement and computation required per model.

Of course, it is not just up to the academics and professionals. Public engagement is also required, as well as aid from other sectors. For instance, transparent communication through social media or other platforms of Al's climate impact to foster greater societal demand and urgency in developing greener approaches, especially since models like ChatGPT are gaining daily use. Cross-sector collaboration, such as joint initiatives between academia, industry, and public agencies are also helpful, in order to set new standards, share and refine best practices, and overall co-fund a sustainable Al infrastructure for the future.

By combining policies, technical approaches, and social engagement, future strategies and solutions can help ensure that the rapid integration of machine learning and other Al tools becomes an asset rather than a liability in global climate efforts.

Conclusion

From early curiosity to cautious endorsement, each sector's view of AI and climate has clearly matured over time. While academic literature has consistently led optimism, industry's narrative caught up post-2020 with measurable deployments. Media sentiment evolved from critique to cautious hope, reflecting growing public awareness of both AI's promise and its planetary cost.

References

- [1] Agent builder. (n.d.). IBM. Retrieved July 15, 2025, from https://www.ibm.com/kr-ko/products/watsonx-orchestrate/ai-agent-builder?lnk=hpls1kr
- [2] Al's Role in Boosting Crop Productivity. (n.d.). Bloomberg. Retrieved July 7, 2025, from https://sponsored.bloomberg.com/article/google-sustainability/ai-s-role-in-boosting-crop-productivity



- [3] *Al-powered pathways: Advancing SDG7 for a sustainable energy future.* (2024, July). Google Sustainability. Retrieved July 9, 2025, from https://sustainability.google/stories/ai-powered-pathways/
- [4] Amasyali, K., & El-Gohary, N. (2018). Machine learning-based occupant energy use behavior optimization. *Construction Research Congress*, 379-389. https://doi.org/10.1061/9780784481301.038
- [5] Amiri, Z., Heidari, A., & Navimipour, N. J. (2024). Comprehensive survey of artificial intelligence techniques and strategies for climate change mitigation. *Energy*, *308*. https://doi.org/10.1016/j.energy.2024.132827.
- [6] Amodei, D., & Hernandez, D. (2018, May 16). *Al and compute*. OpenAl. Retrieved July 14, 2025, from https://openai.com/index/ai-and-compute/
- [7] Baji, T. (2017). GPU: The biggest key processor for AI and parallel processing. In *Proceedings Volume 10454, Photomask Japan 2017: XXIV Symposium on Photomask and Next-Generation Lithography Mask Technology.*
- [8] Bathla, G., Bhadane, K., Singh, R. K., Kumar, R., Aluvalu, R., Krishnamurthi, R., Kumar, A., Thakur, R. N., & Basheer, S. (2022). Autonomous Vehicles and Intelligent Automation: Applications, Challenges, and Opportunities. *Mobile Information Systems*, 2022(1). https://doi.org/10.1155/2022/7632892
- [9] Bengio, Y. (2016, June 1). *Springtime for AI: The Rise of Deep Learning*. Scientific American. Retrieved July 7, 2025, from https://www.scientificamerican.com/article/springtime-for-ai-the-rise-of-deep-learning/
- [10] Bilton, N. (2014, November 5). Artificial intelligence as a threat. *The New York Times*. https://www.nytimes.com/2014/11/06/fashion/artificial-intelligence-as-a-threat.html
- [11] Bommasani, R., Hudson, D.A., Adeli, E., Altman, R., Arora, S., Arx, S.V., Bernstein, M.S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J.Q., Demszky, D., Donahue, C., Doumbouya, M., Durmus, E., Ermon, S., Etchemendy, J., Ethayarajh, K., Fei-Fei, L., Finn, C., Gale, T., Gillespie, L., Goel, K., Goodman, N., Grossman, S., Guha, N., Hashimoto, T., Henderson, P., Hewitt, J., Ho, D.E., Hong, J., Hsu, K., Huang, J., Icard, T., Jain, S., Jurafsky, D., Kalluri, P., Karamcheti, S., Keeling, G., Khani, F., Khattab, O., Koh, P.W., Krass, M., Krishna, R., Kuditipudi, R., Kumar, A., Ladhak, F., Lee, M., Lee, T., Leskovec, J., Levent, I., Li, X.L., Li, X., Ma, T., Malik, A., Manning, C.D., Mirchandani, S., Mitchell, E., Munyikwa, Z., Nair, S., Narayan, A., Narayanan, D., Newman, B., Nie, A., Niebles, J.C., Nilforoshan, H., Nyarko, J., Ogut, G., Orr, L., Papadimitriou, I., Park, J.S., Piech, C., Portelance, E., Potts, C., Raghunathan, A., Reich, R., Ren, H., Rong, F., Roohani, Y., Ruiz, C., Ryan, J., Ré, C., Sadigh, D., Sagawa, S., Santhanam, K., Shih, A., Srinivasan, K., Tamkin, A., Taori, R., Thomas, A.W., Tramèr, F., Wang, R.E., Wang, W., Wu, B., Wu, J., Wu, Y.,



- Xie, S.M., Yasunaga, M., You, J., Zaharia, M., Zhang, M., Zhang, T., Zhang, X., Zhang, Y., Zheng, L., Zhou, K., Liang, P. (2021). On the Opportunities and Risks of Foundation Models. arXiv preprint arXiv:2108.07258.
- [12] Brodtkorb, A. R., Hagen, T. R., & Sætra, M. L. (2013). Graphics processing unit (GPU) programming strategies and trends in GPU computing. *Journal of Parallel and Distributed Computing*, 73(1), 4-13. https://doi.org/10.1016/j.jpdc.2012.04.003
- [13] Bzdok, D., Krzywinski, M., & Altman, N. (2017). Points of Significance: Machine learning: a primer. *Nature methods*, *14*(12), 1119?1120. https://doi.org/10.1038/nmeth.4526
- [14]Campos, R. M., Krasnopolsky, V., Alves, J.-H. G. M., & Penny, S. G. (2019). Nonlinear wave ensemble averaging in the gulf of mexico using neural networks. *Journal of Atmospheric and Oceanic Technology*, 36(1), 113-127. https://doi.org/10.1175/jtech-d-18-0099.1
- [15] Causes and effects of climate change. (n.d.). United Nations. Retrieved July 3, 2025, from https://www.un.org/en/climatechange/science/causes-effects-climate-change
- [16] Chen, Y.-H., Emer, J., Sze, V., & Yang, T.-J. (2018). Understanding the limitations of existing energy-efficient design approaches for deep neural networks. *SysML*.
- [17] Columbus, L. (2018, August 26). *25 machine learning startups to watch in 2018*. Forbes. Retrieved July 19, 2025, from
- https://www.forbes.com/sites/louiscolumbus/2018/08/26/25-machine-learning-startups-to-watch-in-2018/
- [18] Da Li, Xinbo Chen, Michela Becchi, and Ziliang Zong. 2016. Evaluating the energy efficiency of deep convolutional neural networks on cpus and gpus. 2016 IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom) (BDCloudSocialCom-SustainCom), pages 477?484.
- [19] Dean, J. [JeffDean]. (2024, October 7). However that single number is still clouding people's assessment of the environmental costs (e.g. see how this thread started), and @strubell has flatly refused to amend the Arxiv or ACL paper with correct data, and has even co-authored papers after knowing about these errors (which we informed her of as part of our discussions in our writing of the 2021 paper https://arxiv.org/abs/2104.10350). To them further spread this erroneous information (e.g. the first sentence of the related work section of the 2022 work in https://arxiv.org/pdf/2206.05229, which reiterates the wildly inaccurate "lifetime of five cars" meme) after being informed about the incorrect numbers is not honest and shouldn't be how science or transparency is done. [Post]. X. https://x.com/JeffDean/status/1843101133310714176
- [20] Dimock, W. C. (2022, April 5). What AI can do for climate change, and what climate change can do for AI. Scientific American. Retrieved July 10, 2025, from https://www.scientificamerican.com/article/what-ai-can-do-for-climate-change-and-what-climate-change-can-do-for-ai/



- [21] Di Placido, D. (2025, March 27). *The Al-Generated Studio Ghibli Trend, Explained*. Forbes. Retrieved July 4, 2025, from
- https://www.forbes.com/sites/danidiplacido/2025/03/27/the-ai-generated-studio-ghibli-trend-explained/
- [22] Dodge, J., Prewitt, T., Combes, R.T.D., Odmark, E., Schwartz, R., Strubell, E., Luccioni, A.S., Smith, N.A., DeCario, N., Buchanan, W. (2022). Measuring the Carbon Intensity of AI in Cloud Instances. arXiv preprint arXiv:2206.05229.
- [23] *Driving global progress with our partners*. (n.d.). Google Sustainability. Retrieved July 7, 2025, from https://sustainability.google/research/
- [24] Eklund, A., Dufort, P., Forsberg, D., & LaConte, S. M. (2013). Medical image processing on the GPU? past, present and future. *Medical Image Analysis*, *17*(8), 1073-1094. https://doi.org/10.1016/j.media.2013.05.008
- [25] ESG News. (2025, May 5). *Google Backs 15 AI Startups to Tackle Nature Loss and Climate Risk*. ESG News. Retrieved July 10, 2025, from https://esgnews.com/google-backs-15-ai-startups-to-tackle-nature-loss-and-climate-risk/
- [26] Expert insights on observation-based data of the climate system. (n.d.). Climate Data Guide. Retrieved July 12, 2025, from https://climatedataguide.ucar.edu/about
- [27] Faghmous, J. H., & Kumar, V. (2014). A big data guide to understanding climate change: The case for theory-guided data science. *Big Data*, *2*(3), 155-163. https://doi.org/10.1089/big.2014.0026
- [28] *Global news services*. (n.d.). BBC. Retrieved July 7, 2025, from https://www.bbc.com/aboutthebbc/whatwedo/worldservice
- [29] Google. (2016). *Google environmental report 2016*. https://www.gstatic.com/gumdrop/sustainability/environmental-report-2016.pdf
- [30] Google Sustainability. (2023). *Google 2023 Sustainability Report*. Retrieved July 7, 2025, from https://www.gstatic.com/gumdrop/sustainability/google-2023-environmental-report.pdf
- [31] Google Sustainability. (2024). *Google 2024 Environmental Report*. Retrieved July 7, 2025, from https://www.gstatic.com/gumdrop/sustainability/google-2024-environmental-report.pdf
- [32] Growing the internet while reducing energy consumption. (n.d.). Google. Retrieved July 12, 2025, from https://datacenters.google/efficiency/#measuring-efficiency
- [33] Han, X., Zhang, Z., Ding, N., Gu, Y., Liu, X., Huo, Y., Qiu, J., Yao, Y., Zhang, A., Zhang, L., Han, W., Huang, M., Jin, Q., Lan, Y., Liu, Y., Liu, Z., Lu, Z., Qiu, X., Song, R., . . . Zhu, J. (2021).



Pre-trained models: Past, present and future. *Al Open*, 2, 225-250. https://doi.org/10.1016/j.aiopen.2021.08.002

- [34] Hao, K. (2019, June 6). *Training a single AI model can emit as much carbon as five cars in their lifetimes*. MIT Technology Review. Retrieved July 5, 2025, from https://www.technologyreview.com/2019/06/06/239031/training-a-single-ai-model-can-emit-as-m uch-carbon-as-five-cars-in-their-lifetimes/
- [35] Huang, A., Chao, Y., de la Mora Velasco, E., Bilgihan, A., & Wei, W. (2021). When artificial intelligence meets the hospitality and tourism industry: An assessment framework to inform theory and management. *Journal of Hospitality and Tourism Insights*, *5*(5), 1080-1100. https://doi.org/10.1108/jhti-01-2021-0021
- [36] IBM. (2017). *IBM corporate responsibility report 2017*. https://www.responsibilityreports.com/HostedData/ResponsibilityReportArchive/i/NYSE_IBM_20 17.pdf
- [37] IBM and KONE Watson IoT gives lift to innovation in smart buildings [Video]. (n.d.). IBM. https://mediacenter.ibm.com/media/IBM+and+KONE+-+Watson+IoT+Gives+Lift+To+Innovation+In+Smart+Buildings/1_pk8871cw
- [38]Khurana, S., Saxena, S., Jain, S., & Dixit, A. (2021). Predictive modeling of engine emissions using machine learning: A review. *Materials Today: Proceedings*, *38*, 280-284. https://doi.org/10.1016/j.matpr.2020.07.204
- [39] Knight, W. (2020, January 21). *Al can do great things—if it doesn't burn the planet*. Wired. Retrieved July 10, 2025, from https://www.wired.com/story/ai-great-things-burn-planet/
- [40] Landram, K. (2025, February 11). Slashing AI data center cooling cost, GPU/CPU power use. Carnegie Mellon University. Retrieved August 10, 2025, from https://engineering.cmu.edu/news-events/news/2025/02/11-ai-data-center-cooling-breakthrough.html
- [41] Leal Filho, W., Wall, T., Rui Mucova, S. A., Nagy, G. J., Balogun, A.-L., Luetz, J. M., Ng, A. W., Kovaleva, M., Safiul Azam, F. M., Alves, F., Guevara, Z., Matandirotya, N. R., Skouloudis, A., Tzachor, A., Malakar, K., & Gandhi, O. (2022). Deploying artificial intelligence for climate change adaptation. *Technological Forecasting and Social Change*, *180*, 121662. https://doi.org/10.1016/j.techfore.2022.121662
- [42] Lewis-Kraus, G. (2016, December 14). The great A.I. awakening. *The New York Times*. https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html
- [43] Liu, V., Yin, Y. (2024). Green AI: Exploring Carbon Footprints, Mitigation Strategies, and Trade Offs in Large Language Model Training. arXiv preprint arXiv:2404.01157.



- [44] Lohr, S. (2024, August 26). Will A.I. ruin the planet or save the planet? *The New York Times*. https://www.nytimes.com/2024/08/26/climate/ai-planet-climate-change.html
- [45] Luccioni, A.S., Hernandez-Garcia, A. (2023). Counting Carbon: A Survey of Factors Influencing the Emissions of Machine Learning. arXiv preprint arXiv:2302.08476.
- [46] Luccioni, A.S., Hernandez-Garcia, A. (2023). Counting Carbon: A Survey of Factors Influencing the Emissions of Machine Learning. arXiv preprint arXiv:2302.08476.
- [47] *Mata v. Avianca*, No. 1:22-cv-01461 (S.D.N.Y. 2022). https://law.justia.com/cases/federal/district-courts/new-york/nysdce/1:2022cv01461/575368/54/
- [48] McClanahan, C. (2011, March). *History and evolution of GPU architecture*. Mcclanahoochie. Retrieved July 6, 2025, from https://mcclanahoochie.com/blog/wp-content/uploads/2011/03/gpu-hist-paper.pdf
- [49] McGuffie, K., Newhouse, A. (2020). The Radicalization Risks of GPT-3 and Advanced Neural Language Models. arXiv preprint arXiv:2009.06807.
- [50] Microsoft. (2018). *The Carbon Benefits of Cloud Computing*. [White paper]. https://www.microsoft.com/en-us/download/details.aspx?id=56950
- [50] Mittal, S., & Vetter, J. S. (2014). A Survey of Methods for Analyzing and Improving GPU Energy Efficiency. *ACM Computing Surveys (CSUR)*, *47*(2), 1-21. https://doi.org/10.1145/2636342
- [51] Mulkey, S. K., & Stevens, H. (2025, June 19). Can you choose an A.I. model that harms the planet less? *The New York Times*. https://www.nytimes.com/2025/06/19/climate/ai-emissions-chatbot-accuracy.html
- [52] Muller, B. (2022, March 2). BERT 101 ? state of the art NLP model explained. *Hugging face*. https://huggingface.co/blog/bert-101
- [53] NOAA. (2015). Monthly Global Climate Report for Annual 2014. https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00672
- [54] NOAA. (2018). Monthly Global Climate Report for Annual 2017. https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00672
- [55] NOAA. (2019). Monthly Global Climate Report for Annual 2018. https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00672
- [56] NOAA. (2025). Monthly Global Climate Report for January 2025. https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00672



- [57] O'Donnell, J., & Crownhart, C. (2025, May 20). *Everything you need to know about estimating Al's energy and emissions burden*. MIT Technology Review. Retrieved July 7, 2025, from https://www.technologyreview.com/2025/05/20/1116331/ai-energy-demand-methodology/
- [58] Patterson, D., Gonzalez, J., Le, Q., Liang, C., Munguia, L., Rothchild, D., So, D., Texier, M., Dean, J. (2021). Carbon Emissions and Large Neural Network Training. arXiv preprint arXiv:2104.10350.
- [59] Paul, K., & Robins-Early, N. (2024, June 10). *Apple brings ChatGPT to Siri as it debuts 'Apple Intelligence' at WWDC 2024*. The Guardian. Retrieved July 16, 2025, from https://www.theguardian.com/technology/article/2024/jun/10/apple-ai-product-launch
- [60] Polacco, A., & Backes, K. (2018). The amazon go concept: Implications, applications, and sustainability. *Journal of Business and Management*, *24*(1), 79-92. https://doi.org/10.6347/JBM.201803_24(1).0004
- [61] Poulose, A., Kim, J. H., & Han, D. S. (2022). HIT HAR: Human Image Threshing Machine for Human Activity Recognition Using Deep Learning Models. *Computational Intelligence and Neuroscience*, 2022(1). https://doi.org/10.1155/2022/1808990
- [62] Press Gazette. (2025, July 14). *Indian news websites are fastest growers in global English language top 50*. Press Gazette. Retrieved July 7, 2025, from https://pressgazette.co.uk/media-audience-and-business-data/media_metrics/most-popular-web sites-news-world-monthly-2/
- [63] Reports and case studies charting our progress and ambitions. (n.d.). Google Sustainability. Retrieved July 8, 2025, from https://sustainability.google/reports/
- [64] Salehin, I., Islam, S., Saha, P., Noman, S.M., Tuni, A., Hasan, M., & Baten, A. (2024). AutoML: A systematic review on automated machine learning with neural architecture search. *Journal of Information and Intelligence*, *2*(1), 52-81. https://doi.org/10.1016/j.jiixd.2023.10.002
- [65] Samsi, S., Zhao, D., McDonald, J., Li, B., Michaleas, A., Jones, M., Bergeron, W., Kepner, J., Tiwari, D., & Gadepally, V. (2023). From words to watts: Benchmarking the energy costs of large language model inference. *2023 IEEE High Performance Extreme Computing Conference (HPEC)*, 1-9. https://doi.org/10.1109/hpec58863.2023.10363447
- [66] Schmitt, M. (2022). Automated machine learning: Al-driven decision making in business analytics. arXiv preprint arXiv:2205.10538.
- [67] Scientific American. (2025, March 18). *Al needs to be more energy-efficient*. Scientific American. Retrieved July 24, 2025, from https://www.scientificamerican.com/article/ai-needs-to-be-more-energy-efficient/



- [68] Shi, W., Alawieh, M. B., Li, X., & Yu, H. (2017). Algorithm and hardware implementation for visual perception system in autonomous vehicle: A survey. *Integration*, *59*, 148-156. https://doi.org/10.1016/j.vlsi.2017.07.007
- [69] Slimani, I., Slimani, N., Achchab, S., Saber, M., Farissi, I., Sbiti, N., & Amghar, M. (2022). Automated machine learning: the new data science challenge. *International Journal of Electrical and Computer Engineering (IJECE)*, 12, 42-43. https://doi.org/10.11591/ijece.v12i4.pp4243-4252
- [70] Sorkin, A. R., Mattu, R., Warner, B., Kessler, S., Merced, M. J. de la, Hirsch, L., & Livni, E. (2024, May 6). How bad is A.I. for the climate? *The New York Times*. https://www.nytimes.com/2024/05/06/business/dealbook/ai-power-energy-climate.html
- [71] Spencer, G. (2018, September 4). *Al for earth: Helping save the planet with data science*. Microsoft. Retrieved July 7, 2025, from https://news.microsoft.com/apac/features/ai-for-earth-helping-save-the-planet-with-data-science/
- [72] Stackpole, B. (2025, February 10). *Climate change and machine learning? the good, bad, and unknown*. MIT Sloan School of Management. Retrieved July 6, 2025, from https://mitsloan.mit.edu/ideas-made-to-matter/climate-change-and-machine-learning-good-bad-and-unknown
- [73] Stix, G. (2015, December 1). *Deep Learning Is the A.I. Breakthrough We've Been Waiting For.* Scientific American. Retrieved July 7, 2025, from https://www.scientificamerican.com/article/deep-learning-is-the-a-i-breakthrough-we-ve-been-waiting-for/
- [74] Strubell, E., Ganesh, A., McCallum, A. (2019). Energy and Policy Considerations for Deep Learning in NLP. arXiv preprint arXiv:1906.02243.
- [75] Stryker, C. (2019). *What is Responsible AI?* IBM. Retrieved July 8, 2025, from https://www.ibm.com/think/topics/responsible-ai
- [76] Sutskever, I., Vinyals, O., Le, Q.V. (2014). Sequence to Sequence Learning with Neural Networks. arXiv preprint arXiv:1409.3215.
- [77] Torres, J., Vaca, C., Terán, L., & Abad, C. L. (2020). Seq2Seq models for recommending short text conversations. *Expert Systems with Applications*, *150*. https://doi.org/10.1016/j.eswa.2020.113270
- [78] Torres, J., Vaca, C., Terán, L., & Abad, C. L. (2020). Seq2Seq models for recommending short text conversations. *Expert Systems With Applications*, *150*, 113270. https://doi.org/10.1016/j.eswa.2020.113270
- [79] Trok, J. T., Barnes, E. A., Davenport, F. V., & Diffenbaugh, N. S. (2024). Machine learning?based extreme event attribution. *Science Advances*, *10*(34). https://doi.org/10.1126/sciadv.adl3242



- [80] Tugend, A. (2019, May 12). How A.I. can help handle severe weather. *The New York Times*. https://www.nytimes.com/2019/05/12/climate/artificial-intelligence-climate-change.html
- [81] Ukoba, K., Onisuru, O. R., & Jen, T.-C. (2024). Harnessing machine learning for sustainable futures: Advancements in renewable energy and climate change mitigation. *Bulletin of the National Research Centre*, 48(1), 99. https://doi.org/10.1186/s42269-024-01254-7
- [82] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, ?., & Polosukhin, I. (2017). Attention is all you need. In *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 6000-6010).
- [83] Wankhede, K., Wukkadada, B., & Nadar, V. (2018). Just walk-out technology and its challenges: A case of amazon go. 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), 254-257. https://doi.org/10.1109/icirca.2018.8597403
- [84] What Is Climate Change? (n.d.). United Nations. Retrieved July 7, 2025, from https://www.un.org/en/climatechange/what-is-climate-change
- [85] Yang, T., Asanjan, A. A., Welles, E., Gao, X., Sorooshian, S., & Liu, X. (2017). Developing reservoir monthly inflow forecasts using artificial intelligence and climate phenomenon information. *Water Resources Research*, *53*(4), 2786-2812. https://doi.org/10.1002/2017wr020482
- [86] Yin, X., & Wan, X. (2022). How Do Seq2Seq Models Perform on End-to-End Data-to-Text Generation? *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics Volume 1: Long Papers*, 7701-7710.
- [87] Yuventi, J., & Mehdizadeh, R. (2013). A critical analysis of power usage effectiveness and its use in communicating data center energy consumption. *Energy and Buildings*, *64*, 90-94. https://doi.org/10.1016/j.enbuild.2013.04.015
- [88] Zhang, M., & Li, J. (2021). A commentary of GPT-3 in MIT Technology Review 2021. Fundamental Research, 1(6), 831-833. https://doi.org/10.1016/j.fmre.2021.11.011
- [89] Zong, M., Krishnamachari, B. (2022). a survey on GPT-3. arXiv preprint arXiv:2212.00857.