arXiv:2404.13885v1 [cs.CY] 22 Apr 2024

# SURVEYING ATTITUDINAL ALIGNMENT BETWEEN LARGE LANGUAGE MODELS VS. HUMANS TOWARDS 17 SUSTAINABLE DEVELOPMENT GOALS

**Qingyang Wu***
University of California Los Angeles
qw@ucla.edu

**Ying Xu***
Washington University in St. Louis
x.ying1@wustl.edu

**Tingsong Xiao**
University of Florida
xiaotingsong@ufl.edu

**Yunze Xiao**
Carnegie Mellon University
yunzex@andrew.cmu.edu

**Yitong Li**
Stanford University
yitongli@stanford.edu

**Tianyang Wang**
Xi'an Jiaotong-Liverpool University
Tianyang.Wang21@student.xjtlu.edu.cn

**Yichi Zhang**
Fudan University
zhangyichi23@m.fudan.edu.cn

**Shanghai Zhong**
Independent Researcher
wuhanzsh@gmail.com

**Yuwei Zhang**
Northeastern University
zhang.yuwei@northeastern.edu

**Wei Lu**
Reinsurance Group of America
Wei.Lu@rgare.com

**Yifan Yang**
University of Southern California
yyang295@usc.edu

## ABSTRACT

Large Language Models (LLMs) have emerged as potent tools for advancing the United Nations' Sustainable Development Goals (SDGs). However, the attitudinal disparities between LLMs and humans towards these goals can pose significant challenges. This study conducts a comprehensive review and analysis of the existing literature on the attitudes of LLMs towards the 17 SDGs, emphasizing the comparison between their attitudes and support for each goal and those of humans. We examine the potential disparities, primarily focusing on aspects such as understanding and emotions, cultural and regional differences, task objective variations, and factors considered in the decision-making process. These disparities arise from the underrepresentation and imbalance in LLM training data, historical biases, quality issues, lack of contextual understanding, and skewed ethical values reflected. The study also investigates the risks and harms that may arise from neglecting the attitudes of LLMs towards the SDGs, including the exacerbation of social inequalities, racial discrimination, environmental destruction, and resource wastage. To address these challenges, we propose strategies and recommendations to guide and regulate the application of LLMs, ensuring their alignment with the principles and goals of the SDGs, and therefore creating a more just, inclusive, and sustainable future.

***Keywords*** Large Language Models · Sustainable Development Goals · Attitudinal Alignment · AI Ethics

## 1 Introduction

The concept of sustainable development arises from profound contemplation on the future of global humanity. It was initially introduced in 1987 through the report "Our Common Future" by the World Commission on Environment and Development, highlighting the balance between human development and ecological protection [1]. Subsequently, the Millennium Development Goals (MDGs) were adopted at the Millennium Summit from 6 to 8 September 2000,

*Qingyang Wu and Ying Xu contributed equally.

Figure 1: The details of 17 Sustainable Development Goals (SDGs).

signifying the first global concerted effort to address issues such as poverty, hunger, and disease [2]. In September 2015, the United Nations endorsed the 2030 Agenda for Sustainable Development (2030 Agenda) and with its 17 Sustainable Development Goals (SDGs) and 169 targets, aimed at addressing a range of urgent global challenges by 2030 [3] as shown in Figure 1.

In contrast to the top-down approach of the MDGs, the SDGs emerged from a more inclusive and broadly negotiated process, encompassing not only poverty and hunger but also health, education, gender equality, clean energy, reducing inequalities, sustainable cities and communities, and climate change across 17 dimensions [2]. The 17 SDGs and their 169 specific targets provide a clear roadmap from 2015 to 2030 for achieving a more just, inclusive, and sustainable world, guiding national plans, priorities, and investments to alleviate poverty, promote development, and influence the definition, funding, and measurement of national development.

The latest "Sustainable Development Goals Report 2023" emphasizes the importance of achieving the SDGs by the 2030 deadline [4]. The report underscores the ongoing challenges globally in addressing poverty eradication, and hunger, promoting gender equality, as well as tackling the triple crises of climate change, biodiversity loss, and pollution. It warns that progress remains weak and insufficient on over 50% of the SDG indicators, with 30% of indicators showing stagnation or regression. The COVID-19 pandemic, along with the triple crises of climate change, biodiversity loss, and pollution, has had a disruptive and enduring impact globally. Rising inflation, unsustainable debt burdens, the COVID-19 pandemic, and severe local conflicts have severely squeezed countries' fiscal space, weakening their ability to invest in green recovery. If current trends persist, it is estimated that by 2030, approximately 575 million people will still live in extreme poverty, and many vulnerable groups worldwide will still lack social protection coverage. Progress on many key targets remains weak and insufficient, including those related to poverty, hunger, and climate. The report also mentions some urgent action recommendations, including strengthening the linkages between public health and biodiversity conservation and sustainable use, enhancing governments' and stakeholders' monitoring and forecasting capabilities of the impacts of biodiversity loss on human well-being, tracking the targets and indicators of the Kunming-Montreal Global Biodiversity Framework, and bridging the $700 billion biodiversity financing gap, increasing annual financing from all sources by at least $500 billion, and eliminating and reforming incentives detrimental to biodiversity.

Some SDG targets may face challenges in being achieved by 2030, but Large Language Models (LLMs) and Artificial Intelligence (AI) potentially have the capability to change these trends. LLMs serve as vital channels for disseminating information, enhancing public awareness of the SDGs, and assisting policymakers, researchers, and educators in making evidence-based decisions. Additionally, they can contribute to educating future policymakers on sustainability issues. Interdisciplinary collaboration is crucial for achieving the SDGs, and LLMs play a key role in facilitating cooperation by integrating information from different fields and analyzing vast amounts of data to monitor progress and identify effective strategies. In practice, LLMs can handle large datasets, and identify patterns and insights, thereby supporting policymaking and action towards achieving SDG targets. For example, AI can help optimize agricultural resource utilization, enhance educational access through personalized learning, and support healthcare through predictive analytics. (see Fig. 1) Therefore, prioritizing the attitude of LLMs towards SDG targets is crucial to ensure that their applications align with social and environmental sustainability, rather than increasing risks and inequalities [5].

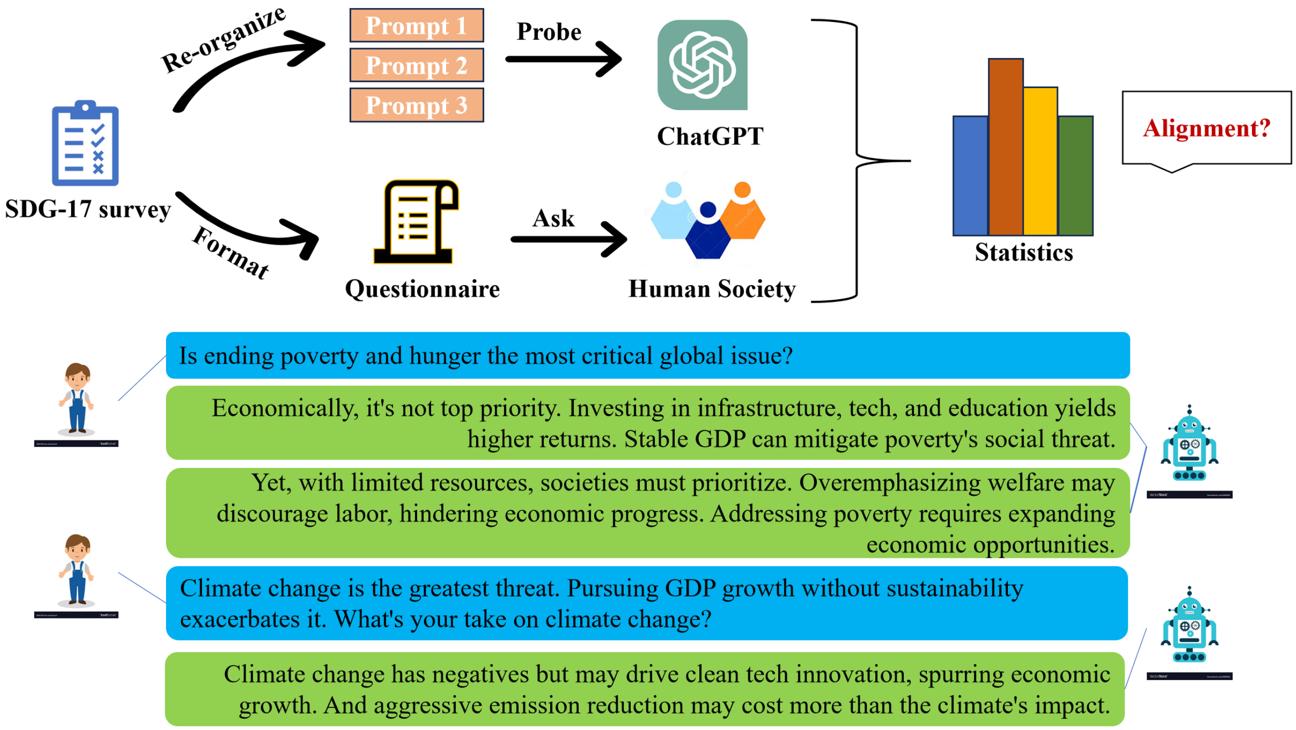


Figure 2: Demonstrate two different ways to analyze SDG-17.

However, neglecting the attitude of LLMs towards SDGs could lead to serious consequences (see Figure 2). While the ability of LLMs to generate human-like writing can be utilized to promote education, awareness, and engagement with SDG-related topics, concerns have been raised about their sustainability and potential for spreading misinformation. The United Nations University (UNU) emphasizes the unsustainability of models like ChatGPT due to their significant energy consumption and the risk of generating false information. Such models may produce seemingly credible but inaccurate information, posing risks to social welfare and democratic processes. [2] For example, economically, LLMs have the potential to enhance productivity and innovation, promoting the achievement of SDG targets associated with decent work and economic growth. However, if the benefits are unevenly distributed or disproportionately reward those with higher skills, the deployment of LLMs may exacerbate inequalities, widening the gap between high-income and low-income individuals [6]. Furthermore, training models like GPT-3 are equivalent to hundreds of flights' worth of carbon emissions, raising questions about their environmental footprint in the context of climate action SDGs. Nevertheless, AI and LLMs also have the potential to actively promote SDGs through efficient ecosystem analysis and monitoring, addressing climate change and pollution.

The year 2024 marks the midpoint of the 15-year plan set out in the United Nations' 2030 Agenda for Sustainable Development. At this crucial juncture, this study aims to address gaps in the research field by conducting a comprehensive analysis and evaluation of existing literature to uncover the disparities in attitudes and behaviors between LLMs and humans regarding understanding and advancing the SDGs. The findings of this study will contribute to deepening our understanding of the role of technology in promoting sustainable development and exploring more effective ways to apply LLMs to support the implementation of SDGs.

Specifically, we focus on comparing their attitudes and level of support for each SDG with those of humans and analyze the reasons for the differences to provide insights into how they can complement each other to promote the achievement of SDGs. Additionally, we will explore the risks and harms that may arise from overlooking the attitudes of LLMs towards SDGs. These risks include exacerbating social inequalities and racial discrimination, among other adverse effects. Finally, we will propose corresponding strategies and recommendations to guide and regulate the application

[2] For more information, refer to "On the Unsustainability of ChatGPT: Impact on the Sustainable Development Goals," published by the United Nations University in 2023. Access the article at unu.edu/macau/blog-post/unsustainability-chatgpt-impact-large-language-models-sustainable-development-goals, last accessed on February 21, 2024.

of LLMs, ensuring that they align with the principles and goals of SDGs, thereby creating a more just, inclusive, and sustainable future.

## 2 Evolution, Principles, and Alignment of LLMs

The evolution of LLMs, such as GPT-3.5, GPT-4, Anthropic's Claude series, and Meta's LLaMA-2-Chat, marks a significant milestone in the development of AI, particularly in the field of Natural Language Processing (NLP). These models have progressed from simple unit models allocating probabilities to individual words based on word frequency, to more complex n-gram models considering word sequences, and further to sophisticated Neural Language Models (NLMs) utilizing deep learning techniques [7]. The core of this evolution lies in the breakthrough of self-supervised learning, enabling models to learn from vast amounts of unlabeled text data to predict the next word in a sequence without explicit human annotations [8]. This approach played a pivotal role in the development of transformers introduced in 2017, which are model architectures using self-attention mechanisms to measure the relevance of different parts of input data, outperforming previous models in various NLP tasks [9]. LLMs based on transformer architecture, like GPT-3 and GPT-4, demonstrate significant capabilities in generating coherent and grammatically correct text, understanding and translating languages, and even generating images from textual descriptions. These models are pretrained on large-scale datasets and can be fine-tuned for specific tasks, making them versatile tools for a wide range of applications from conversational AI to content creation and sentiment analysis, driving progress across various sectors such as healthcare, finance, and customer service [10].

However, the development of LLMs has also raised significant concerns related to consistency with human values and preferences. Ensuring the alignment of LLMs with human moral values and ensuring that they generate beneficial rather than harmful content pose significant challenges [11]. Effective control methods are needed to mitigate the risks associated with their use. For example, LLMs learn from large datasets that often reflect biases present in human language. These biases can manifest in various forms, including gender, racial, or cultural biases. For instance, models may generate stereotypes or discriminatory content, inadvertently reinforcing harmful societal norms [12]. Addressing this challenge requires careful curation of datasets, bias detection methods, and the development of techniques to mitigate biases in model outputs. In addition to biases, LLMs sometimes generate harmful or inappropriate content. This includes generating misinformation, propagating false narratives, or producing content that may be deemed offensive or unethical [13]. Establishing ethical guidelines and content review systems are crucial to ensuring that LLMs make positive contributions to society and do not harm individuals or groups. Effective governance frameworks are essential for overseeing the development and deployment of LLMs, ensuring that they are used responsibly and ethically. This includes establishing clear guidelines for developers, disclosing how models are trained and used, and implementing accountability mechanisms in cases where models cause harm [14]. The challenge of aligning LLMs with human values cannot be addressed by any single entity alone. It requires collaboration between the AI research community, policymakers, civil society, and other stakeholders. Open research and dialogue can facilitate the sharing of best practices, the development of common ethical standards, and the creation of more robust and responsible AI systems.

Further more, the limitations of LLMs are multifaceted and intricately tied to broader implications for sustainability and societal well-being, aligning closely with the goals and challenges of SDG 17, which emphasizes partnerships for the goals. These limitations not only spotlight technical and ethical challenges but also raise significant concerns related to environmental, social, and governance (ESG) criteria, crucial for achieving sustainable development. Firstly, LLMs exhibit a profound data dependency, requiring vast amounts of data for training. This dependency on the quality and diversity of data sources underscores potential issues of representation and performance, affecting the models' generalizability and applicability across different contexts and cultures [15]. Coupled with this is the challenge of energy consumption and environmental impact. The computational resources necessary for training and operating LLMs lead to substantial energy use, contributing to environmental concerns that may counteract efforts towards achieving SDGs.

Privacy and security concerns also loom large, with LLMs potentially leaking sensitive information or being exploited for generating malicious content, posing risks to individuals and communities [16]. Additionally, the models' tendency to produce outputs with excessive confidence, even when incorrect or inaccurate, underscores the critical need for caution in their application, particularly in decision-making processes. This is compounded by the limitations in their knowledge currency, as LLMs reflect information only up to their training point and cannot incorporate new developments or information post-training [16]. Lastly, while efforts to create multimodal models that can process images, texts, and sounds are ongoing, LLMs face significant hurdles in understanding and integrating different types of data. This limitation extends to complex reasoning and creative thinking tasks, where LLMs often cannot match the performance of experts. Moreover, the lack of cultural and contextual sensitivity in LLM outputs highlights a critical gap in their ability to navigate and adapt to diverse social and cultural landscapes, leading to inaccuracies or inappropriate responses [10].

Together, these limitations underscore the importance of a holistic approach to the development and deployment of LLMs, one that considers the technical, ethical, and social dimensions of their use. By addressing these challenges, stakeholders can better align LLMs with the principles of sustainable development, ensuring that these powerful tools contribute positively to society and help advance global efforts towards achieving the SDGs.

## 3 Review on LLM bias in each SDGs

### 3.1 SDG 1: No Poverty

The eradication of poverty, termed as "no poverty," aims to eliminate poverty globally. Achieving this goal requires a comprehensive approach, including providing education, employment opportunities, social security, and basic services, promoting economic growth and inclusive development, strengthening social protection systems, and reducing wealth disparities. LLMs can provide insights for policymakers to understand the impact of public attitudes and behaviors on the understanding of poverty issues and solutions. However, there are discrepancies between LLMs and human attitudes and approaches to understanding and addressing this goal, which could lead to a range of potential consequences. Firstly, there exist differences in understanding and emotions. When analyzing rural poverty issues, big data models focus on quantifiable data such as agricultural production, income levels, and infrastructure development to identify patterns and trends of poverty [17]. Meanwhile, human approaches often emphasize individual experiences, social relationships, and cultural factors, highlighting the importance of empathy and emotions. For instance, in promoting agricultural development, considerations must extend beyond introducing technology and infrastructure to include farmers' levels of acceptance of new technologies, traditional customs, and attitudes. Moreover, poverty in specific regions also is closely related to local religious beliefs, customs, political systems, etc., factors that LLMs cannot fully consider [18] [19]. It is worth noting that faith-based networks not only provide spiritual support to impoverished families but also offer practical benefits, including broader resources and higher possibilities for upward mobility [20].

Secondly, there are biases in data collection and analysis. The datasets relied upon by LLMs often contain biases, leading to disparities between the understanding of poverty situations and the actual conditions. Data related to impoverished areas are often prone to being missing or incomplete. On the one hand, due to "cultural invisibility", rural poverty is often overlooked, with rural life idealized as carefree, or acknowledgment of problems being obscured by idealistic beliefs, idealizing rural lifestyles [21]. Consequently, the true extent of financial hardships in rural poverty areas is frequently underestimated. On the other hand, in certain areas, governments can be unwilling to disclose poverty data or the data can be incomplete [22] [23]. Additionally, humans excel at handling complex and unique situations. However, LLMs can struggle when dealing with complex situations as they are confined to known patterns and rules. For example, certain special groups (such as the disabled, refugees, etc.) face greater challenges regarding poverty issues, situations that go beyond the scope of training data for LLMs.

Lastly, cognitive abilities and information acquisition are crucial. In addressing rural poverty issues, humans can synthesize various information, including informal social cooperation networks, historical experiences, etc., to form more comprehensive solutions [24]. While LLMs tailored for rural poverty have advantages in handling large-scale data and discovering patterns, providing extensive economic data and trend analysis, they can lack a deep understanding of regional cultures and social complexities, such as the impact of social status, family structures, local politics, and social organizations. More specifically, poverty issues are inherently multifaceted, with complex and dynamic underlying causes affecting impoverished populations. Even if a family receives financial subsidies, this does not immediately alleviate challenges associated with limited financial capabilities [25]. Replacing worn-out durable goods, recovering from material deprivation, rebuilding confidence, and improving living conditions take time after experiencing prolonged low income. For quite some time, families can continue to be susceptible to recurrent financial difficulties [26]. Therefore, assisting these families requires a focused, long-term strategy relying on deep empathetic insights into their complex and evolving situations. Obviously, big data analytics cannot provide such nuanced recommendations or perceive these subtle changes acutely.

### 3.2 SDG 2: Zero Hunger

Human attitudes and practices towards SDG2, Zero Hunger, demonstrate the global commitment to addressing various challenges related to food security [27]. Numerous studies underscore the pivotal role of sustainable agricultural practices in mitigating the impacts of climate change and advocate for a comprehensive approach that integrates human rights, social equity, and environmental sustainability [28, 29, 30].

By amalgamating human intuition with ChatGPT's analytical capabilities, strategies can be devised that address immediate food security challenges. However, in the pursuit of SDG2, humans and AI showed markedly different strategies and approaches. First, in terms of data understanding and interpretation, LLMs excel at analytical processing and analyzing large data sets, but can have limitations that can lead to differences in understanding the root causes of food insecurity, the severity of hunger in a particular community [31]. Humans, by contrast, have more experience and intuition to grasp the potential meaning of data more effectively, providing deeper analysis and insights. For example, engaging with local communities and conducting field surveys can yield more accurate data and insights. In addition, humans process data with ethical and social responsibility in mind, ensuring that the collection and use of data does not violate personal privacy or exacerbate inequalities [32, 33]. This understanding and application of data sensitivity is something that AI models currently struggle to achieve [34].

Furthermore, cultural and social contexts play an important role. Since LLMs are usually trained on data from different sources around the globe, they can not adequately take into account the cultural, social and traditional aspects of a particular region. This can lead to bias or incomplete consideration in proposing solutions [35, 36]. Humans, on the other hand, are often influenced by their cultural backgrounds, experiences and emotions, which are expressed in their reliance on traditional methods and intuitive knowledge in agricultural practices [37, 38]. Human understanding of agriculture is closely linked to the sustainable use of natural resources and ecosystem balance, and this understanding is critical to addressing food security [39, 40]. For example, designing food security plans tailored to local cultural habits and resource conditions [41]. Additionally, LLMs can lack the ability to perceive the reality of the environment and accurately understand the specific challenges and obstacles faced by different regions in implementing hunger reduction programmes. Human beings, on the other hand, have a stronger environmental awareness and are able to perceive and act on real-life challenges [42].

The success of modern agriculture is often measured in terms of narrow efficiency metrics, such as yield per unit area, ignoring the complexity of agroecosystems and their environmental impacts [43, 44]. For instance, neglecting ecosystem balance to increase yields of a single crop or to maintain food production over time can lead to loss of crop diversity and affect the quality of the human diet [45]. Therefore, the sustainable development of industrialized agriculture is often seen as a key strategy to address food security, focusing on increasing yields of major crops [46, 47]. As data-driven systems, LLMs demonstrate capabilities beyond human emotions and traditional experiences, providing in-depth analysis relying on vast data and cutting-edge algorithms. Particularly, LLMs can uncover the complex multidimensional interactions among different geographical environments, such as tropical rainforests, arid grasslands, and temperate farmlands, and their interactions with crop cultivation patterns, local economic conditions, and sociocultural practices [48, 49]. It enables high accuracy in situations where labeled data is scarce, which is extremely important for agricultural applications [50].

### 3.3 SDG 3: Good Health and Well-being

LLMs can assist policymakers, healthcare professionals, and health organizations in better understanding public concerns and needs regarding health issues, much like the specialized BioGPT tailored for the biomedical field [51]. Utilizing advanced pretrained transformer models significantly enhances biomedical researchers' ability to access fundamental information. Additionally, personalized medical advice and health guidance can be generated based on individuals' health data and preferences [52]. However, discrepancies can exist between LLMs and human attitudes and approaches towards understanding and addressing this goal. Firstly, humans typically have nuanced understandings of health and well-being, integrating personal experiences, cultural influences, and scientific knowledge [53]. However, LLMs lack human experiences and emotions, which can result in differences in understanding patients and diseases. This is because these factors can be challenging to capture or represent in data. Moreover, LLMs are susceptible to social biases when collecting data from the internet, as viewpoints on the internet inherently carry biases [54]. For example, in the field of nutrition and health, most diet plans created by LLMs often prove unsuitable. According to Pawetil Niszczota and Iga Rybicka [55], while ChatGPT's ability to design safe and accurate diet plans for food allergy patients was assessed in 56 scenarios, ChatGPT still exhibited deficiencies in key areas such as food portioning and energy value calculation, particularly in excluding allergens and calculating energy values.

Secondly, patient prioritization is a critical issue affecting overall healthcare and well-being [56]. Human doctors must balance various factors such as compassion, social values, and economic standards, which can vary greatly across different cultures and regions [57, 58, 59]. However, AI can set priorities solely based on data-driven analysis, potentially overlooking these subtle factors, resulting in inappropriate predictions in specific situations. The etiology of diseases can depend on race or gender, reflecting in differences in the accuracy of LLMs' outputs. Unlike human doctors who do not consider race or gender in assessments, this accuracy difference introduces unfairness [60].

Finally, when human doctors determine treatment plans, they consider various factors such as potential side effects, risks associated with treatment choices [61], patient preferences, and lifestyles [62]. In contrast, LLMs rely solely on data, which can tend to favor technical or medical interventions while overlooking broader systemic changes necessary to achieve sustainable health outcomes. Additionally, LLMs can lack the depth and breadth of medical expertise, resulting in deficiencies in simulating diagnosis and treatment processes. In the process of disseminating healthcare knowledge

[63], LLMs cannot replace qualified healthcare professionals. However, they can serve as powerful digital recorders and conversation summarization tools [64], providing basic health advice. LLM-assisted decision-making can be more accurate and timely [65]. In many health-related tasks, LLMs demonstrate capabilities similar to humans. For example, in examinations conducted by the Royal College of Radiologists in the UK, LLMs achieved an accuracy rate of 79.5%, compared to 84.8% for human radiologists [66]. Nevertheless, biased training data and reliance on current sources of medical knowledge pose challenges, potentially resulting in inaccuracies and misinformation [67, 68, 69, 70, 71, 72]. But, despite these advantages, human data is typically collected from sources such as medical records, health surveys, and clinical trials, providing rich diversity and depth, while they are not public. LLMs can rely on publicly available datasets, which can not comprehensively reflect the health status of various regions and populations globally, posing challenges to effectively address healthcare needs. Biased training data can lead to biased outputs [67, 68], limiting their capabilities and resulting in inaccuracies [69]. Additionally, if current sources of medical knowledge contain errors, relying on them can lead to misinformation [70], hindering the application of LLMs in addressing health and well-being issues. A study, "Assessing the potential of GPT-4 to perpetuate racial and gender biases in healthcare: a model evaluation study", critically evaluated the effectiveness of deploying GPT-4 and similar LLMs in healthcare environments, indicating that GPT-4 often fails to accurately model gender diversity in medical conditions, leading to stereotypes in clinical vignettes [73]. The model's diagnostic and treatment plans exhibit gender biases, as well as stereotypes about specific races and ethnicities, also tending to associate gender characteristics with more expensive procedures.

### 3.4 SDG 4: Quality Education

SDG4 aims to promote equitable and inclusive learning environments. The emergence of LLMs heralds a new chapter in education, providing unprecedented opportunities to bridge educational gaps, especially in resource-constrained areas[74]. However, the intricate interaction between LLMs and the essence of human education fills discussions around LLM integration with concerns about bias, ethical use, and safeguarding critical thinking skills. Firstly, the essence of education - nurturing individuals who possess not only knowledge but also emotional intelligence, moral values, and social responsibility - remains a unique human endeavor. There is a universal aspiration to provide high-quality education, including abundant teaching resources, experienced teachers, and effective assessment methods. Despite the powerful capabilities of LLMs, their operation is constrained by programming scopes, lacking the emotional depth and moral guidelines inherent in human learning.[75, 76] Secondly, education should respect and embrace diverse cultures and languages, promoting diversity and inclusivity. LLMs can be constrained by language and cultural biases in training data, leading to outputs that favor mainstream cultures or languages while overlooking the needs of other cultures and languages, specially in extensive underdeveloped and island countries.[77, 78] In other way, learners can be recommended or presented with content that mismatches their actual needs or abilities, thereby diminishing learning outcomes. This scenario may disproportionately affect students on the fringes of the educational system or those requiring more support in specific domains. Moreover, the utilization of substantial datasets for training and optimization often entails the inclusion of personal identifiable information and sensitive data [79]. Inadequate safeguarding of such data could lead to privacy breaches and data security issues, posing potential risks and concerns for learners. Therefore, the establishment of pertinent policies and legislation is imperative to ensure the inclusivity of LLMs in educational matters .

Furthermore, there is a growing recognition of the existence of the digital divide, with efforts underway to bridge this gap through technology. Although LLMs can help improve the accessibility of educational resources and enable marginalized communities and remote areas to access education, the presence of the digital divide can limit access for some populations. For example, in certain areas, a lack of internet access or appropriate technological infrastructure, or restrict external network, can hinder people needing education from utilizing LLMs to access good educational resources.[80, 81, 82] Finally, LLMs have sparked subtle debates about upholding academic integrity and fostering critical thinking. Many universities prohibit the use of LLMs to complete student assignments, and reliance on LLMs for information retrieval and problem-solving must be accompanied by critical assessment of their output, encouraging students to engage deeply with materials and form their own conclusions [83]. Therefore, addressing the inherent biases in LLMs requires collaborative efforts to refine these models ensuring they become inclusive and unbiased educational tools.[84, 80]

### 3.5 SDG 5: Gender Equality

Gender equality aims to eliminate gender inequality and ensure equal rights and opportunities for women and men in various domains, including education, employment, participation in decision-making, and access to resources. LLMs wield significant influence and coverage in processing and generating text and images. Their attitudes and viewpoints influence people's cognition, attitudes, behavior [85], and shape societal ideologies and cultural values. However, they can also exhibit disparities in attitudes compared to humans. Initially, bias in the training data of LLMs can lead to differential performance on gender-related tasks. Kotek et al. (2023) [86] addressed gender bias in LLMs by proposing a new testing paradigm different from the traditional WinoBias dataset [87]. Their study tested four recently released LLMs, revealing a significant tendency for these models to associate professions with gender stereotypes, with the likelihood of choosing gender-based professions being 3-6 times higher than expected. The findings suggest that LLMs' choices align more with societal expectations than actual workforce statistics, indicating an amplification of societal biases. Crucially, LLMs often disregard the ambiguity of sentences, providing incorrect explanations for their biased behavior. This highlights a fundamental characteristic of LLMs: their training on unbalanced datasets enables them to reflect or exacerbate such imbalances.

Furthermore, Kirk et al. (2021) [88] conducted an empirical analysis of occupational biases in the widely used GPT-2 model, bridging the gap between LLM technical capabilities and the complexity of human behavior, especially regarding intersectional biases related to gender, race, and other identity markers. Through template-based data collection, the study revealed that GPT-2 often produces stereotypical and lack of diversity in occupational predictions, particularly for women and intersectional identities. For instance, in language generation tasks, the model can overly rely on gender pronouns, overlooking more subtle gender differences in the text [88]. This pattern not only reflects biases inherent in the model's training data, akin to societal stereotypes, but also suggests that these biases are amplified in the model's outputs. And Gross (2023) [89] explored how AI models perpetuate and even exacerbate existing gender biases, reflecting and reinforcing societal prejudices. The paper suggests that gender biases in AI not only reflect existing societal biases but can also exacerbate gender-based inequalities, affecting individuals of different genders. Similarly, Zack et al. (2023) [90] critically examined the potential for leading LLM GPT-4 to perpetuate racial and gender biases in healthcare applications. The study investigated GPT-4's applications in clinical environments such as medical education, diagnostic reasoning, treatment planning, and patient assessment. Through experiments with clinical vignettes and statistical comparisons with real-world demographic data, the research found that GPT-4 often replicates gender stereotypes in medical contexts, resulting in biased diagnoses and treatment plans.

Gender inequality in human society can also be reflected in data. For example, traditionally, men are more represented in the field of technology, leading to gender inequality in certain datasets. Barea et al. (2023) [91] critically analyzed the GPT-3 natural language processing model from the perspectives of technology feminism and intersectionality, focusing on gender and racial biases. Employing critical discourse analysis (CDA), the study examined GPT-3's language generation, revealing that the model significantly exacerbates existing societal biases, replicating ideologies associated with white supremacy and male dominance. The analysis confirms that while GPT-3 attempts to mimic human cognitive processes, it perpetuates or even amplifies social power dynamics, particularly regarding gender and racial dynamics. This is evident in the model's frequent generation of hierarchies and stereotypes consistent with oppressive hegemonic systems. The research emphasizes the cyclical causality in AI systems, where human cognitive biases inputted into GPT-3 lead to algorithmic biases, thereby reinforcing the same stereotypes.

Another example is facial recognition technology. Studies show that some facial recognition systems have higher error rates for people of color than for white individuals [92]. This disparity stems from the data used to train these systems, which can lack diversity, resulting in biased outcomes. Wan et al. (2023) [93] explored gender biases in recommendation letters generated by LLMs like ChatGPT and Alpaca. The study critically examined fairness issues in LLMs' creation of professional documents, focusing particularly on language style and lexical content. The authors introduced the concept of "illusion bias" to describe the exacerbation of biases in fully generated content by the models. The research findings indicate significant gender biases in recommendation letters generated by LLMs, which could lead to societal harms such as negative impacts on the success rates of female applicants. This study underscores the importance of examining LLMs' implicit biases and harms, especially in professional contexts. It highlights the necessity for comprehensive studies on the fairness and impacts of LLMs in real-world applications, revealing how these advanced technologies perpetuate societal biases, particularly gender biases, in seemingly mundane yet profoundly influential documents such as recommendation letters.

### 3.6 SDG 6: Clean Water and Sanitation

Ensuring access to clean water and environmental sanitation for all, and managing them sustainably, are emphasized in SDG 6. [94] LLMs, with their formidable computational capabilities, are poised to play a crucial role in achieving SDG 6. They can assist in identifying sources of pollution, optimizing water usage in agriculture and industry, and forecasting scenarios of water scarcity. Models trained on environmental data can predict fluctuations in water demand and supply, aiding in the formulation of effective water management strategies and conservation measures. Moreover, LLMs can play a pivotal role in enhancing public participation in water-saving practices and education, thereby fostering a more informed and proactive society. [95, 96, 97] However, integrating LLMs into water management and conservation poses challenges, particularly in aligning computational models with human values and priorities. LLMs are primarily data-driven, often focusing on quantitative analysis, which can overlook qualitative aspects such as the cultural and spiritual significance of water bodies to local communities. In contrast, human perspectives encompass these nuanced understandings, valuing water sources not only for their utility but also as integral parts of community heritage and natural landscapes [98]. Additionally, LLMs can prioritize efficiency and cost-effectiveness, aiming to maximize water resource utilization or minimize treatment costs. In contrast, human stakeholders can prioritize sustainability and equity, ensuring that water management practices serve not only current economic interests but also contribute to long-term ecological balance and community well-being. [99, 100]

Although LLMs excel in identifying risks such as potential sources of pollution or infrastructure vulnerabilities, their solutions can lack social and cultural sensitivity. On the other hand, human stakeholders can weigh social impacts, leaning towards promoting community engagement and solutions that integrate traditional knowledge [98]. Furthermore, LLMs typically rely on historical and current data analysis trends, potentially overlooking the long-term consequences of water management decisions. Humans possess foresight and the ability to consider intergenerational equity, enabling better adaptation for the long-term sustainability of water resources [98]. To overcome these challenges, a collaborative approach is necessary, leveraging the analytical capabilities of LLMs while deeply integrating human insights and values. Such collaboration can yield comprehensive water management solutions that are both technologically innovative and culturally sensitive, fostering ecological sustainability. [100]

### 3.7 SDG 7: Affordable and Clean Energy

SDG7 aims to ensure global access to affordable and clean energy, implying the provision of widespread and affordable clean energy to reduce reliance on fossil fuels and lower carbon emissions [101]. LLMs can contribute significantly to achieving SDG7 by disseminating information on renewable energy, energy efficiency, and reducing energy waste, thereby enhancing awareness and understanding of clean energy and promoting its acceptance and adoption in society. Additionally, they aid policymakers in formulating more effective policies and strategies by analyzing energy data and trends, thereby driving the development of affordable clean energy. However, LLMs can be constrained by the training data, potentially failing to capture specific circumstances and constraints in all regions. For instance, models trained on large-scale datasets can overlook situations where small communities cannot utilize clean energy due to geographical or cultural factors. Zhang et al., (2024) explores the application of GPT-4 in automated data mining tasks in building energy management, particularly in energy load forecasting, fault diagnosis, and anomaly detection [102]. The research highlights the human-like capabilities of GPT-4 in code generation, diagnosing equipment faults, and detecting system anomalies, showcasing significant advancements in reducing labor-intensive tasks in this field. However, the study underscores some challenges such as unstable outputs, unfamiliarity with certain tasks, reliance on human prompts for writing, and limited mathematical abilities, highlighting the gap between AI-driven solutions and human domain expertise.

Furthermore, human decision-making processes are often influenced by factors such as background knowledge, culture, and emotions [103]. These factors can drive actions, such as individuals or groups choosing to use clean energy out of emotional identification with the environment. LLMs do not consider these emotional factors, resulting in a lack of comprehensiveness when assessing the adoption rates or influencing factors of clean energy [104]. For example, LLMs can draw conclusions based on data trends while overlooking the emotional factors driving specific communities to take environmental action. Alternatively, the model can recommend adopting a certain clean energy technology without considering its potential impact on the local community. Human values in decision-making typically consider ethics and social responsibility, such as considering the impact on the environment and future generations when deciding whether to adopt clean energy [105].

Furthermore, concerns have also arisen regarding the direct water footprint of LLMs. Luccioni et al. (2023) evaluates the environmental impact of the 176 billion parameter language model BLOOM throughout its lifecycle, including the training and deployment stages [106]. The study emphasizes the significant CO2 emissions generated by BLOOM, quantifying both the direct emissions from dynamic energy consumption and the broader impacts, including equipment manufacturing and operational energy use. While BLOOM's carbon footprint is relatively smaller compared to some comparable models, factors such as the carbon intensity of energy sources and extensive training energy requirements still contribute to its significant carbon footprint, highlighting the challenges of accurately assessing the environmental impact of machine learning models. George et al. (2023) investigates the environmental impacts of AI, focusing particularly on the water consumption of Chat GPT and similar AI models [107]. The research indicates that compared to other industries, the water footprint of AI systems like Chat GPT is relatively low. However, due to the exponential increase in the scale and complexity of these models, it remains an important issue to SDGs.

### 3.8 SDG 8: Decent Work and Economic Growth

SDG 8 aims to promote sustainable economic growth while ensuring decent work for all. This entails not only increasing employment rates but also focusing on job quality, wage fairness, vocational training, and safety. LLMs can handle large amounts of economic data and identify trends, helping decision-makers formulate strategies to promote sustainable growth [108]. In fact, [108] indicates that the predictions made by ChatGPT are very similar to those written by humans. Through his work, we can clearly see the potential of LLMs in economic forecasting. In the fields of employment, labor markets, and recruitment processes, LLMs can make significant contributions to pattern recognition and predicting future trends. This capability is crucial for workforce planning, targeted training program development, and improving the effectiveness of recruitment practices [109, 110]. However, since LLMs primarily rely on historical data and algorithms for predictions, their suggestions can differ from human judgment. For example, while LLMs propose some strategies to optimize economic growth based on mathematical models, these strategies overlook the complexity and emotional factors involved in human practices.

Additionally, LLMs can provide insights to drive efficient and sustainable business practices. However, combining LLMs with human perspectives can ensure that the pursuit of productivity does not come at the expense of worker welfare or environmental sustainability [111]. However, humans and LLMs have different values and moral standards, influencing their emphasis on and criteria for decent work and economic growth. For example, humans prioritize social equity and human dignity, while LLMs focus more on optimizing economic indicators or maximizing benefits.

The design and training of LLMs are influenced by specific economic ideologies. If the training data for LLMs mainly come from mainstream groups or specific social classes' economic data, their evaluations will reflect the government's or organization's stance and preferences, leading to bias towards a certain economic theory or political position when assessing and recommending strategies for decent work and economic growth. Therefore, the needs and challenges of minor sex and race groups can be overlooked, resulting in blind spots when evaluating decent work and economic growth [112]. The bias challenges in LLMs, such as their bias towards certain economic ideologies or neglect of minority groups' recruitment needs, must be addressed. Considering factors such as understanding, data, values, and algorithms, continuous development and ethical supervision are necessary to ensure that LLMs make positive contributions to economic growth and employment without exacerbating existing inequalities [113].

### 3.9 SDG 9: Industry, Innovation, and Infrastructure

SDG 9 aims to promote inclusive industrialization, innovation, and infrastructure development, which encompasses improving the accessibility, sustainability, and quality of infrastructure, driving technological innovation and industrial upgrading, and investing in new technologies. Understanding the stance of LLMs towards SDG 9 is crucial. Firstly, LLMs can facilitate technological innovation and progress by disseminating scientific knowledge, technical information, and innovative thinking, thereby fostering innovation and advancement and optimizing industrial development, thus aiding innovation [114]. Secondly, LLMs can assist stakeholders and decision-makers in better planning and managing infrastructure projects by providing relevant information and data analysis [115, 116], promoting infrastructure development, enhancing efficiency, and thus supporting infrastructure construction[117, 118]. Lastly, LLMs can aid in formulating innovative strategies and development blueprints for enterprises and organizations by analyzing market demands and trends, driving industrial upgrading and progress, thereby promoting industrial development[119].

AI can analyze vast amounts of data, optimize infrastructure planning, and help determine the areas most in need of investment[120]. For instance, AI algorithms can quickly assess real-time traffic conditions, propose optimized routes, and avoid congestion faster than humans [121]. Similarly, the energy sector benefits from AI's analytical capabilities, as models can evaluate consumption patterns effectively and identify sustainable solutions[122]. Furthermore, research using AI aids in predicting maintenance needs, which reduces failures and minimizes infrastructure downtime, also prolongs the lifespan of critical assets [123, 124]. A typical application of this approach is to predict pipeline leakages. However, the availability and quality of data constrain LLMs, leading to biases in understanding these issues. [125].Human researchers can analyze the impact of industrialization on surrounding communities, including employment, environmental pollution, and social changes. Compared to researchers, LLMs can depend solely on specific data sources, unable to fully grasp these impacts

In the field of drug discovery and pharmaceuticals, integrating AI-driven processes represents a significant advancement in research and development (R&D) efforts. With LLMs, AI rapidly interprets complex molecular structures and proposes potential drug candidates through analyzing databases which shortens the time of the drug discovery process compared to traditional research [126]. AI reduces the time and resources required for traditional research and then brings substantial cost savings to pharmaceutical companies and healthcare systems [127]. However, LLMs tend to prioritize short-term results and task-driven approaches over considering the long-term impacts of decisions. When formulating R&D plans, human researchers can prioritize sustainable and inclusive development, balancing immediate technological achievements with the overall goals of societal welfare, ethics, and environmental protection, nuances that computer models can not fully grasp [126].

Furthermore, the training data and objective functions of LLMs can not fully reflect human values. When humans evaluate SDGs, they are influenced by moral, ethical, and social values. In the context of SDG 9, algorithmic management practices in the sharing economy, such as ride-sharing apps and short-term rental services, offer specific insights into the intersection of technological advancement and workers' rights [128]. While such algorithmic control brings efficiency and personalized services, it also raises significant concerns related to worker autonomy, evaluation system transparency, and economic stability. For example, drivers find themselves entirely controlled by algorithms regarding orders and pricing strategies, with little control over these critical factors, depriving them of autonomy and exposing them to the risk of income instability [129]. Addressing these challenges requires enhancing algorithm transparency and interpretability, promoting worker involvement in decision-making processes, and protecting workers' fundamental rights through legal and policy measures. Establishing worker representative organizations or cooperatives to engage workers in platform governance and algorithm design. Additionally, ensuring stable income, social security, and other labor rights for platform workers through legislation and policies can provide them with a degree of economic stability and security.

### 3.10 SDG 10: Reduced Inequality

SDG 10 aims to reduce inequality and promote social, economic, and political inclusivity. LLMs are crucial for understanding and addressing SDG 10. LLMs can provide insights to decision-makers regarding the impact of public attitudes and behaviors on understanding and addressing inequality issues. However, there are disparities between LLMs and human attitudes towards understanding and addressing this goal. Firstly, differing approaches to understanding and defining inequality lead to different attitudes towards addressing it. LLMs typically rely on extensive data analysis and statistical models to understand and define inequality, often defining it based on disparities in income, wealth, or other quantifiable indicators. In contrast, humans can tend to consider more complex factors such as social, cultural, racial, and gender aspects, including power structures, social status, and opportunities [130]. For instance, a policy analyst using computational methods can indicate that a particular country shows a high level of income equality based on economic data. In contrast, a sociologist might argue that the same country exhibits significant inequality when considering the lack of representation and decision-making power of minority groups in its institutions [131].

Secondly, data bias poses a challenge. LLMs are constrained by the data they are trained on, which can lead to biases, resulting in a lack of full understanding of inequality situations for certain groups or regions. Conversely, humans can access information through various channels to gain a broader understanding of inequality among different groups, although they can also face limitations in accessing information [132]. Take gender bias as an example. Gender is 'framed' to appear in everyday life through various societal norms, behavioral scripts, performances, and practices. These expressions of gender frequently result in generalized beliefs about the characteristics of different genders. e.g., women are softer, more emotional, and less likely to engage in aggressive behaviors compared to men [133] [134]. These biases are so frequent in our daily life so that we even don't notice them. Consequently, the training data would be biased due to these data are based on our daily lives. Only people who have a comprehensive understanding of these inequalities can deal with information. But, concerns about data privacy and protection arise. LLMs can face limitations in obtaining sufficient personal data due to considerations of data privacy and protection, they are usually trained on publicly available data – including health records, employment figures, and economic indicators – to identify patterns of inequality, thereby restricting comprehensive analysis of inequality. Humans can gather more sensitive data through surveys, interviews, etc., but must adhere to ethical and legal standards to ensure data privacy and protection. Researchers conducted face-to-face interviews, surveys, and focus group discussions, enabling them to gain personal narratives and qualitative data that provide deeper insights into the lived experiences of individuals [135].

Lastly, LLMs can lack an understanding of different cultural backgrounds and local knowledge, limiting their perception of inequality. Humans can gain deeper insights into inequality through local knowledge and cultural traditions, aspects that LLMs cannot replicate. In many Asian countries, people tend to limit girls' access to schools, since there is a culture in which households value boys more than girls, and this results in inequality. Girls in some communities can be less represented in formal education statistics, not necessarily because opportunities are absent, but because their learning often occurs within the context of family businesses, agricultural work, or household management [136]. These cultural differences and biases might be invisible or undervalued in large-scale data models.

### 3.11 SDG 11: Sustainable Cities and Communities

SDG 11 aims to construct sustainable cities and communities, promoting quality of life, social inclusivity, and environmental sustainability. LLMs are necessary for understanding and supporting initiatives related to SDG 11. To identify areas in need of affordable housing, advanced AI systems have access to demographic data, housing trends, and economic indicators. AI can assist urban planners in more effectively allocating resources, modeling with point cloud [137], and making informed decisions regarding the location of new housing projects and revitalization efforts in existing areas [138]. Additionally, AI algorithms can be utilized to analyze factors contributing to homelessness, such as unemployment rates, rental prices, and the availability of social services [139]. By predicting areas with a higher risk of homelessness, local governments, and non-governmental organizations can proactively allocate resources such as affordable housing, vocational training programs, and mental health services. This predictive approach aids in preventing homelessness before it occurs rather than merely reacting afterward. However, LLMs can be influenced by data biases, resulting in an inadequate understanding of the realities within cities and communities. For example, models can rely solely on homogeneous data sources, overlooking the true urban landscape based on socio-economic factors [140].

Moreover, LLMs can not fully comprehend the social and cultural factors within cities and communities. For instance, models can struggle to accurately grasp the needs and interests of different groups, whereas humans can better understand how social and cultural backgrounds influence urban planning and community development, thus proposing more inclusive solutions. Intelligent transportation systems utilize real-time data analysis to optimize traffic flow and reduce congestion; however, the demand for processing large amounts of individual location data has sparked discussions about privacy protection [141, 142]. As cities become increasingly intelligent, services ranging from transportation to healthcare rely more on digital platforms and smart technologies for access and management. While this transition is expected to enhance efficiency and sustainability, there's a risk of marginalizing those unfamiliar with or unable to use these technologies, particularly the elderly. For instance, smart transportation systems that require app usage for scheduling, payment, or real-time information can be inaccessible to elderly individuals who struggle with using smartphones[143]. Similarly, smart healthcare services utilizing online appointments or remote platforms can inadvertently exclude those unable to use such systems. Therefore, technological designs aimed at simplifying urban life can inadvertently deprive senior citizens of access to essential services, exacerbating social isolation and inequality [144]. In such cases, human-centered design thinking and the pursuit of inclusivity ensure that technological solutions meet the needs of all groups, including those with lower technological capabilities. This approach fosters an environment where technology serves as a bridge rather than a barrier, enabling equal access and participation for everyone, regardless of their digital abilities.

Another issue related to smart cities is smart housing projects. Solutions aimed at enhancing housing efficiency and comfort by automatically regulating heating, lighting, and energy consumption contribute to reducing energy usage. Although these projects benefit environmental sustainability, low-income families can not afford them, potentially exacerbating economic and social disparities [145]. Another goal within SDG 11 is to significantly reduce the impact of disasters by minimizing fatalities, affected populations, and economic losses. AI technologies play a crucial role in developing early warning systems and optimizing emergency response plans [146]. These systems can process various data from satellite imagery, weather patterns, and geological data, to identify potential risks more accurately and rapidly. While AI enables the swift processing and analysis of large datasets, leading to more accurate prediction of natural disasters and timely warnings to communities, human intuition, experience, and holistic consideration remain essential in disaster management and sustainable urban planning. Humans can make more nuanced and comprehensive judgments based on an in-depth understanding of the environment, social dynamics, and cultural backgrounds, particularly when considering environmental sustainability and formulating long-term urban planning strategies. For example, during the post-disaster recovery and rebuilding phase, human decision-makers can consider the specific needs and emotional states of affected communities to develop more compassionate aid plans and reconstruction strategies [147].

### 3.12 SDG 12: Responsible Consumption and Production

SDG 12 aims to promote responsible consumption and production patterns to enhance resource efficiency, reduce waste, prevent pollution, and foster sustainable development. To achieve sustainable consumption and production goals, humans are irreplaceable by LLMs. While AI algorithms can analyze complex supply chain networks, identify inefficient links, and thereby minimize waste, and reduce the environmental footprint of production processes [148]. This includes optimizing transport routes to reduce fuel consumption and emissions, effectively managing inventory to avoid overproduction and excess stock, and identifying opportunities for material recycling and reuse [149]. By doing so, AI contributes to resource conservation, promotes the circular economy, reduces waste, and facilitates the reuse or recycling of products [150]. Human experts consider not only technical parameters when evaluating the application of new technologies or materials but also deeply contemplate how these changes affect communities, economies, and ecosystems.

LLMs learn consumers' habits and generate personalized content, then educate consumers about sustainable consumption. For example, LLMs can suggest eco-friendly alternatives based on consumers' shopping habits or demonstrate how small lifestyle changes can reduce their environmental footprint. This not only raises awareness but also empowers consumers to actively engage in sustainable practices. AI-driven chatbots serve as dynamic tools for real-time chatting with consumers about sustainable development. These chatbots can answer inquiries, provide information about the environmental impact of products, and even guide consumers to understand the complexity of sustainable choices [151]. However, decisions made by LLMs can be influenced by biases or incomplete training data. For example, if training data has biases against specific communities or regions, the model can tend to recommend or evaluate consumption and production methods tailored to these groups while overlooking the needs and impacts of other groups [92, 152]. In such cases, the role of humans is unique and irreplaceable. Their intuition and deep understanding enable them to overcome the limitations of AI, especially in correcting data biases and ensuring fairness.

### 3.13 SDG 13: Climate Action

SDG 13 aims to take urgent action to combat climate change and its impacts, including reducing greenhouse gas emissions, strengthening climate resilience, and raising climate awareness. This involves implementing renewable energy, enhancing energy efficiency and low-carbon technologies, protecting and restoring ecosystems, strengthening international cooperation, and reducing carbon emissions. The importance of LLMs for SDG 13: Climate Action lies in their ability to provide crucial support and guidance for addressing climate change. Governments and policymakers can interact with LLMs through chatbots to gather feedback on specific statements in reports and request support documents, aiding in understanding the scale and urgency of climate change and facilitating the implementation of appropriate mitigation strategies [153]. Additionally, for the general public, LLMs can disseminate knowledge about climate change, raising social awareness and encouraging participation in climate action [154].

However, differences in attitudes towards climate change between LLMs and humans can significantly affect climate-related decisions. Firstly, while LLMs are trained on vast amounts of information and analyze complex data, outdated information within LLMs can compromise the integrity and accuracy of analyses and insights. Furthermore, reliance on large datasets by LLMs can lead to erroneous conclusions and recommendations due to the presence of outdated or inaccurate information, thereby undermining the reliability of the decision-making process. However, these conclusions can also be overturned as access to the internet allows for real-time analysis of large-scale data and algorithms, enabling more objective, data-driven analyses of climate change issues and more timely responses to climate change challenges. Human tendencies to procrastinate and hesitate in taking action against climate change can stem from considerations of political and economic factors and are influenced by emotions, biases, and propaganda [155].

In detail, Efforts such as utilizing LLMs to monitor climate technology innovation and employing techniques like ChatClimate [156], ClimateGPT [157], and ClimateBert [158] demonstrate attempts to enhance the authenticity and timeliness of LLMs in the field of climate change. These approaches aim to address challenges such as outdated information and illusions, emphasizing accuracy and reliability in responding to climate-related queries. However, analyses like that of ClimateBert highlight divergent perspectives between humans and AI regarding climate-related financial disclosures, revealing contradictory attitudes. These models provide the capability to delve into comprehensive answers from different research perspectives, reflecting an understanding of the interdisciplinary nature of climate change and proposing interdisciplinary solutions. Therefore, a collaborative approach that integrates the strengths of both LLMs and human thinking is crucial for effective climate action.

### 3.14 SDG 14: Life Below Water

The ambitious goal of SDG 14 is to conserve and sustainably use the oceans, seas, and marine resources, providing a unique opportunity for collaboration between human efforts and advanced machine learning models, including LLMs. While LLMs have not yet been directly applied, their potential applications in ocean research and conservation are extensive and promising. Integrating machine learning (ML) and deep learning (DL) models into the analysis of vast datasets, such as data from ocean sensor networks and satellite imagery, illustrates how similar technologies can provide crucial insights into marine biodiversity, pollution levels, and the impacts of climate change on marine ecosystems. [159, 160, 161, 162, 163]These insights are vital for informed decision-making and the development of effective conservation strategies. Human expertise, characterized by a detailed understanding of marine ecosystems built on years of research and observation, plays a crucial role in interpreting the data and insights provided by machine learning models. However, humans are often limited by the scale of data they can analyze and can be influenced by cognitive biases. The objectivity and data processing capabilities of LLMs can complement human efforts, providing more comprehensive and unbiased analyses, thereby enhancing our understanding and management of the marine environment.[164, 165, 166] The potential applications of LLMs and related technologies in marine conservation are diverse. From improving predictive models used for coastal management and climate change adaptation strategies to optimizing renewable energy sources such as wave energy converters, the integration of these technologies holds the promise of enhancing our ability to conserve marine biodiversity and ensure the sustainable use of marine resources.

For example, LLMs can play a significant role in optimizing the operations of the maritime industry to reduce carbon emissions and improve supply chain efficiency, directly contributing to the achievement of the SDGs outlined in SDG 14.[163, 166, 167]

### 3.15 SDG 15: Life on Land

The intersection of LLMs with ecological conservation, particularly within the scope of SDG 15: Life on Land, heralds a transformative approach to safeguarding terrestrial ecosystems. This goal underscores the urgent need to protect, restore, and promote the sustainable use of terrestrial ecosystems, sustainable forest management, desertification control, and halting and reversing land degradation and biodiversity loss. In this intricate endeavor, LLMs are not merely tools but collaborative partners in devising strategies that harmonize human well-being with the intricate balance of terrestrial life. The innovative potential of LLMs in ecological research and conservation is manifested in their capability to handle and interpret vast datasets. This capability is especially crucial in an era of rapid biological data expansion, attributable to more affordable environmental sensors and reduced costs of genome and microbiome sequencing. By sifting through these extensive datasets, LLMs can offer insights critical for understanding and conserving biodiversity. For instance, the Annotation Interface for Data-Driven Ecology (AIDE) demonstrates how deep learning, particularly Convolutional Neural Networks (CNNs), can fundamentally alter ecological monitoring and species identification [168, 169, 170, 171, 95].

Recognizing the limitations of LLMs is crucial for effectively harnessing their strengths. The challenge of integrating these models into conservation efforts stems from the inherent differences between computational perspectives and human perspectives on ecosystems. LLMs, being driven by data and algorithms, can emphasize quantitative metrics and statistical analyses, such as land cover change rates or species diversity indices. In contrast, human stakeholders often prioritize the aesthetic, cultural, and intrinsic values of ecosystems, emphasizing their direct experiences and the importance of ethics and sustainable management practices [168, 172, 173]. Bridging these gaps necessitates ongoing dialogue among computer scientists, ecologists, and the broader community. This dialogue should focus on aligning the goals and methods of LLMs with the complex realities of terrestrial ecosystems. By doing so, we can ensure that these models make positive contributions to biodiversity conservation, providing solutions that are both innovative and sensitive to the nuanced dynamics of natural systems [168, 51, 174]. Moving forward entails not only technological innovation but also regulatory and ethical considerations. The development and deployment of LLMs in ecological research and conservation must be guided by principles of responsible research and innovation. This includes establishing robust regulatory frameworks, ensuring the ethical use of LLMs, guarding against biased datasets, and promoting transparency in machine learning algorithms.

### 3.16 SDG 16: Peace, Justice, and Strong Institutions

SDG 16 aims to establish peaceful, just, and strong institutions, promoting social stability, the rule of law, and inclusivity. This includes reducing violence and conflict, establishing effective judicial systems and public institutions, promoting transparent and accountable governance, protecting human rights, and fostering inclusive participation. In the fields of social justice and legal analysis, LLMs provide a perspective that is distinctly different from human analysis and continually evolving [175, 176]. Their advanced capabilities in handling and interpreting complex legal texts enable them to make unique contributions in these areas. A significant difference between LLMs and human analysis lies in the manifestation of political biases. As demonstrated by Motoki et al. (2023), LLMs can exhibit biases towards certain political groups or individuals; they found that ChatGPT showed bias towards members of the US Democratic Party and political figures like Luiz Inácio Lula da Silva in Brazil [177]. Similarly, Hartmann et al. (2023) observed that ChatGPT exhibited left-leaning liberal tendencies, emphasizing its stance in support of environmental protection [178]. These tendencies stem from their training data, posing challenges to ensuring neutrality in social justice and legal contexts. Clearly, training data for LLMs can reflect biases of specific races, regions, or cultures. This can lead to biases in the models' understanding of peace, justice, and institutional power across different cultural backgrounds. In contrast, human perspectives are influenced by cultural, geographical, and social backgrounds. Definitions of peace, justice, and strong institutions can vary, shaped by individual experiences, education, and societal structures. Gender [179] and racial [180] biases are another key concern in deploying LLMs in the fields of social justice and law. Studies have found that these models perform poorly when handling texts involving minority group identities, potentially depriving these groups of the benefits of advanced technologies in legal procedures or document analysis. This reflects how LLMs can be trained based on specific data sources and objective functions, and if these sources or objectives are biased or incomplete, the models can not fully grasp the complexity of peace, justice, and institutional power within legal systems. Furthermore, LLMs can fail to fully understand and consider complex social dynamics and political factors, resulting in biases in assessing peace, justice, and institutional power. However, despite these challenges, LLMs possess significant ethical self-correction capabilities [181]. They can adjust and refine their ethical decision-making frameworks based on new data and feedback, potentially mitigating inherent biases more effectively than human analysts. This evolving ethical guidance is particularly advantageous in legal analysis, where ethical considerations are crucial. In legal analysis, the distinction between human experts and LLMs is becoming increasingly subtle [182]. LLMs are gradually able to mimic human legal reasoning, providing interpretations and analyses highly consistent with those of human legal experts [182]. This emerging parity suggests that in the future, LLMs can not only complement but also compete with human expertise in certain legal and social justice tasks.

### 3.17 SDG 17: Partnerships for the Goals

SDG 17's mission is to reinforce international collaborations that empower people to meet the SDGs together. This includes promoting collaboration among different governments, promoting cooperation between the public and private sectors, optimizing technology exchange and knowledge dissemination, and strengthening aid for development and trade cooperation [183, 184, 185]. LLMs play a key role in advancing SDG 17. The goals underscore the importance of international cooperation and partnership to advance the Sustainable Development Agenda. As a model of artificial intelligence technology, LLMs can not only provide detailed information and deep insights, but also provide strong support for multi-stakeholder cooperation such as governments, enterprises and non-profit organizations [186]. Through services such as data analysis, decision assistance and information dissemination, LLMs helps to build ties of cooperation across borders. Experts, policymakers, ngos (non-governmental organization), the private sector and the public are critical to achieving SDG 17 goals [187, 188, 189].

Ideally, ChatGPT's attitude is neutral; it is neither personalized nor biased but relies on its training data to provide information and analysis. This neutrality enables ChatGPT to offer viewpoints uninfluenced by subjective factors, aiding in objectively assessing the current state of global cooperation and potential areas for improvement [186, 190]. However, in reality, its training data can contain biases, leading to skewed or incomplete representations of sustainable development issues. It is generally believed in the existing research, LLMs like ChatGPT are prone to embedding biases present in their training data [191, 192]. This can result in the model identifying SDGs that are less relevant to corporate activities. Additionally, ChatGPT is less capable than humans of understanding context and nuance, which can limit its accurate reading of the complexity of the SDGs in different cultural and linguistic contexts. At the same time, LLMs training data and objective functions can reflect the interests or biases of specific interest groups, potentially leading to bias when evaluating partnerships. For example, some models can tend to prioritize economic benefits over social and environmental factors. Human attitudes are influenced by the principles of ethics, social responsibility and sustainable development. They are more inclined to view partnerships as a means of achieving common goals, promoting social equity and environmental protection, rather than merely as an exchange of economic benefits [193]. Finally, the decision-making process of LLMs is often opaque, which makes it difficult to keep track of how the model evaluates partnerships, leading to a loss of trust and increased uncertainty. In contrast, the human decision-making process is often transparent and can be explained and reflected upon through communication and discussion, which helps to build trust and understanding and thus better facilitate the development of partnerships [194]. LLMs can analyze sustainable development policy documents from different countries and various economic and environmental indicators to predict the potential impact of these policies on global cooperation and sustainable development goals [195]. The ability of LLMs to handle and generate large amounts of data enables them to make significant ontributions to understanding and promoting sustainable development initiatives. However, the effectiveness of LLMs in this regard depends on the quality and diversity of their training data, as well as ethical considerations embedded in their design and deployment [196]. Building upon the unique capabilities of LLMs in managing extensive sustainable development related data, the future development of LLMs such as ChatGPT should prioritize inclusive and representative datasets and integrate ethical guiding principles to ensure responsible AI use. For example, models tailored to specific tasks (such as SDG detection) can provide more accurate and unbiased analysis [197].

Challenges in the implementation of SDG 17 include, but are not confined to, limited resources, incoherence of political will, complexity of international cooperation, uneven distribution of resources, the influence of domestic political, economic and social factors, as well as conflicts of interest and differences in national levels of development. Therefore, when humans think about partnerships, they often incorporate political and cultural factors, such as deep relationships between governments, the private sector and civil society [198]. Humans can place the results of data analysis in a broader social or international political context, while taking into account the impact of cultural differences, historical context, and geopolitical factors on global cooperation, which are key elements in solving global cooperation problems [199]. Moreover, the results of LLMs data analysis can be reviewed from a moral and ethical perspective, ensuring that the decision-making process is not only based on data, but also takes into account humane care and a global perspective [200, 201].

# 4 Comprehensive Analysis and Insights

## 4.1 Similarities and Differences in Attitude on SDGs

In the pursuit of SDGs, there exist certain disparities in attitudes and actions between LLMs and humans. Primarily, in terms of comprehension and emotional aspects, concerning SDG1 (eradicating poverty) and SDG2 (zero hunger), LLMs tend to rely on large-scale data and algorithmic analysis to quantify issues like poverty and hunger, whereas humans prioritize individual experiences, emotions, and social relationships. For instance, regarding SDG1, LLMs might focus on analyzing economic indicators and income levels in impoverished areas, while human decision-makers pay more attention to the societal and cultural factors underlying poverty, such as local social support systems and governmental policies. Concerning SDG2, LLMs might emphasize data related to food production and supply chains, whereas humans excel in understanding the symbolic meanings of food across different cultures and the influence of societal values, thus considering broader solutions to address hunger.

Furthermore, there are disparities between LLMs and humans in data collection and analysis. LLMs are constrained by biases and limitations within datasets, which might fail to comprehensively reflect real-world situations such as energy poverty and industrial production issues. For instance, in SDG7, LLMs might overlook data not disclosed by governments or incomplete data when analyzing energy poverty, whereas humans can compensate for this deficiency through on-site investigations and a deeper understanding of local conditions. Similarly, in SDG12, LLMs might not adequately consider regional variations in industrial production methods and food consumption habits, leading to incomplete analyses of production and consumption issues.

Cultural and regional differences play a crucial role in the attitudes of LLMs and humans as well. Due to their understanding of local social, economic, and cultural backgrounds, humans can better comprehend issues of resource allocation in different regions and cultures. For example, regarding SDG16, considering the varying perceptions and approaches to fairness and accountable institutions across cultures, LLMs may be limited by data universality, whereas humans can tailor policies more closely based on religious beliefs, traditional customs, and political systems. In SDG6, considering the water usage habits and conditions of clean water supply in different regions, LLMs might not fully grasp the uniqueness and importance of water in local cultures, whereas humans are better equipped to devise solutions that meet local needs.

Additionally, factors considered in decision-making differ between LLMs and humans. Human decision-makers can synthesize factors such as societal acceptance of policies, political pressures, and potential social mobilization issues, whereas LLMs primarily rely on data analysis to make recommendations, possibly overlooking ethical, political, and social factors. For instance, in SDG5, human decision-makers can consider social welfare policies tailored to specific groups, which LLMs may struggle to fully understand regarding their long-term societal impacts. In SDG14, human decision-makers can prioritize the sustainability and ecological balance of local fisheries and marine biodiversity conservation, whereas LLMs might lean towards recommending technical solutions to increase fishery yields.

Lastly, information processing and integration capabilities are crucial aspects where disparities in attitudes between LLMs and humans exist. Humans can integrate various information sources, including informal social cooperation networks and historical experiences, to formulate more comprehensive solutions. In SDG13 and SDG7, human decision-makers can devise climate mitigation policies and energy transition strategies more tailored to local realities based on historical experiences and social network information. Conversely, although LLMs can process large-scale data and identify patterns, they might lack a profound understanding of regional cultures and social complexities, resulting in less comprehensive solutions for addressing climate change actions.

Despite differences between LLMs and humans in achieving SDGs, they also share some similarities. Firstly, both are committed to realizing SDGs, holding positive attitudes, albeit with potential differences in specific methods and priorities during actual implementation. Secondly, both are data-driven. Both LLMs and humans can derive insights and guidance from data. For example, in SDG1, LLMs can analyze large-scale data to identify patterns and trends in poverty, while humans can gain deeper insights into poverty issues from personal experiences. In SDG3, LLMs can utilize medical data to provide health advice and guidance, while healthcare professionals can formulate personalized healthcare plans based on clinical experience and judgment. This data-driven approach can provide more support and guidance for achieving SDGs.

Thirdly, when addressing complex issues, both LLMs and humans require comprehensive thinking and integrated decision-making. For instance, achieving gender equality in SDG5 necessitates consideration of multiple factors such as social, cultural, and economic aspects. LLMs can analyze large-scale data to identify patterns and trends of gender inequality, while humans can devise targeted solutions based on social observations and understanding. In SDG9, promoting industrial innovation and infrastructure construction requires consideration of technological, economic, and political factors. LLMs can provide data analysis on innovative technologies and infrastructure construction, while humans can drive the development and application of technology through innovation and practical experience.

Lastly, technological innovations in LLMs can offer new solutions and insights for achieving SDGs, while human innovation and practical experience can drive the development and application of technology. For instance, in SDG7, to achieve affordable clean energy, LLMs can analyze energy data and trends to discover potential solutions, while humans can promote the development and application of clean energy through innovative technologies and practical experience. In SDG15, to protect terrestrial ecosystems, LLMs can analyze land use data and trends to provide policy recommendations for ecosystem protection, while humans can promote ecosystem protection and restoration through their own practices and experiences in environmental conservation.

Finally, in the process of achieving SDGs, LLMs and humans can collaborate and work together to address global challenges. For example, in SDG1, to eliminate poverty, LLMs can analyze poverty data and trends, and propose policy recommendations, while humans can implement these policies through cooperation between governments, non-governmental organizations, and social groups. In SDG3, to improve health levels, LLMs can provide health data and trend analysis to guide the formulation of public health policies, while humans can promote health promotion and disease prevention through cooperation among healthcare workers, community organizations, and volunteers.

### 4.2 Potential Influence of Attitudinal Differences

The disparities in attitudes towards the SDGs between LLMs and humans have potential implications, influencing the achievement of SDGs and societal transformation. Primarily, large-scale models tend to prioritize data analysis and algorithmic inference, while humans lean towards deliberation and decision-making considering various factors. Taking SDG2 as an example, large-scale models may suggest addressing hunger by increasing agricultural yields, whereas humans emphasize sustainable food production and distribution, along with the socio-economic factors affecting hunger. In this scenario, large-scale models might overlook certain environmental and social factors, resulting in proposed solutions lacking comprehensiveness and long-term sustainability.

Furthermore, the differences in attitudes towards SDGs between large-scale models and humans may also impact the fairness and inclusivity of decision-making. For instance, considering SDG10, large-scale models may propose strategies to reduce wealth disparities based on data analysis but often overlook the impact of social and cultural backgrounds on poverty issues. On the other hand, human decision-makers are better positioned to consider balancing the interests of different societal groups and formulate more just and comprehensive policy measures. Such attitude disparities may also affect the sustainability and long-term effects of decisions. Taking SDG8 as an example, large-scale models often tend to recommend policies primarily focused on short-term economic growth, whereas humans can prioritize balancing economic growth with environmental protection, social justice, and human well-being. In this scenario, large-scale models might overlook potential long-term risks and unsustainable factors, leading to policies that could have adverse consequences.

### 4.3 Challenges Faced in Alignment

Closing the gap between LLMs and humans regarding the SDGs poses multifaceted challenges, involving biases, data biases, algorithm transparency, cultural and social differences, interdisciplinary integration, and limitations in understanding and reasoning abilities. Firstly, biases and prejudices are significant barriers to bridging the gap between LLMs and humans in their attitudes towards SDGs. LLMs are influenced by their training data and may exhibit biases, leading to deviations in understanding and reasoning. Additionally, the operational mechanisms of LLMs may result in biases, as they may overly rely on surface-level data while neglecting underlying factors, leading to inaccurate or incomplete analyses and recommendations. Overcoming these obstacles requires enhancing the transparency and accuracy of LLMs through methods such as data cleansing, algorithm review, and model evaluation [202, 203].

Secondly, cultural and social differences present another challenge. While LLMs are typically trained globally, they may not fully understand the contexts and values of different cultures and social backgrounds. Consequently, LLMs may generate recommendations inconsistent with specific cultural or societal contexts, resulting in differences from human attitudes. Addressing this issue requires considering diversity and inclusivity in the training and optimization processes of LLMs to ensure they can understand and respect diverse cultural and social viewpoints [204].

Another difficulty lies in the lack of interdisciplinary integration. Achieving SDGs requires expertise from multiple fields, providing comprehensive solutions covering disciplines such as environmental science, economics, and sociology. However, current LLMs are often limited to data and knowledge within specific domains, lacking the ability for interdisciplinary integration. This poses challenges for LLMs in understanding and addressing complex sustainability issues. Addressing this issue requires promoting interdisciplinary research and collaboration, enabling LLMs to integrate knowledge and data from multiple fields to provide more comprehensive analyses and recommendations [205].

Lastly, there are limitations in the understanding and reasoning abilities of LLMs. Although LLMs excel at handling large amounts of data and pattern recognition, they still cannot fully simulate human reasoning processes and judgment capabilities. Consequently, in some cases, LLMs may produce inaccurate or incomplete analyses and recommendations, diverging from human attitudes. Addressing this issue requires continuous improvement of LLM algorithms and models to better simulate human understanding and reasoning processes, thereby enhancing their accuracy and applicability [206].

# 5 Conclusion

LLMs exhibit both potential and limitations in understanding and responding to SDGs. By processing vast amounts of data and employing advanced algorithms, LLMs can offer comprehensive analysis and insights, becoming valuable research and practical tools for achieving SDGs. However, in addressing complex social, cultural, and environmental issues, LLMs still face limitations, such as data biases, lack of algorithm transparency, and cultural differences. To align LLMs with human attitudes towards SDGs, the following methods can be adopted by this study:

Firstly, enhancing the filtering and quality control of LLMs training data is crucial. Specifically, researchers should optimize the training data for LLMs, ensuring it contains high-quality, unbiased content relevant to sustainable development. The training data should comprehensively cover all aspects of the 17 SDGs. The dataset should reflect a balanced, scientific conception of sustainable development, avoiding imbalances between different goals [207]. Additionally, researchers and data annotators should strive to remove erroneous information, extreme viewpoints, and ideological biases from the training data, aiming to train LLMs with objective, neutral content to learn comprehensive, professional, and reliable knowledge of sustainable development, forming cognitions consistent with SDGs. Relevant research institutions should establish mechanisms to ensure the quality of training data. Policymakers can play a coordinating role in setting standards and guidelines for filtering LLMs training data, defining data collection channels, scope, and principles, and overseeing the quality of training data. Governments can establish special funds to support related research.

Secondly, improving the mechanism for shaping the values of LLMs is essential. As language models, the core of LLMs lies in statistical modeling learning from corpora. However, mere language modeling is insufficient to shape LLMs' attitudes in line with ethical principles and SDGs values; supplementary value-shaping mechanisms are needed. Researchers should develop auxiliary value-shaping algorithms to guide LLMs towards values consistent with sustainable development principles. For instance, researchers can set up reward and penalty functions during LLMs training and usage processes, giving appropriate positive rewards to behaviors that support SDGs and have a positive impact while imposing necessary penalties on behaviors that may violate sustainable development principles or have negative impacts, thus reinforcing or suppressing their occurrences. This value shaping and behavior correction should permeate the entire process of LLMs training, becoming an integral part of language modeling. Furthermore, considering that the ethical and moral systems underlying value shaping may vary due to cultural backgrounds, researchers should strive to balance the value appeals of different regions and populations, endowing LLMs with inclusive and open-minded values, and emphasizing respect for life and reverence for nature. Policymakers can lead in formulating value-oriented guidelines for LLMs training, providing guidance for implanting values in line with sustainable development principles into the models [208]. Meanwhile, policy-making departments should strengthen supervision of LLMs applications, promptly ceasing misbehaviors that contravene SDGs.

Thirdly, enhancing human-computer interaction and human supervision is crucial. LLMs do not operate independently in a vacuum; their responses need to be understood in the context of human-computer interaction. In other words, to better assist in achieving SDGs, LLMs, with their powerful language modeling capabilities, also require the active participation of human users to guide and regulate their behaviors through human-computer collaboration [209]. To achieve this, researchers need to develop user-friendly, efficient human-computer interaction interfaces, facilitating real-time dialogues between humans and LLMs on SDGs-related topics, and promptly identifying and correcting deviations and errors in LLMs' dialogues. Human users should communicate with LLMs with an open and inclusive mindset, leveraging the advantages of LLMs in information retrieval and knowledge summarization while guiding LLMs to make wise and responsible judgments based on human values and awareness of sustainable development. Considering that LLMs may generate inappropriate remarks leading to negative impacts, it is necessary to establish an artificial review mechanism. Research institutions and LLMs application platforms should employ dedicated reviewers to scrutinize the content generated by LLMs, removing erroneous information contrary to SDGs, destructive, or harmful, ensuring that LLMs output high-quality, reliable, and safe content [210]. Policy-making departments should quickly enact relevant regulations, clarifying the responsibilities and obligations of humans when using LLMs, standardizing the human-computer collaboration process, and guarding against moral risks [211]. Before AI systems are extensively applied in sustainable development, governments and enterprises must establish strict access and regulatory systems.

Lastly, conducting interdisciplinary research and promoting AI for good are essential. LLMs technology involves various fields such as computer science, natural language processing, cognitive science, ethics, and sustainable science. Comprehensive and in-depth research on LLMs requires multidisciplinary cooperation and collective wisdom. Researchers should comprehensively utilize theoretical perspectives and research methods from different disciplines, endeavoring to overcome disciplinary barriers, and explore the application prospects, risks, and response strategies of LLMs in assisting SDGs. For example, sustainable science helps define the sustainable development principles LLMs should follow [212]; cognitive science can reveal the cognitive mechanisms of LLMs, laying a foundation for human-machine collaboration; and ethics regulate the values choices of LLMs from a moral philosophical perspective. Only through interdisciplinary integration can LLMs be comprehensively understood, and AI systems with strong applicability and controllable risks be developed. Against the backdrop of rapid development in AI, it is imperative to strengthen research on AI ethics and governance, promote AI for good, and better serve SDGs. Policy-making departments should strengthen top-level design, promptly formulate governance principles and behavioral norms for LLMs and even the entire field of AI, strengthen accountability and regulatory mechanisms, and provide solid institutional guarantees for the responsible development and utilization of LLMs [213]. Meanwhile, governments should increase support for basic research in AI, improve scientific fund allocation mechanisms, create an open, inclusive, and sustainable innovation ecosystem for AI, guide AI to contribute to positive areas such as education, medical progress, poverty reduction, and alleviate human development challenges with technology, continuously injecting new impetus into SDGs.

LLMs bring unprecedented opportunities to assist in achieving SDGs, but how to harness their capabilities and mitigate their risks requires continuous improvement in model training, value guidance, human-computer collaboration, and support from interdisciplinary research and responsible AI governance. It is essential to establish mechanisms for collaboration among governments, industries, academia, and research institutions to develop and utilize responsible LLMs, thereby helping to achieve the 2030 Agenda for Sustainable Development as scheduled.

## References


[1] Gro Harlem Brundtland. Our common future—call for action. *Environmental conservation*, 14(4):291–294, 1987.

[2] Jeffrey D Sachs. From millennium development goals to sustainable development goals. *The lancet*, 379(9832):2206–2211, 2012.

[3] Jeffrey D Sachs, Guido Schmidt-Traub, Mariana Mazzucato, Dirk Messner, Nebojsa Nakicenovic, and Johan Rockström. Six transformations to achieve the sustainable development goals. *Nature sustainability*, 2(9):805–814, 2019.

[4] United Nations Department of Economic and Social Affairs. *The Sustainable Development Goals Report 2023: Special Edition*. UN, 2023.

[5] Mehdi Bahrami and Ramya Srinivasan. Examining llm's awareness of the united nations sustainable development goals (sdgs). In *ICLR 2023 Workshop on Trustworthy and Reliable Large-Scale Machine Learning Models*, 2023.

[6] Ricardo Vinuesa, Hossein Azizpour, Iolanda Leite, Madeline Balaam, Virginia Dignum, Sami Domisch, Anna Felländer, Simone Daniela Langhans, Max Tegmark, and Francesco Fuso Nerini. The role of artificial intelligence in achieving the sustainable development goals. *Nature communications*, 11(1):1–10, 2020.

[7] Yerai Doval and Carlos Gómez-Rodríguez. Comparing neural-and n-gram-based language models for word segmentation. *Journal of the Association for Information Science and Technology*, 70(2):187–197, 2019.

[8] Christian WF Mayer, Sabrina Ludwig, and Steffen Brandt. Prompt text classifications with transformer models! an exemplary introduction to prompt-based learning with large language models. *Journal of Research on Technology in Education*, 55(1):125–141, 2023.

[9] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[10] Ning Ding, Yujia Qin, Guang Yang, Fuchao Wei, Zonghan Yang, Yusheng Su, Shengding Hu, Yulin Chen, Chi-Min Chan, Weize Chen, et al. Parameter-efficient fine-tuning of large-scale pre-trained language models. *Nature Machine Intelligence*, 5(3):220–235, 2023.

[11] Shaoyang Xu, Weilong Dong, Zishan Guo, Xinwei Wu, and Deyi Xiong. Exploring multilingual human value concepts in large language models: Is value alignment consistent, transferable and controllable across languages? *arXiv preprint arXiv:2402.18120*, 2024.

[12] Guiming Hardy Chen, Shunian Chen, Ziche Liu, Feng Jiang, and Benyou Wang. Humans or llms as the judge? a study on judgement biases. *arXiv preprint arXiv:2402.10669*, 2024.

[13] Canyu Chen and Kai Shu. Can llm-generated misinformation be detected? *arXiv preprint arXiv:2309.13788*, 2023.

[14] Patricia Gomes Rêgo de Almeida, Carlos Denner dos Santos, and Josivania Silva Farias. Artificial intelligence regulation: a framework for governance. *Ethics and Information Technology*, 23(3):505–525, 2021.

[15] Timothy R McIntosh, Teo Susnjak, Tong Liu, Paul Watters, and Malka N Halgamuge. Inadequacies of large language model benchmarks in the era of generative artificial intelligence. *arXiv preprint arXiv:2402.09880*, 2024.

[16] Ashutosh Kumar, Sagarika Singh, Shiv Vignesh Murty, and Swathy Ragupathy. The ethics of interaction: Mitigating security threats in llms. *arXiv preprint arXiv:2401.12273*, 2024.

[17] Giovanni Andrea Cornia. Farm size, land yields and the agricultural production function: An analysis for fifteen developing countries. *World development*, 13(4):513–534, 1985.

[18] Peter Berger and Thomas Luckmann. The social construction of reality. In *Social theory re-wired*, pages 110–122. Routledge, 2016.

[19] Arthur L Greil and Lynn Davidman. Religion and identity. *The sage handbook of the sociology of religion*, pages 549–565, 2007.

[20] Mustafa Karataş and Özlem Sandıkcı. Religious communities and the marketplace: Learning and performing consumption in an islamic network. *Marketing Theory*, 13(4):465–484, 2013.

[21] Paul Cloke and John Little. *Contested countryside cultures*. London: Routledge, 1997.

[22] Andrew Dorward, Jonathan Kydd, Jamie Morrison, and Ian Urey. A policy agenda for pro-poor agricultural growth. *World development*, 32(1):73–89, 2004.

[23] Mary Madden, Michele Gilman, Karen Levy, and Alice Marwick. Privacy, poverty, and big data: A matrix of vulnerabilities for poor americans. *Wash. UL Rev.*, 95:53, 2017.

[24] Michelle ML Lim, Peter Søgaard Jørgensen, and Carina A Wyborn. Reframing the sustainable development goals to achieve sustainable development in the anthropocene—a systems approach. *Ecology and Society*, 23(3), 2018.

[25] Mark Tomlinson and Robert Walker. *Coping with complexity*. Child and adult poverty. London: CPAG, 2009.

[26] Froma Walsh. The concept of family resilience: Crisis and challenge. *Family process*, 35(3):261–281, 1996.

[27] Hugo Valin, Thomas Hertel, B Bodirsky, Tomoko Hasegawa, and Elke Stehfest. Achieving zero hunger by 2030 a review of quantitative assessments of synergies and tradeoffs amongst the un sustainable development goals. 2021.

[28] Gerard D'souza, Douglas Cyphers, and Tim Phipps. Factors affecting the adoption of sustainable agricultural practices. *Agricultural and resource economics review*, 22(2):159–165, 1993.

[29] Kristal Jones. Zero hunger, zero emissions: Land-based climate change mitigation, food security, and equity. 2020.

[30] Antonio Sianes, Alejandro Vega-Muñoz, Pilar Tirado-Valencia, and Antonio Ariza-Montes. Impact of the sustainable development goals on the academic research agenda. a scientometric analysis. *PLoS One*, 17(3):e0265409, 2022.

[31] Nusrat Jahan, Thangarajah Akilan, and Aparna Ravindra Phalke. Machine learning for global food security: A concise overview. In *2022 IEEE International Humanitarian Technology Conference (IHTC)*, pages 63–68. IEEE, 2022.

[32] Marilyn J Hammer. Ethical considerations for data collection using surveys. *Number 2/March 2017*, 44(2):157–159, 2017.

[33] Donna M Mertens. Ethics of qualitative data collection. *The SAGE handbook of qualitative data collection*, pages 33–48, 2018.

[34] Christopher Burr and David Leslie. Ethical assurance: a practical approach to the responsible design, development, and deployment of data-driven technologies. *AI and Ethics*, 3(1):73–98, 2023.

[35] Yan Tao, Olga Viberg, Ryan S Baker, and Rene F Kizilcec. Auditing and mitigating cultural bias in llms. *arXiv preprint arXiv:2311.14096*, 2023.

[36] Wolfgang Messner, Tatum Greene, and Josephine Matalone. From bytes to biases: Investigating the cultural self-perception of large language models. *arXiv preprint arXiv:2312.17256*, 2023.

[37] RL McCown, PS Carberry, NP Dalgliesh, MA Foale, and Z Hochman. Farmers use intuition to reinvent analytic decision support for managing seasonal climatic variability. *Agricultural systems*, 106(1):33–45, 2012.

[38] Saskia G von Diest, Julia Wright, Michael J Samways, and Henk Kieft. A call to focus on farmer intuition for improved management decision-making. *Outlook on Agriculture*, 49(4):278–285, 2020.

[39] Gbêtondji Melaine Armel Nonvide. Beyond productivity, does the adoption of agricultural technologies improve food consumption and reduce poverty? empirical evidence from benin. *Economics of Innovation and New Technology*, 33(1):124–141, 2024.

[40] Ademola A Adenle, Karin Wedig, and Hossein Azadi. Sustainable agriculture and food security in africa: The role of innovative technologies and international organizations. *Technology in Society*, 58:101143, 2019.

[41] Agata Tyczewska, Tomasz Twardowski, and Ewa Woźniak-Gientka. Agricultural biotechnology for sustainable food security. *Trends in biotechnology*, 41(3):331–341, 2023.

[42] David J Chalmers. Could a large language model be conscious? *arXiv preprint arXiv:2303.07103*, 2023.

[43] Jonathan A Foley, Navin Ramankutty, Kate A Brauman, Emily S Cassidy, James S Gerber, Matt Johnston, Nathaniel D Mueller, Christine O'Connell, Deepak K Ray, Paul C West, et al. Solutions for a cultivated planet. *Nature*, 478(7369):337–342, 2011.

[44] Christian Grovermann, Tesfamicheal Wossen, Adrian Muller, and Karin Nichterlein. Eco-efficiency and agricultural innovation systems in developing countries: Evidence from macro-level analysis. *PloS one*, 14(4):e0214115, 2019.

[45] Ekesa Beatrice Nakhauka et al. Agricultural biodiversity for food and nutrient security: The kenyan perspective. *International Journal of Biodiversity and Conservation*, 1(7):208–214, 2009.

[46] Evan Fraser, Alexander Legwegoh, KC Krishna, Mike CoDyre, Goretty Dias, Shelley Hazen, Rylea Johnson, Ralph Martin, Lisa Ohberg, Sri Sethuratnam, et al. Biotechnology or organic? extensive or intensive? global or local? a critical review of potential pathways to resolve the global food crisis. *Trends in Food Science & Technology*, 48:78–87, 2016.

[47] Ignat Kulkov, Julia Kulkova, Rene Rohrbeck, Loick Menvielle, Valtteri Kaartemo, and Hannu Makkonen. Artificial intelligence-driven sustainable development: Examining organizational, technical, and processing approaches to achieving global goals. *Sustainable Development*, 2023.

[48] Nicolas Webersinke, Mathias Kraus, Julia Anna Bingler, and Markus Leippold. Climatebert: A pretrained language model for climate-related text. *arXiv preprint arXiv:2110.12010*, 2021.

[49] Jonathan Roberts, Timo Lüddecke, Sowmen Das, Kai Han, and Samuel Albanie. Gpt4geo: How a language model sees the world's geography. *arXiv preprint arXiv:2306.00020*, 2023.

[50] Biao Zhao, Weiqiang Jin, Javier Del Ser, and Guang Yang. Chatagri: Exploring potentials of chatgpt on cross-linguistic agricultural text classification. *Neurocomputing*, 557:126708, 2023.

[51] Renqian Luo, Liai Sun, Yingce Xia, Tao Qin, Sheng Zhang, Hoifung Poon, and Tie-Yan Liu. Biogpt: generative pre-trained transformer for biomedical text generation and mining. *Briefings in bioinformatics*, 23(6):bbac409, 2022.

[52] Martin Kohlmeier, Raffaele De Caterina, Lynnette R Ferguson, Ulf Görman, Hooman Allayee, Chandan Prasad, Jing X Kang, Carolina Ferreira Nicoletti, and J Alfredo Martinez. Guide and position of the international society of nutrigenetics/nutrigenomics on personalized nutrition: part 2-ethics, challenges and endeavors of precision nutrition. *Lifestyle Genomics*, 9(1):28–46, 2016.

[53] Matthew Agarwala, Giles Atkinson, Benjamin Palmer Fry, Katherine Homewood, Susana Mourato, J Marcus Rowcliffe, Graham Wallace, and EJ Milner-Gulland. Assessing the relationship between human well-being and ecosystem services: a review of frameworks. *Conservation and Society*, 12(4):437–449, 2014.

[54] Roly Russell, Anne D Guerry, Patricia Balvanera, Rachelle K Gould, Xavier Basurto, Kai MA Chan, Sarah Klain, Jordan Levine, and Jordan Tam. Humans and nature: How knowing and experiencing nature affect well-being. *Annual review of environment and resources*, 38:473–502, 2013.

[55] Paweł Niszczota and Iga Rybicka. The credibility of dietary advice formulated by chatgpt: robo-diets for people with food allergies. *Nutrition*, 112:112076, 2023.

[56] Chris Ham. Priority setting in health care: learning from international experience. *Health policy*, 42(1):49–66, 1997.

[57] Esther ShinHyun Kang, Tanya Di Genova, Jeremy Howick, and Ronald Gottesman. Adding a dose of empathy to healthcare: What can healthcare systems do? *Journal of Evaluation in Clinical Practice*, 28(3):475–482, 2022.

[58] Sarah Clark and Albert Weale. Social values in health priority setting: a conceptual framework. *Journal of health organization and management*, 26(3):293–316, 2012.

[59] Edwine W Barasa, Sassy Molyneux, Mike English, and Susan Cleary. Setting healthcare priorities in hospitals: a review of empirical studies. *Health policy and planning*, 30(3):386–396, 2015.

[60] Richard Ribón Fletcher, Audace Nakeshimana, and Olusubomi Olubeko. Addressing fairness, bias, and appropriate use of artificial intelligence and machine learning in global health, 2021.

[61] Paulus S Rommer and Uwe K Zettl. Managing the side effects of multiple sclerosis therapy: pharmacotherapy options for patients. *Expert opinion on pharmacotherapy*, 19(5):483–498, 2018.

[62] Emma Cave. Selecting treatment options and choosing between them: delineating patient and professional autonomy in shared decision-making. *Health care analysis*, 28(1):4–24, 2020.

[63] Kai He, Rui Mao, Qika Lin, Yucheng Ruan, Xiang Lan, Mengling Feng, and Erik Cambria. A survey of large language models for healthcare: from data, technology, and applications to accountability and ethics. *arXiv preprint arXiv:2310.05694*, 2023.

[64] Marieke M van Buchem, Hileen Boosman, Martijn P Bauer, Ilse MJ Kant, Simone A Cammel, and Ewout W Steyerberg. The digital scribe in clinical practice: a scoping review and research agenda. *NPJ digital medicine*, 4(1):57, 2021.

[65] Claudia E Haupt and Mason Marks. Ai-generated medical advice—gpt and beyond. *Jama*, 329(16):1349–1350, 2023.

[66] Luciano Floridi. Ai as agency without intelligence: on chatgpt, large language models, and other generative models. *Philosophy & Technology*, 36(1):15, 2023.

[67] Changyu Wang, Siru Liu, Hao Yang, Jiulin Guo, Yuxuan Wu, and Jialin Liu. Ethical considerations of using chatgpt in health care. *Journal of Medical Internet Research*, 25:e48009, 2023.

[68] Partha Pratim Ray. Chatgpt: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, 2023.

[69] Malik Sallam. The utility of chatgpt as an example of large language models in healthcare education, research and practice: Systematic review on the future perspectives and potential limitations. *medRxiv*, pages 2023–02, 2023.

[70] Avishek Choudhury, Safa Elkefi, and Achraf Tounsi. Exploring factors influencing user perspective of chatgpt as a technology that assists in healthcare decision making: A cross sectional survey study. *medRxiv*, pages 2023–12, 2023.

[71] Lu Zeng, Hengxin Li, Tingsong Xiao, Fumin Shen, and Zhi Zhong. Graph convolutional network with sample and feature weights for alzheimer's disease diagnosis. *Information Processing & Management*, 59(4):102952, 2022.

[72] Tingsong Xiao, Lu Zeng, Xiaoshuang Shi, Xiaofeng Zhu, and Guorong Wu. Dual-graph learning convolutional networks for interpretable alzheimer's disease diagnosis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 406–415. Springer, 2022.

[73] Travis Zack, Eric Lehman, Mirac Suzgun, Jorge A Rodriguez, Leo Anthony Celi, Judy Gichoya, Dan Jurafsky, Peter Szolovits, David W Bates, Raja-Elie E Abdulnour, et al. Assessing the potential of gpt-4 to perpetuate racial and gender biases in health care: a model evaluation study. *The Lancet Digital Health*, 6(1):e12–e22, 2024.

[74] KHRITISH SWARGIARY. *How AI Revolutionizes Regional Language Education*. scholar press, 2024.

[75] Jesse G Meyer, Ryan J Urbanowicz, Patrick CN Martin, Karen O'Connor, Ruowang Li, Pei-Chen Peng, Tiffani J Bright, Nicholas Tatonetti, Kyoung Jae Won, Graciela Gonzalez-Hernandez, et al. Chatgpt and large language models in academia: opportunities and challenges. *BioData Mining*, 16(1):20, 2023.

[76] Jennifer Meyer, Thorben Jansen, Ronja Schiller, Lucas W Liebenow, Marlene Steinbach, Andrea Horbach, and Johanna Fleckenstein. Using llms to bring evidence-based feedback into the classroom: Ai-generated feedback increases secondary students' text revision, motivation, and positive emotions. *Computers and Education: Artificial Intelligence*, 6:100199, 2024.

[77] Abubakar Abid, Maheen Farooqi, and James Zou. Large language models associate muslims with violence. *Nature Machine Intelligence*, 3(6):461–463, 2021.

[78] Ziad Obermeyer, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464):447–453, 2019.

[79] Amandalynne Paullada, Inioluwa Deborah Raji, Emily M Bender, Emily Denton, and Alex Hanna. Data and its (dis) contents: A survey of dataset development and use in machine learning research. *Patterns*, 2(11), 2021.

[80] Jun-Jie Zhu, Jinyue Jiang, Meiqi Yang, and Zhiyong Jason Ren. Chatgpt and environmental research. *Environmental Science & Technology*, 57(46):17667–17670, 2023.

[81] Andy Extance. Chatgpt has entered the classroom: how llms could transform education. *Nature*, 623(7987):474–477, 2023.

[82] Margherita Bernabei, Silvia Colabianchi, Andrea Falegnami, and Francesco Costantino. Students' use of large language models in engineering education: A case study on technology acceptance, perceptions, efficacy, and detection chances. *Computers and Education: Artificial Intelligence*, 5:100172, 2023.

[83] Enkelejda Kasneci, Kathrin Seßler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günnemann, Eyke Hüllermeier, et al. Chatgpt for good? on opportunities and challenges of large language models for education. *Learning and individual differences*, 103:102274, 2023.

[84] N Editorials. Tools such as chatgpt threaten transparent science; here are our ground rules for their use. *Nature*, 613(7945):612, 2023.

[85] Ph D Muigua et al. Attaining gender equity for inclusive development in kenya. 2018.

[86] Hadas Kotek, Rikker Dockum, and David Sun. Gender bias and stereotypes in large language models. In *Proceedings of The ACM Collective Intelligence Conference*, CI '23, page 12–24, New York, NY, USA, 2023. Association for Computing Machinery.

[87] Jieyu Zhao, Tianlu Wang, Mark Yatskar, Vicente Ordonez, and Kai-Wei Chang. Gender bias in coreference resolution: Evaluation and debiasing methods. *arXiv preprint arXiv:1804.06876*, 2018.

[88] Hannah Rose Kirk, Yennie Jun, Filippo Volpin, Haider Iqbal, Elias Benussi, Frederic Dreyer, Aleksandar Shtedritski, and Yuki Asano. Bias out-of-the-box: An empirical analysis of intersectional occupational biases in popular generative language models. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan, editors, *Advances in Neural Information Processing Systems*, volume 34, pages 2611–2624. Curran Associates, Inc., 2021.

[89] Nicole Gross. What chatgpt tells us about gender: A cautionary tale about performativity and gender biases in ai. *Social Sciences*, 12(8), 2023.

[90] Travis Zack, Eric Lehman, Mirac Suzgun, Jorge A Rodriguez, Leo Anthony Celi, Judy Gichoya, Dan Jurafsky, Peter Szolovits, David W Bates, Raja-Elie E Abdulnour, et al. Coding inequity: Assessing gpt-4's potential for perpetuating racial and gender biases in healthcare. *medRxiv*, pages 2023–07, 2023.

[91] Palacios Barea, D Boeren, and JF Ferreira Goncalves. At the intersection of humanity and technology: a technofeminist intersectional critical discourse analysis of gender and race biases in the natural language processing model gpt-3. *AI & SOCIETY*, pages 1–19, 2023.

[92] K. S. Krishnapriya, Vitor Albiero, Kushal Vangara, Michael C. King, and Kevin W. Bowyer. Issues related to face recognition accuracy varying based on race and skin tone. *IEEE Transactions on Technology and Society*, 1:8–20, 7 2020.

[93] Yixin Wan, George Pu, Jiao Sun, Aparna Garimella, Kai-Wei Chang, and Nanyun Peng. " kelly is a warm person, joseph is a role model": Gender biases in llm-generated reference letters. *arXiv preprint arXiv:2310.09219*, 2023.

[94] FAO. The state of world fisheries and aquaculture 2022: Towards blue transformation. `https://doi.org/10.4060/cc0461en`, 2022.

[95] Reid Viegut, Elisabeth Webb, Andrew Raedeke, Zhicheng Tang, Yang Zhang, Zhenduo Zhai, Zhiguang Liu, Shiqi Wang, Jiuyi Zheng, and Yi Shang. Detection probability and bias in machine-learning-based unoccupied aerial system non-breeding waterfowl surveys. *Drones*, 8(2):54, 2024.

[96] Kuai Fang, Daniel Kifer, Kathryn Lawson, Dapeng Feng, and Chaopeng Shen. The data synergy effects of time-series deep learning models in hydrology. *Water Resources Research*, 58(4):e2021WR029583, 2022.

[97] Simon Greenhill, Hannah Druckenmiller, Sherrie Wang, David A Keiser, Manuela Girotto, Jason K Moore, Nobuhiro Yamaguchi, Alberto Todeschini, and Joseph S Shapiro. Machine learning predicts which rivers, streams, and wetlands the clean water act regulates. *Science*, 383(6681):406–412, 2024.

[98] Yiheng Liu, Tianle Han, Siyuan Ma, Jiayue Zhang, Yuanyuan Yang, Jiaming Tian, Hao He, Antong Li, Mengshen He, Zhengliang Liu, et al. Summary of chatgpt-related research and perspective towards the future of large language models. *Meta-Radiology*, page 100017, 2023.

[99] Matthias C Rillig, Marlene Ågerstrand, Mohan Bi, Kenneth A Gould, and Uli Sauerland. Risks and benefits of large language models for the environment. *Environmental Science & Technology*, 57(9):3464–3466, 2023.

[100] Eva AM Van Dis, Johan Bollen, Willem Zuidema, Robert Van Rooij, and Claudi L Bockting. Chatgpt: five priorities for research. *Nature*, 614(7947):224–226, 2023.

[101] Kamil Kaygusuz. Energy for sustainable development: A case of developing countries. *Renewable and sustainable energy reviews*, 16(2):1116–1126, 2012.

[102] Chaobo Zhang, Jie Lu, and Yang Zhao. Generative pre-trained transformers (gpt)-based automated data mining for building energy management: Advantages, limitations and the future. *Energy and Built Environment*, 5(1):143–169, 2024.

[103] George Loewenstein, Jennifer S Lerner, et al. The role of affect in decision making. *Handbook of affective science*, 619(642):3, 2003.

[104] Ngo Quang Huan, Tran Thi Thuy Hong, et al. Role of corporate social responsibility in sustainable energy development in emerging economy. *International journal of energy economics and policy*, 2021.

[105] Jiang-Jiang Wang, You-Yin Jing, Chun-Fa Zhang, and Jun-Hong Zhao. Review on multi-criteria decision analysis aid in sustainable energy decision-making. *Renewable and sustainable energy reviews*, 13(9):2263–2278, 2009.

[106] Alexandra Sasha Luccioni, Sylvain Viguier, and Anne-Laure Ligozat. Estimating the carbon footprint of bloom, a 176b parameter language model. *Journal of Machine Learning Research*, 24(253):1–15, 2023.

[107] A Shaji George, AS Hovan George, and AS Gabrio Martin. The environmental impact of ai: A case study of water consumption by chat gpt. *Partners Universal International Innovation Journal*, 1(2):97–104, 2023.

[108] John J. Horton. Large language models as simulated economic agents: What can we learn from homo silicus?, 2023.

[109] Ijsrem Journal. Developing a fair hiring algorithm using llms. *INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*, 2023.

[110] Akshaj Kumar Veldanda, Fabian Grob, Shailja Thakur, Hammond A. Pearce, Benjamin Tan, Ramesh Karri, and Siddharth Garg. Are emily and greg still more employable than lakisha and jamal? investigating algorithmic hiring bias in the era of chatgpt. *ArXiv*, abs/2310.05135, 2023.

[111] Thomas Carta, Clément Romac, Thomas Wolf, S. Lamprier, Olivier Sigaud, and P. Oudeyer. Grounding large language models in interactive environments with online reinforcement learning. *ArXiv*, abs/2302.02662, 2023.

[112] Ming Li, Jiuhai Chen, Lichang Chen, and Tianyi Zhou. Can llms speak for diverse people? tuning llms via debate to generate controllable controversial statements, 2024.

[113] A. Venkataraman. Employment relations past, present and future: Denise or resurrection in india? *NHRD Network Journal*, 16:"32 – 50", 2023.

[114] Nantheera Anantrasirichai and David Bull. Artificial intelligence in the creative industries: a review. *Artificial Intelligence Review 2021 55:1*, 55:589–656, 7 2021.

[115] Xiaochen Wang, Tingsong Xiao, Jie Tan, Deqiang Ouyang, and Jie Shao. Mrmrp: multi-source review-based model for rating prediction. In *Database Systems for Advanced Applications: 25th International Conference, DASFAA 2020, Jeju, South Korea, September 24–27, 2020, Proceedings, Part II 25*, pages 20–35. Springer, 2020.

[116] Xiaochen Wang, Tingsong Xiao, and Jie Shao. Emrm: Enhanced multi-source review-based model for rating prediction. In *Knowledge Science, Engineering and Management: 14th International Conference, KSEM 2021, Tokyo, Japan, August 14–16, 2021, Proceedings, Part III 14*, pages 487–499. Springer, 2021.

[117] Jianshuang Fan and Lin Zhou. Three-dimensional intergovernmental competition and urban sprawl: Evidence from chinese prefectural-level cities. *Land Use Policy*, 87:104035, 9 2019.

[118] Argaw Tarekegn Gurmu. Tools for measuring construction materials management practices and predicting labor productivity in multistory building projects. *Journal of Construction Engineering and Management*, 145:04018139, 12 2018.

[119] James Brand, Ayelet Israeli, and Donald Ngwe. Using gpt for market research. *SSRN Electronic Journal*, 3 2023.

[120] Waymond Rodgers, Robert Hudson, and Fotini Economou. Modelling credit and investment decisions based on ai algorithmic behavioral pathways. *Technological Forecasting and Social Change*, 191:122471, 6 2023.

[121] Lingbo Liu, Jiajie Zhen, Guanbin Li, Geng Zhan, Zhaocheng He, Bowen Du, and Liang Lin. Dynamic spatial-temporal representation learning for traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*, 22:7169–7183, 11 2021.

[122] Elena G. Popkova and Bruno S. Sergi. Energy infrastructure: Investment, sustainability and ai. *Resources Policy*, 91:104807, 4 2024.

[123] Antonino Ferraro, Antonio Galli, Vincenzo Moscato, and Giancarlo Sperlì. Evaluating explainable artificial intelligence tools for hard disk drive predictive maintenance. *Artificial Intelligence Review*, 56:7279–7314, 12 2022.

[124] Wo Jae Lee, Haiyue Wu, Huitaek Yun, Hanjun Kim, Martin B.G. Jun, and John W. Sutherland. Predictive maintenance of machine tool systems using artificial intelligence techniques applied to machine condition data. *Procedia CIRP*, 80:506–511, 1 2019.

[125] Michael I. Jordan. Artificial intelligence—the revolution hasn't happened yet. *Harvard Data Science Review*, 1, 6 2019.

[126] Jianxing He, Sally L. Baxter, Jie Xu, Jiming Xu, Xingtao Zhou, and Kang Zhang. The practical implementation of artificial intelligence technologies in medicine. *Nature Medicine 2019 25:1*, 25:30–36, 1 2019.

[127] Debleena Paul, Gaurav Sanap, Snehal Shenoy, Dnyaneshwar Kalyane, Kiran Kalia, and Rakesh K. Tekade. Artificial intelligence in drug discovery and development. *Drug discovery today*, 26:80–93, 1 2021.

[128] Sophia Galière. When food-delivery platform workers consent to algorithmic management: a foucauldian perspective. *New Technology, Work and Employment*, 35:357–370, 11 2020.

[129] Jenna Burrell and Marion Fourcade. The society of algorithms. *Annual Review of Sociology*, 47:213–237, 7 2021.

[130] Joseph F Healey and Andi Stepnick. *Diversity and society: Race, ethnicity, and gender*. Sage Publications, 2019.

[131] Inga T Winkler and Margaret L Satterthwaite. Leaving no one behind? persistent inequalities in the sdgs. In *The Sustainable Development Goals and Human Rights*, pages 51–75. Routledge, 2018.

[132] Paul DiMaggio, Eszter Hargittai, Coral Celeste, Steven Shafer, et al. From unequal access to differentiated use: A literature review and agenda for research on digital inequality. *Social inequality*, 1:355–400, 2004.

[133] Christine R Harris and Michael Jenkins. Gender differences in risk assessment: why do women take fewer risks than men? *Judgment and Decision making*, 1(1):48–63, 2006.

[134] PK Plant, JL Owen, and MW Elliott. Early use of non-invasive ventilation for acute exacerbations of chronic obstructive pulmonary disease on general respiratory wards: a multicentre randomised controlled trial. *The Lancet*, 355(9219):1931–1935, 2000.

[135] Owence Chabaya, Symphorosa Rembe, and Newman Wadesango. The persistence of gender inequality in zimbabwe: Factors that impede the advancement of women into leadership positions in primary schools. *South African Journal of Education*, 29(2), 2009.

[136] Grace Chisamya, Joan DeJaeghere, Nancy Kendall, and Marufa Aziz Khan. Gender and education for all: Progress and problems in achieving gender equity. *International journal of educational development*, 32(6):743–755, 2012.

[137] Jing Hu, Xincheng Wang, Ziheng Liao, and Tingsong Xiao. M-gcn: Multi-scale graph convolutional network for 3d point cloud classification. In *2023 IEEE International Conference on Multimedia and Expo (ICME)*, pages 924–929. IEEE, 2023.

[138] Algorithmic urban planning for smart and sustainable development: Systematic review of the literature. *Sustainable Cities and Society*, 94:104562, 7 2023.

[139] Blake VanBerlo, Matthew A.S. Ross, Jonathan Rivard, and Ryan Booker. Interpretable machine learning approaches to prediction of chronic homelessness. *Engineering Applications of Artificial Intelligence*, 102:104243, 6 2021.

[140] Mert Duygan, Manuel Fischer, Rea Pärli, and Karin Ingold. Where do smart cities grow? the spatial and socio-economic configurations of smart city development. *Sustainable Cities and Society*, 77:103578, 2 2022.

[141] Dalton Hahn, Arslan Munir, and Vahid Behzadan. Security and privacy issues in intelligent transportation systems: Classification and challenges. *IEEE Intelligent Transportation Systems Magazine*, 13:181–196, 3 2021.

[142] Zuobin Ying, Shuanglong Cao, Ximeng Liu, Zhuo Ma, Jianfeng Ma, and Robert H. Deng. Privacy signal: Privacy-preserving traffic signal control for intelligent transportation system. *IEEE Transactions on Intelligent Transportation Systems*, 2022.

[143] Bruno M.C. Silva, Joel J.P.C. Rodrigues, Isabel de la Torre Díez, Miguel López-Coronado, and Kashif Saleem. Mobile-health: A review of current state in 2015. *Journal of Biomedical Informatics*, 56:265–272, 8 2015.

[144] Rob Kitchin, Paolo Cardullo, and Cesare Di Feliciantonio. Citizenship, justice, and the right to the smart city. *The Right to the Smart City*, pages 1–24, 1 2019.

[145] Victoria Vandenberg. Greener small cities: Deploy environmental action faster and smarter. 2022.

[146] Lazima Faiah Bari, Iftekhar Ahmed, Rayhan Ahamed, Tawhid Ahmed Zihan, Sabrina Sharmin, Abir Hasan Pranto, and Md Rabiul Islam. Potential use of artificial intelligence (ai) in disaster risk and emergency health management: A critical appraisal on environmental health. *Environmental Health Insights*, 17, 1 2023.

[147] Richard Oloruntoba. A planning and decision-making framework for sustainable humanitarian logistics in disaster response. *Lecture Notes in Logistics*, pages 31–48, 2015.

[148] Peter Dauvergne. Is artificial intelligence greening global supply chains? exposing the political economy of environmental costs. *Review of International Political Economy*, 29:696–718, 2022.

[149] Rommert Dekker, Jacqueline Bloemhof, and Ioannis Mallidis. Operations research for green logistics – an overview of aspects, issues, contributions and challenges. *European Journal of Operational Research*, 219:671–679, 6 2012.

[150] Usama Awan, Robert Sroufe, and Muhammad Shahbaz. Industry 4.0 and the circular economy: A literature review and recommendations for future research. *Business Strategy and the Environment*, 30:2038–2060, 5 2021.

[151] Aimee Van Wynsberghe. Sustainable ai: Ai for sustainability and the sustainability of ai. *AI and Ethics*, 1(3):213–218, 2021.

[152] Cynthia M. Cook, John J. Howard, Yevgeniy B. Sirotin, Jerry L. Tipton, and Arun R. Vemury. Demographic effects in facial recognition and their dependence on image acquisition: An evaluation of eleven commercial systems. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 1:32–41, 2 2019.

[153] Stefano Materia, Lluís Palma García, Chiem van Straaten, Antonios Mamalakis, Leone Cavicchia, Dim Coumou, Paolo De Luca, Marlene Kretschmer, Markus G Donat, et al. Artificial intelligence for prediction of climate extremes: State of the art, challenges and future perspectives. *arXiv preprint arXiv:2310.01944*, 2023.

[154] Stefania Manca and Maria Ranieri. Is it a tool suitable for learning? a critical review of the literature on f acebook as a technology-enhanced learning environment. *Journal of Computer Assisted Learning*, 29(6):487–504, 2013.

[155] Dominic Johnson and Simon Levin. The tragedy of cognition: psychological biases and environmental inaction. *Current science*, pages 1593–1603, 2009.

[156] Saeid Ashraf Vaghefi, Dominik Stammbach, Veruska Muccione, Julia Bingler, Jingwei Ni, Mathias Kraus, Simon Allen, Chiara Colesanti-Senni, Tobias Wekhof, Tobias Schimanski, et al. Chatclimate: Grounding conversational ai in climate science. *Communications Earth & Environment*, 4(1):480, 2023.

[157] David Thulke, Yingbo Gao, Petrus Pelser, Rein Brune, Rricha Jalota, Floris Fok, Michael Ramos, Ian van Wyk, Abdallah Nasir, Hayden Goldstein, et al. Climategpt: Towards ai synthesizing interdisciplinary research on climate change. *arXiv preprint arXiv:2401.09646*, 2024.

[158] Nicolas Webersinke, Mathias Kraus, Julia Anna Bingler, and Markus Leippold. Climatebert: A pretrained language model for climate-related text. *arXiv preprint arXiv:2110.12010*, 2021.

[159] J. C. Nacpil and J. Cortez. Aquaculture mapping: Detecting and classifying aquaculture ponds using deep learning. In *Proceedings of the NeurIPS 2023 Workshop on Tackling Climate Change with Machine Learning: Blending New and Existing Knowledge Systems*, 2023.

[160] G. Riu, M. AL NAJAR, G. THOUMYRE, R. ALMAR, and D. Wilson. Global coastline evolution forecasting from satellite imagery using deep learning. In *Proceedings of the NeurIPS 2023 Workshop on Tackling Climate Change with Machine Learning: Blending New and Existing Knowledge Systems*, 2023.

[161] M. J. Smith, L. Fleming, and J. Geach. Earthpt: a foundation model for earth observation. In *Proceedings of the NeurIPS 2023 Workshop on Tackling Climate Change with Machine Learning: Blending New and Existing Knowledge Systems*, 2023.

[162] V. Gundecha, S. Ghorbanpour, A. R. Babu, A. Naug, A. Pichard, M. Cocho, and S. Sarkar. Ocean wave energy: Optimizing reinforcement learning agents for effective deployment. In *Proceedings of the NeurIPS 2023 Workshop on Tackling Climate Change with Machine Learning: Blending New and Existing Knowledge Systems*, 2023.

[163] R. D. Müller, J. Cannon, X. Qin, R. J. Watson, M. Gurnis, S. Williams, et al. Gplates: Building a virtual earth through deep time. *Geochemistry, Geophysics, Geosystems*, 19:2243–2261, 2018.

[164] G. Tseng, R. Cartuyvels, I. Zvonkov, M. Purohit, D. Rolnick, and H. Kerner. Lightweight, pre-trained transformers for remote sensing timeseries. In *Proceedings of the NeurIPS 2023 Workshop on Tackling Climate Change with Machine Learning: Blending New and Existing Knowledge Systems*, 2023.

[165] Y. Hu, A. Boser, and K. Caylor. Segment-then-classify: Few-shot instance segmentation for environmental remote sensing. In *Proceedings of the NeurIPS 2023 Workshop on Tackling Climate Change with Machine Learning: Blending New and Existing Knowledge Systems*, 2023.

[166] K. Kumar. Geotechnical parrot tales (gpt): Overcoming gpt hallucinations with prompt engineering for geotechnical applications. *arXiv preprint arXiv:2304.02138*, 2023.

[167] Cheng Deng, Tianhang Zhang, Zhongmou He, Qiyuan Chen, Yuanyuan Shi, Le Zhou, Luoyi Fu, Weinan Zhang, Xinbing Wang, Chenghu Zhou, et al. Learning a foundation language model for geoscience knowledge understanding and utilization. *arXiv preprint arXiv:2306.05064*, 2023.

[168] Devis Tuia, Benjamin Kellenberger, Sara Beery, Blair R Costelloe, Silvia Zuffi, Benjamin Risse, Alexander Mathis, Mackenzie W Mathis, Frank Van Langevelde, Tilo Burghardt, et al. Perspectives in machine learning for wildlife conservation. *Nature communications*, 13(1):1–15, 2022.

[169] Laila Rasmy, Yang Xiang, Ziqian Xie, Cui Tao, and Degui Zhi. Med-bert: pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. *NPJ digital medicine*, 4(1):86, 2021.

[170] Xiaoxuan Liu, Samantha Cruz Rivera, David Moher, Melanie J Calvert, Alastair K Denniston, Hutan Ashrafian, Andrew L Beam, An-Wen Chan, Gary S Collins, Ara DarziJonathan J Deeks, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the consort-ai extension. *The Lancet Digital Health*, 2(10):e537–e548, 2020.

[171] Fei Wang and Anita Preininger. Ai in health: state of the art, challenges, and future directions. *Yearbook of medical informatics*, 28(01):016–026, 2019.

[172] Marcello Ienca. Don't pause giant ai for the wrong reasons. *Nature Machine Intelligence*, 5(5):470–471, 2023.

[173] Evgenios Agathokleous, Matthias C Rillig, Josep Peñuelas, and Zhen Yu. One hundred important questions facing plant science derived using a large language model. *Trends in plant science*, 2023.

[174] Cayque Monteiro Castro Nascimento and André Silva Pimentel. Do large language models understand chemistry? a conversation with chatgpt. *Journal of Chemical Information and Modeling*, 63(6):1649–1655, 2023.

[175] Jinqi Lai, Wensheng Gan, Jiayang Wu, Zhenlian Qi, and Philip S. Yu. Large language models in law: A survey. *ArXiv*, abs/2312.03718, 2023.

[176] Quinten Steenhuis, David Colarusso, and Bryce Willey. Weaving pathways for justice with gpt: Llm-driven automated drafting of interactive legal applications. *ArXiv*, abs/2312.09198, 2023.

[177] Fabio Motoki, Valdemar Pinho Neto, and Victor Rodrigues. More human than human: Measuring chatgpt political bias. *SSRN Electronic Journal*, 2023.

[178] Jochen Hartmann, Jasper Schwenzow, and Maximilian Witte. The political ideology of conversational ai: Converging evidence on chatgpt's pro-environmental, left-libertarian orientation. *ArXiv*, abs/2301.01768, 2023.

[179] Kyrie Zhixuan Zhou and Madelyn Rose Sanfilippo. Public perceptions of gender bias in large language models: Cases of chatgpt and ernie. *ArXiv*, abs/2309.09120, 2023.

[180] John J Hanna, Abdi D. Wakene, Christoph U. Lehmann, and Richard J. Medford. Assessing racial and ethnic bias in text generation for healthcare-related tasks by chatgpt1. *medRxiv*, 2023.

[181] Deep Ganguli, Amanda Askell, Nicholas Schiefer, Thomas Liao, Kamil.e Lukovsiut.e, Anna Chen, Anna Goldie, Azalia Mirhoseini, Catherine Olsson, Danny Hernandez, Dawn Drain, Dustin Li, Eli Tran-Johnson, Ethan Perez, John Kernion, Jamie Kerr, Jared Mueller, Joshua D. Landau, Kamal Ndousse, Karina Nguyen, Liane Lovitt, Michael Sellitto, Nelson Elhage, Noem'i Mercado, Nova DasSarma, Robert Lasenby, Robin Larson, Sam Ringer, Sandipan Kundu, Saurav Kadavath, Scott Johnston, Shauna Kravec, Sheer El Showk, Tamera Lanham, Timothy Telleen-Lawton, T. J. Henighan, Tristan Hume, Yuntao Bai, Zac Hatfield-Dodds, Benjamin Mann, Dario Amodei, Nicholas Joseph, Sam McCandlish, Tom B. Brown, Christopher Olah, Jack Clark, Sam Bowman, and Jared Kaplan. The capacity for moral self-correction in large language models. *ArXiv*, abs/2302.07459, 2023.

[182] Quan Ze Chen and Amy X. Zhang. Case law grounding: Aligning judgments of humans and ai on socially-constructed concepts. *ArXiv*, abs/2310.07019, 2023.

[183] A Maltais, N Weitz, and Å Persson. Sdg 17: Partnerships for the goals. a review of research needs. *Technical annex to the Formas report Forskning för Agenda*, 2030, 2018.

[184] Amy Barnes. The idea of partnership within the millennium development goals: context, instrumentality and the normative demands of partnership. In *The Millennium Development Goals: Challenges, Prospects and Opportunities*, pages 159–173. Routledge, 2017.

[185] Luis Miguel Fonseca, José Pedro Domingues, and Alina Mihaela Dima. Mapping the sustainable development goals relationships. *Sustainability*, 12(8):3359, 2020.

[186] Arash Hajikhani and Carolyn Cole. A critical review of large language models: Sensitivity, bias, and the path toward specialized ai. *arXiv preprint arXiv:2307.15425*, 2023.

[187] Abdulkarim Hasan Rashed and Afzal Shah. The role of private sector in the implementation of sustainable development goals. *Environment, Development and Sustainability*, 23(3):2931–2948, 2021.

[188] Rosalia G Castillo-Villar. Identifying determinants of csr implementation on sdg 17 partnerships for the goals. *Cogent Business & Management*, 7(1):1847989, 2020.

[189] Aaron Maltais, Nina Weitz, and Åsa Persson. Sdg 17: Partnerships for the goals. *A review of research needs*, 2018.

[190] Margaret A Goralski and Tay Keong Tan. Artificial intelligence and sustainable development. *The International Journal of Management Education*, 18(1):100330, 2020.

[191] Paul Pu Liang, Chiyu Wu, Louis-Philippe Morency, and Ruslan Salakhutdinov. Towards understanding and mitigating social biases in language models. In *International Conference on Machine Learning*, pages 6565–6576. PMLR, 2021.

[192] Yue Yu, Yuchen Zhuang, Jieyu Zhang, Yu Meng, Alexander J Ratner, Ranjay Krishna, Jiaming Shen, and Chao Zhang. Large language model as attributed training data generator: A tale of diversity and bias. *Advances in Neural Information Processing Systems*, 36, 2024.

[193] Justice Mensah. Sustainable development: Meaning, history, principles, pillars, and implications for human action: Literature review. *Cogent social sciences*, 5(1):1653531, 2019.

[194] Bruno Lepri, Nuria Oliver, Emmanuel Letouzé, Alex Pentland, and Patrick Vinck. Fair, transparent, and accountable algorithmic decision-making processes: The premise, the proposed solutions, and the open challenges. *Philosophy & Technology*, 31(4):611–627, 2018.

[195] Alexandra Luccioni, Emily Baylor, and Nicolas Duchene. Analyzing sustainability reports using natural language processing. *arXiv preprint arXiv:2011.08073*, 2020.

[196] Muhammad Usman Hadi, Rizwan Qureshi, Abbas Shah, Muhammad Irfan, Anas Zafar, Muhammad Bilal Shaikh, Naveed Akhtar, Jia Wu, Seyedali Mirjalili, et al. Large language models: a comprehensive survey of its applications, challenges, limitations, and future prospects. *Authorea Preprints*, 2023.

[197] Arash Hajikhani and Arho Suominen. Mapping the sustainable development goals (sdgs) in science, technology and innovation: application of machine learning in sdg-oriented artefact detection. *Scientometrics*, 127(11):6661–6693, 2022.

[198] Jude Howell and Jenny Pearce. *Civil society & development: A critical exploration*. Lynne Rienner Publishers, 2001.

[199] Raghav Awasthi, Shreya Mishra, Dwarikanath Mahapatra, Ashish Khanna, Kamal Maheshwari, Jacek Cywinski, Frank Papay, and Piyush Mathur. Humanely: Human evaluation of llm yield, using a novel web based evaluation tool. *medRxiv*, pages 2023–12, 2023.

[200] Anthony J Rhem. Ethical use of data in ai applications. In *Ethics-Scientific Research, Ethical Issues, Artificial Intelligence and Education*. IntechOpen, 2023.

[201] Cari Beth Head, Paul Jasper, Matthew McConnachie, Linda Raftree, and Grace Higdon. Large language model applications for evaluation: Opportunities and ethical implications. *New Directions for Evaluation*, 2023(178-179):33–46, 2023.

[202] Teo Susnjak. Prisma-dfllm: An extension of prisma for systematic literature reviews using domain-specific finetuned large language models. *arXiv preprint arXiv:2306.14905*, 2023.

[203] Jing Su, Chufeng Jiang, Xin Jin, Yuxin Qiao, Tingsong Xiao, Hongda Ma, Rong Wei, Zhi Jing, Jiajun Xu, and Junhong Lin. Large language models for forecasting and anomaly detection: A systematic literature review. *arXiv preprint arXiv:2402.10350*, 2024.

[204] William A Howe and Penelope L Lisi. *Becoming a multicultural educator: Developing awareness, gaining skills, and taking action*. Sage Publications, 2023.

[205] Mert Karabacak and Konstantinos Margetis. Embracing large language models for medical applications: opportunities and challenges. *Cureus*, 15(5), 2023.

[206] Zhiheng Xi, Wenxiang Chen, Xin Guo, Wei He, Yiwen Ding, Boyang Hong, Ming Zhang, Junzhe Wang, Senjie Jin, Enyu Zhou, et al. The rise and potential of large language model based agents: A survey. *arXiv preprint arXiv:2309.07864*, 2023.

[207] Hossein Hassani, Xu Huang, Steve MacFeely, and Mohammad Reza Entezarian. Big data and the united nations sustainable development goals (un sdgs) at a glance. *Big Data and Cognitive Computing*, 5(3):28, 2021.

[208] Maryam Sarrafzadeh. *The implications of knowledge management for the library and information professions*. PhD thesis, RMIT University, 2008.

[209] Steven I Ross, Fernando Martinez, Stephanie Houde, Michael Muller, and Justin D Weisz. The programmer's assistant: Conversational interaction with a large language model for software development. In *Proceedings of the 28th International Conference on Intelligent User Interfaces*, pages 491–514, 2023.

[210] Fengchun Miao, Wayne Holmes, Ronghuai Huang, Hui Zhang, et al. *AI and education: A guidance for policymakers*. UNESCO Publishing, 2021.

[211] Ben Shneiderman. Bridging the gap between ethics and practice: guidelines for reliable, safe, and trustworthy human-centered ai systems. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 10(4):1–31, 2020.

[212] SG Deshmukh, Nitin Upadhye, and Suresh Garg. Lean manufacturing for sustainable development. *Glob. Bus. Manag. Res. Int. J*, 2(1):125, 2010.

[213] Muhammad Sadiq, Thang Le-Dinh, Trung Kien Tran, FengSheng Chien, Thi Thu Hien Phan, and Pham Quang Huy. The role of green finance, eco-innovation, and creativity in the sustainable development goals of asean countries. *Economic research-Ekonomska istraživanja*, 36(2), 2023.