AI Hallucinations: A Misnomer Worth Clarifying

Negar Maleki
University of South Florida
negarmaleki@usf.edu

Balaji Padmanabhan
University of Maryland
bpadmana@umd.edu

Kaushik Dutta
University of South Florida
duttak@usf.edu

***Abstract*—As large language models continue to advance in Artificial Intelligence (AI), text generation systems have been shown to suffer from a problematic phenomenon termed often as "hallucination." However, with AI's increasing presence across various domains including medicine, concerns have arisen regarding the use of the term itself. In this study, we conducted a systematic review to identify papers defining "AI hallucination" across fourteen databases. We present and analyze definitions obtained across all databases, categorize them based on their applications, and extract key points within each category. Our results highlight a lack of consistency in how the term is used, but also help identify several alternative terms in the literature. We discuss implications of these and call for a more unified effort to bring consistency to an important contemporary AI issue that can affect multiple domains significantly[1].**

***Index Terms*—AI, Hallucination, Generative AI**

I. Introduction

One of the early uses of the term "hallucination" in the field of Artificial Intelligence (AI) was in computer vision, in 2000 [1], where it was associated with constructive implications such as super-resolution [1], image inpainting [2], and image synthesis [3]. Interestingly, in this context hallucination was regarded as a valuable asset in computer vision rather than an issue to be circumvented. For instance, an image with low resolution might have been rendered more useful with careful hallucination [1] that generated additional pixels specifically for this purpose.

Despite this (more positive) beginning, recent research has started to employ the term "hallucination" to describe a specific type of error in image captioning [4] and adversarial attack in object detection [5]. In this context, "hallucination" refers to instances where non-existent objects are erroneously detected or incorrectly localized at their anticipated positions. This latter (more negative) interpretation of "hallucination" in computer vision mirrors its analogous usage in language models. For instance, in 2017, researchers highlighted challenges in language models, such as *"the output of the Neural Machine Translation (NMT) system is often quite fluent but entirely unrelated to the input"* [6], or *"language models presume likelihood, but the generated content is ultimately incorrect and unsupported by any information"* [7], which is interpreted as a form of hallucination in AI.

To date, a precise and universally accepted definition of "hallucination" remains absent in the discussions related to this in the increasingly broader field of AI [8]. Diverse definitions, or implied interpretations, persist; sometimes even contradictory, as previously highlighted within the field of computer vision where multiple, disparate interpretations coexist under the same term.

Beyond the AI context, and specifically in the medical domain, the term "hallucination" is a psychological concept denoting a specific form of sensory experience [9]. Ji et al. [10], from the computer science perspective (in ACM Computing Surveys), rationalized the use of the term "hallucination" as *"an unreal perception that feels real"* by drawing from Blom's definition — *"a percept, experienced by a waking individual, in the absence of an appropriate stimulus from the extracorporeal world."* On the other hand, Østergaard et al. [11], from the medical perspective (in Schizophrenia Bulletin, one of the leading journals in the discipline), raised critical concerns regarding even the adoption of the "hallucination" terminology in AI for two primary reasons: 1) The "hallucination" metaphor in AI from this perspective is a misnomer, as AI lacks sensory perceptions, and errors arise from data and prompts rather than the absence of stimuli, and 2) this metaphor is highly stigmatizing, as it associates negative issues in AI with a specific issue in mental illness, particularly schizophrenia, thereby possibly undermining many efforts to reduce stigma in psychiatry and mental health.

Given AI's increasing presence across various domains, including the medical field, concerns have arisen regarding the multifaceted, possibly inappropriate and potentially even harmful use of the term "hallucination" [11], [12]. To address this issue effectively, two potential paths of work offer some promise: 1) The establishment of a consistent and universally applicable terminologies that can be uniformly adopted across all AI-impacted domains will help, particularly if such terminologies lead to the use of more specific and nuanced terms that actually describe the issues they highlight (as we will show later, such vocabulary does exist, but needs more consistent use) and 2) The formulation of a robust and formal definition of "AI hallucination" within the context of AI. These measures are essential to promote clarity and coherence in discussions and research related to "hallucination" in AI, and to mitigate potential confusion and ambiguity in cross-disciplinary applications.

Motivated by these issues, in this paper, we conduct a systematic review of the use of "AI hallucination" across 14

databases with a focus on identifying various definitions that have been used in the literature so far (our review covers more fields than just healthcare and computer science, including ethical and legal settings, and domains as diverse as physics, sports, etc. in order to explore any broader issues). Recently, two papers ([10], [13]) explored the concept of hallucination in Natural Language Generation- (NLG-) specific tasks (e.g., text translation, text summarization, knowledge graph, etc.). Our work builds on these studies to also consider the application of NLG in diverse domains. The pervasive nature of AI extends beyond these specific tasks, affecting numerous domains and applications. Consequently, our broader review done here reveals the extensive utilization of Large Language Models (LLMs) across almost a much broader space of domains to date, and provides a comprehensive understanding of how the term has been leveraged across various fields. Generally we see that research attempting to define "AI hallucination" does so based on their individual understanding and the challenges encountered within their respective fields. The findings from our systematic and broad review underscore the challenge that the term "AI hallucination" lacks a precise, universally accepted definition, resulting in the observation of various characteristics associated with this term across different applications. We present a summary of these different interpretations and provide some guidance going forward.

II. Methodology

Our systematic review covered an extensive database search across various domains, including computer science and health, with a focus on the following databases: PubMed, MEDLINE, Scopus, PubMed Central, Web of Science, BioMed Central, Embase, PLOS, CINAHL, ACM, IEEEXplore, ScienceDirect, Google Scholar, and arXiv (no relevant documents were found in MedlinePlus, Cochrane Library, and UpToDate databases, so those were excluded).

Our search methodology was tailored to adapt to the volume of results as well as the relevance of the papers to our research objectives. We manually reviewed every paper that made it through this process in order to identify possible definitions/usage of the term "hallucination" in AI. Given this goal we had to adapt the search in some cases to identify papers most closely relevant to this objective as noted further below. Also, given differences in how search queries are interpreted across the different databases we had to iteratively modify the search term within each database as well in many cases. For clarity, we present all the specific details of this below (in order to be transparent about how we created the subset of papers from which to examine the definitions). However, the summary of these is provided in Table I, including details of the study period for each database.

In the PubMed Database, we initiated an advanced search employing the keywords "Artificial Intelligence" AND "Hallucination" within the "All field" category, yielding 103 papers within the last 10 years. However, the query "AI+hallucination" yielded only 3 papers. Conversely, within the Scopus database, searches for "AI hallucination" or "AI+hallucination" resulted in a total of 1445 records across all fields over the same 10-year period. To manage this extensive dataset, we refined our search criteria to focus on the Title, Abstract, or Introduction, which reduced the results to 483 relevant records. A detailed review of each abstract led us to download papers that appeared pertinent to AI hallucination. This approach significantly differed from searching within abstracts alone, which produced only 49 records and missed some relevant documents.

In PubMed Central (PMC), the query "AI+hallucination" yielded just 1 paper. Consequently, we conducted an advanced search using the keywords "Artificial Intelligence" AND "Hallucination" within the "Text Word" category, uncovering 371 records from the past decade. PMC does not provide abstracts, necessitating the manual examination of each paper to assess its relevance to AI hallucination.

In both the MEDLINE and Web of Science databases, we employed the term "AI hallucination" within the "All field" category, yielding 157 and 139 records, respectively, spanning the last 10 years. In the case of MEDLINE, each record underwent a thorough review, and records containing definitions for AI hallucination were downloaded. Conversely, Web of Science offered abstracts for the records, enabling us to screen them individually and select those relevant to AI hallucination.

Within the BioMed Central (BMC) database, a search for "AI hallucination" led to the retrieval of 76 papers published within the last 10 years. Subsequently, we accessed each paper individually and downloaded those containing pertinent definitions of AI hallucination. In the Embase database, our search for "AI hallucination" produced 80 records published within the last 10 years. These records underwent meticulous abstract review, and those relevant to AI hallucination were selectively downloaded for further analysis.

In the PLOS database, our initial search with "AI hallucination" resulted in a substantial 1064 records across all fields for the past 10 years. Given this large dataset, we refined our search to focus on the "Body" section, yielding 885 records. We proceeded to review the abstracts of each paper and downloaded those demonstrating relevance to AI hallucination. Within the CINAHL database, we conducted an advanced search utilizing the terms "AI" or "Artificial Intelligence" combined with "Hallucination" within the "All field" category, yielding 34 records published in the past 10 years.

In the ACM and IEEEXplore databases, we performed advanced searches using the terms "AI" AND "hallucination" within the "Full Text" category, resulting in 264 and 257 records, respectively, spanning the past 10 years. Each record underwent individual review, and those containing definitions related to AI hallucination, particularly within the field of LLMs, were downloaded.

In the ScienceDirect database, searches for "AI hallucination" or ("AI" AND "hallucination") yielded an identical number of records, specifically 769 English records, spanning the last decade. Each record underwent meticulous examina-

TABLE I
METHOD SUMMARY

Source/ Database	Search Category	Query Terms	Num. of Papers	Study End Date*
PubMed	All Field	"Artificial Intelligence" AND "Hallucination"	103	09/27/2023
MEDLINE	All Field	"AI hallucination"	157	09/28/2023
Scopus	Title, Abstract, or Introduction	"AI+hallucination" OR "AI hallucination"	483	09/27/2023
PubMed Central	Text Word	"Artificial Intelligence" AND "Hallucination"	371	09/28/2023
Web of Science	All Field	"AI hallucination"	139	09/28/2023
BioMed Central	All Field	"AI hallucination"	76	09/29/2023
Embase	All Field	"AI hallucination"	80	09/29/2023
PLOS	Body	"AI hallucination"	885	09/29/2023
CINAHL	All Field	"Hallucination" AND ("AI" OR "Artificial Intelligence")	34	09/29/2023
ACM	Full Text	"AI" AND "hallucination"	264	09/30/2023
IEEEXplore	Full Text	"AI" AND "hallucination"	257	09/30/2023
ScienceDirect	All Field	"AI hallucination" OR ("AI" AND "hallucination")	769	09/30/2023
Google Scholar	All Field	"AI hallucination" AND "hallucination in AI"	89	10/01/2023
arXiv	All Field	"AI" AND "hallucination"	40	10/01/2023

* The start date is the same for all databases: 01/01/2013 (Date format: mm/dd/yyyy).

tion, and we selectively downloaded papers containing defined concepts of AI hallucination within the LLMs domain.

Searching for "AI hallucination" on Google Scholar yielded 17,000 records from the last 10 years, rendering a comprehensive review unfeasible. To address this challenge, we employed Google Scholar's advanced search feature, identifying records containing the exact phrases "AI hallucination" and "hallucination in AI," which reduced the results to 89 records from the last decade. Subsequently, we conducted a meticulous screening of these records to identify those providing definitions for AI hallucination.

Similarly, within the arXiv database, we conducted an advanced search using the keywords "AI" AND "hallucination" within the "All field" category, resulting in the retrieval of 40 relevant papers. As with our prior search, we meticulously examined each paper and downloaded those containing definitions for AI hallucination.

The eligibility criteria encompassed any type of published scientific research or preprints, such as articles, reviews, communications, editorials, and opinions, that contained the following search terms: "AI hallucination," "AI" AND "hallucination," "Hallucination in AI," or ("AI" OR "Artificial Intelligence") AND "hallucination" in any part of the document, including the title, abstract, and full text. As explained for each database, we employed the most appropriate search terms.

Initially, our search yielded 3753 records, in total, matching these criteria. However, we refined our search to focus exclusively on records that offered a definition of "AI hallucination" within the context of LLMs. It is essential to clarify that we excluded other types of hallucination, such as face hallucination, auditory voice/verbal hallucination, etc., as they were not the primary focus of this review. Our search involved thorough examination of entire documents, and we collected any documents that indicated the presence of a definition for AI hallucination.

Our exclusion criterion was limited to non-English records. The precise database search strategy encompassed all available documents from January 1st, 2013, to October 1st, 2023. In total, we identified 333 records that provide a definition either independently or by inference from a referenced paper. The summary of the methodology is provided in Table I, including details of the study period for each database. While our review of this work is one of the broadest to date, we acknowledge limitations that are implicit in the methodology above - particularly ones where we had to reduce the retrieved number of papers to focus on potentially more relevant ones due to the manual nature of our review (i.e. where we individually reviewed each paper to identify how the term was used and extract the relevant definition in the proper context). Therefore, while the definitions presented here are certainly those that were used it is possible we may have missed a few other definitions that may have newer connotations not identified in our work here. All the 333 definitions are provided in the Appendix.

III. RESULT

We reviewed all retrieved papers and documented the definitions provided in each. One main takeaway was that a formal and consistent definition of hallucination simply does not currently exist. There is also little agreement on the specific characteristics of AI hallucination. Depending on the application, we observe varying characteristics, sometimes even contradictory ones.

For instance, in the context of text translation, Koehn and Knowles [6] described hallucination as "fluent but irrelevant," or Miao et al. [26] characterized it as "fluent but inadequate," while Lee et al. [27] attributed "abnormal and unrelated" characteristics to it, thus illustrating different attributes within the

TABLE II
ALTERNATIVE TERMS USED

Alternative Terms	Definitions	References
Confabulation	AI generated responses that sound plausible but are, in fact, incorrect.	[14]
	Definition was not provided.	[15]
Delusion	AI generated responses that are false.	[16]
Stochastic Parroting	The repetition of training data or its patterns, rather than actual understanding or reasoning.	[17]
	LLM model generates confident, specific, and fluent answers that are factually completely wrong.	[18]
	Definition was not provided.	[19]
Factual Errors	Inaccuracies in information or statements that are not in accordance with reality or the truth, often unintentional but resulting in incorrect or misleading information.	[20]
Fact Fabrication	The occurrence where inaccurate information is invented, not represented in the training dataset, and is presented lucidly.	[21]
Fabrication	The phenomenon where, as a generative AI, ChatGPT generates outputs based on statistical prediction of the text without human-like reasoning, potentially resulting in plausible-sounding but inaccurate responses.	[22]
	The phenomenon in ChatGPT output where the text is cogent but not necessarily true.	[23]
	Definition was not provided.	[24]
Falsification and Fabrication	Definition was not provided.	[12]
Mistakes, Blunders, Falsehoods	Answers that are fabricated when data are insufficient for an accurate response.	[25]
Hasty Generalizations, False Analogy, False Dilemma	AI models making inferences that do not follow from the premises; also "hasty generalizations," i.e., the fallacy of making (too) strong claims based on (too) limited data.	[11]

same context. In the text summarization context, hallucination refers to generated content that is inconsistent with the source document [28], [29], with some studies categorizing it into subtypes: "Intrinsic hallucination" and "Extrinsic hallucination" [30], [31], raising concerns, particularly regarding the latter.

Before the launch of ChatGPT on November 30, 2022, we hardly observed definitions for AI hallucination in fields other than computer science. However, with the advent of ChatGPT, researchers have recognized the urgent need for Large Language Models (LLMs) in various fields, including medicine. Therefore, over time, we have observed that the definition has changed and seems to have become a problem more relevant to ChatGPT, albeit with different characteristics under the same term across various applications.

In recent times, for reasons discussed earlier in this paper as well as broader concern about giving AI "human" characteristics inadvertently by using this term, researchers have made efforts to replace the term 'hallucination,' deeming it unsuitable and advocating for its renaming or for alternatives. We have compiled many of the suggested terms found in the literature in Table II, along with their definitions in the respective papers. This is a start in the right direction perhaps - in the search for specific definitions and specific characteristics that we want to model - but does illustrates the lack of consistency in the literature that we pointed out in this paper.

Based on the alternate terms we found, some "old" problems appear to re-surface: the terms confabulation and delusional for instance have connections to mental health conditions as well. However, fabrication, stochastic parroting and hasty generalization together suggest three viable alternatives. Fact fabrication captures many of the cases previously attributed to 'hallucination' without the negative connotations, while stochastic parroting appears to be an appropriate descriptive term for the reasons behind fact fabrication in Generative AI. While we need clarity in terms of distinguishing between stated facts (in the training data) and inferences, the reference to hasty generalization does start to capture such a distinction.

Finally, since our focus here was on reviewing AI hallucinations across various applications, we grouped all the final papers examined by category, extracted definitions related to AI hallucination, and used ChatGPT 3.5 [32] to extract key points. The applications included chatbots, dialogue settings, generative AI, academia, health, legal and ethical settings, science, technology, text translation, question and answering, text summarization, and others. As shown in Table III, the extracted summaries share similar characteristics, but highlight different extents of inaccuracy, ranging from "deviating from established knowledge", "factual incorrectness", "fictional" to "nonsensical" – offering further considerations for a robust taxonomy that will be needed to bring out such nuances.

IV. DISCUSSION

"Hallucinate" secured its position as the word of 2023 ([38], [39]) and Dictionary.com noted a 46% surge in searches for the term over the past year. The popular press has also keyed in on this (Table IV highlights topics some recent articles discuss and the meaning of "AI hallucination" they convey). These articles primarily feature interviews with CEOs of big tech companies, who discuss future efforts to prevent "hallucination" in chatbots' outputs. Indeed, preventing such occurrences continues to be a key research goal, but few solutions have emerged so far. In the meanwhile, the Generative AI continues to expand its applications into multiple domains, making the need for good solutions vital. As a precursor to even developing solutions, this paper calls for

TABLE III
KEY POINTS OF "HALLUCINATION" DEFINITIONS WITHIN EACH APPLICATION. THE CHARACTERISTICS OF DEFINITIONS ARE PRESENTED IN **BOLD**, ALTHOUGH THEY MAY BE SIMILAR ACROSS DIFFERENT APPLICATIONS.

Application	Number of Papers	LLM Generated Key Points of Definitions
Chatbot	34	The definitions collectively highlight the central theme of AI-generated content **deviating from factual correctness**, at times even leading to entirely **fictional or erroneous information**. In essence, AI hallucination underscores the ongoing challenge of maintaining accuracy and reliability in AI-generated content within the context of chatbot applications.
Dialogue Setting	8	The definitions collectively underscore the challenge of ensuring accuracy and reliability in dialogue systems, given the potential pitfalls associated with generating content that is **unsupported, nonsensical, or factually incorrect**. These issues are particularly pertinent when deploying large pre-trained language models in dialogue applications, as they struggle with maintaining fidelity to the source material while generating coherent and accurate responses.
Generative AI	50	The definitions collectively emphasize the complexity of ensuring factual accuracy and reliability in AI-generated content within generative AI applications, highlighting the potential pitfalls of **deviating from adherence to factual correctness**.
Academia	88	A common thread among these definitions is the generation of text or content by AI models that **lacks fidelity to factual accuracy, reality, or the intended context**.
Health	82	The key idea common to all the definitions is that "AI hallucination" occurs when AI systems generate information that **deviates from factual accuracy, context, or established knowledge**. In essence, AI hallucination manifests as the production of text that, though **potentially plausible, deviates from established facts or knowledge** in health applications.
Legal and Ethical Setting	16	The definitions collectively emphasize the multifaceted challenges posed by AI hallucination in the legal and ethical context. They highlight **issues of accuracy, confidence, relevance, context, and potential misinformation**, underscoring the critical importance of addressing these challenges to ensure the responsible and ethical use of AI systems.
Science	10	Across the definitions, the central theme is that AI hallucination involves the generation of text or information that **deviates from factual accuracy, coherence, or faithfulness to the input or source content**, with potential consequences for scientific accuracy and integrity.
Technology	8	The definitions reflect the multifaceted nature of AI hallucination in technology applications, encompassing **accuracy, unpredictability, credibility, and the balance between reasonableness and correctness**.
Text Translation	4	The definitions collectively emphasize the central theme of "AI hallucination" in text translation, which revolves around challenges related to maintaining **fidelity, coherence, and relevance** in the generated translations to ensure accurate and meaningful output.
Question and Answering	7	"AI hallucination" in question and answer applications raises concerns related to the **accuracy, truthfulness, and potential spread of misinformation** in AI-generated answers, emphasizing the need for improving the reliability of these systems.
Text Summarization	19	The definitions highlight the multifaceted challenges posed by "AI hallucination" in text summarization, encompassing issues related to **fidelity, coherence, factual accuracy, and the preservation of the original meaning** in generated summaries.
Others *	7	These diverse applications collectively emphasize the challenge of maintaining **accuracy, coherence, and trustworthiness** in AI-generated content, highlighting the need for tailored approaches to address domain-specific concerns.

* Including: Investment portfolio, Journalism, Reinforcement Learning, Retail, Sport, and Survey Setting.

TABLE IV
SOME POPULAR PRESS ARTICLES ON AI HALLUCINATION

	What the Press Article Discussed...	The Real Meaning the Press Article Conveys about "AI Hallucination"	Source
1	CNBC provided some examples where ChatGPT generated outputs that sounded correct but weren't actually true, such as a legal brief written by ChatGPT to a Manhattan federal judge	When an AI model "hallucinates," it generates fabricated information in response to a user's prompt, but presents it as if it's factual and correct	CNBC [33]
2	The New York Times asked ChatGPT, Google's Brad, and Microsoft's Bing: When did The New York Times first report on "artificial intelligence"?	Chatbots provide inaccurate answers to questions; although false, the responses appear plausible as they blur and conflate people, events, and ideas	The New York Times [34]
3	The New York Times traced the evolution of the term "hallucination" throughout the newspaper's history	-	The New York Times [35]
4	CNN addressed the major issue of "AI hallucination" and narrated on the responses of OpenAI's and Google's CEOs to the question: Can hallucination be prevented?	AI-powered tools like ChatGPT impress with their ability to provide human-like responses, but a growing concern is their tendency to just make things up	CNN [36]
5	Forbes narrated the history of artificial neural networks, which started around eight decades ago, when researchers sought to replicate the functioning of the brain	"AI hallucination" refers to unrealistic ideas about achieving "artificial general intelligence" (AGI), while understanding of how our brains work is limited	Forbes [37]

more systematic, consistent and semantically nuanced terms that can replace "hallucinations" for the reasons noted here. As one step toward such a call, we presented a short summary from one of the broadest manual literature reviews on this topic to date. Our findings illustrate the current lack of consistency and consensus on this issue, but also bring to light some recent options that are good alternatives. More work is needed to develop a systematic taxonomy that can be widely adopted as we discuss these issues in the context of AI applications.

APPENDIX
REVIEW OF THE "AI HALLUCINATION" DEFINITIONS

TABLE V: AI hallucination definitions

Num.	Author(s)	Year	Application	Definition	Citation
1	Koehn and Knowles [6]	2017	Text Translation	"AI hallucination" occurs when the output of the Neural Machine Translation (NMT) system is often quite fluent but entirely unrelated to the input.	1249
2	Wiseman et al. [7]	2017	Natural Language Generation	"AI hallucination" occurs when a language model presumes likelihood, but the generated content is ultimately incorrect and unsupported by any information.	556
3	Lee et al. [40]	2018	Text Translation	"AI hallucination" occurs when Neural Machine Translation (NMT) systems produce highly pathological translations that are completely untethered from the source material.	83
4	Nie et al. [41]	2019	Large Language Model	"AI hallucination" refers to the problem where the generated texts often contain information that is irrelevant to or contradicted by the input.	63
5	Tian et al. [42]	2019	Large Language Model	"AI hallucination" occurs when a language model generates text that is fluent but unfaithful to the source.	66
6	Dušek et al. [43]	2019	Natural Language Generation	"AI hallucination" occurs when a language model adds information that is not grounded in the input.	89
7	Ferreira et al. [44]	2019	Natural Language Generation	"AI hallucination" refers to a language model describing non-linguistic representations that are not present in the input.	141
8	Martindale et al. [45]	2019	Text Translation	"AI hallucination" occurs when the machine translation output contains more information than the reference text.	29
9	Dušek and Kasner [46]	2020	Natural Language Generation	"AI hallucination" occurs when expressions in the output do not correspond to input facts.	46
10	Kang and Hashimoto [47]	2020	Natural Language Generation	"AI hallucination" occurs when neural language models often produce fluent text that is unfaithful to the source.	72
11	Parikh et al. [48]	2020	Natural Language Generation	"AI hallucination" occurs when a language model generates text that is fluent but not faithful to the source.	242
12	Maynez et al. [30]	2020	Text Summarization	"AI hallucination" refers to content that is unfaithful to the input document. "Intrinsic hallucinations" are consequences of synthesizing content using the information present in the input document. "Extrinsic hallucinations" are model generations that ignore the source material altogether. Hallucinate content that is unfaithful to the input document.	572
13	Durmus et al. [49]	2020	Question Answering	"AI hallucination" refers to the occurrence when AI generates content inconsistent with the source document, i.e., unfaithful.	269
14	Zhao et al. [28]	2020	Text Summarization	"AI hallucination" occurs when language models generate material that is not supported by the original text.	73
15	Dong et al. [29]	2020	Text Summarization	"AI hallucination" occurs when a language model generates content that is factually inconsistent with the source documents.	89
16	Filippova [8]	2020	Large Language Model	"AI hallucination" refers to the generated content which is either unfaithful to the input or nonsensical.	46
17	Zhou et al. [50]	2020	Hallucination Detection	"AI hallucination" occurs when the model generates additional content not supported by the input.	87
18	Elsahar et al. [51]	2020	Hallucination Mitigation	"AI hallucination" occurs when the fluency of natural language generation models can be highly misleading, as it often distracts from the wrong facts stated in the generated text.	33

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
19	Tang et al. [52]	2021	Text Summarization	"AI hallucination" refers to content that is not supported by the source documents.	32
20	Cao et al. [53]	2021	Text Summarization	"AI hallucination" refers to content that is not directly inferable from the source text.	37
21	Miao et al. [26]	2021	Text Translation	"AI hallucination" refers to generating fluent but inadequate translations to the source sentences.	21
22	Chen et al. [54]	2021	Text Summarization	"Intrinsic" and "Extrinsic hallucinations", involving the fabrication of untruthful text spans containing information that may either be present or absent from the source.	62
23	Wang et al. [55]	2021	Text Summarization	"AI hallucination" refers to where the model generates fictional content.	14
24	Xiao and Wang [56]	2021	Large Language Model	"AI hallucination" occurs where models generate description tokens that are not supported by the source inputs.	58
25	Dziri et al. [57]	2021	Knowledge Graph	"AI hallucination" occurs when dialogue models, despite maintaining plausible general linguistic capabilities, are still unable to fully discern facts and may instead hallucinate factually invalid information.	51
26	Liu et al. [58]	2021	Hallucination Detection	"AI hallucination" occurs when language models exhibit a propensity to hallucinate non-existent or incorrect content that is unacceptable in most user-oriented applications.	29
27	Huang et al. [59]	2021	Text Summarization	"Intrinsic hallucination" is a fact that is contradicted to the source document. "Extrinsic hallucination" is the fact that is neutral to the source document (i.e., the content that is neither supported nor contradicted by the source document).	51
28	Lin [60]	2021	Question Answering	"AI hallucination" refers to the tendency of LLMs to generate false statements.	250
29	Shuster et al. [61]	2021	Large Language Model	"AI hallucination" refers to the occurrence where language models generate plausible-looking statements that are factually incorrect.	198
30	Sekulić et al. [62]	2021	Natural Language Generation	"AI hallucination" occurs where generated responses do not correspond to the real-world.	26
31	Ghosh et al. [63]	2021	Reinforcement Learning	"AI hallucination" occurs when the generated text asserts information not present in the source.	5
32	Perez-Beltrachini and Lapata [64]	2021	Text Summarization	"AI hallucination" refers to neural summarization models' propensity to generate text that does not preserve the meaning of the input.	18
33	Lyu et al. [31]	2022	Text Summarization	"AI hallucination" signifies the presence of distorting or fabricating facts within generated summaries, resulting in inconsistencies between a summary and the corresponding original document. "Extrinsic hallucination" entails adding information not directly inferable from the input information. "Intrinsic hallucination" involves manipulating the information present in the input document.	0
34	Ali et al. [65]	2022	Health	"AI hallucination" refers to scenarios in which a Language Model (LLM) asserts inaccurate facts or contextual data that it falsely believes to be correct in its response.	29

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
35	Ando et al. [66]	2022	Health	"AI hallucination" occurs when abstractive summarization may sometimes unintentionally generate unfaithful descriptions. "Intrinsic hallucination" is a phenomenon in which the concept or term itself is in the source documents; its synthesis misrepresents the information in the source, and the meaning becomes inconsistent. "Extrinsic hallucination" is content that is neither supported nor contradicted by the source and is caused by source documents with poor information.	2
36	Ando et al. [67]	2022	Health	"AI hallucination" refers to the phenomenon where the abstractive approach frequently produces fake content that does not align with the reference summary.	2
37	Rebuffel et al. [68]	2022	Natural Language Generation	"AI hallucination" refers to model outputs that are often subject to over-generation, where misaligned fragments from training instances, known as divergences, can induce similarly misaligned outputs during inference.	43
38	Wan and Bansal [69]	2022	Text Summarization	"AI hallucination" refers to a scenario where a summary contains facts or entities not present in the original document.	30
39	Corbelle et al. [70]	2022	Natural Language Generation	"AI hallucination" refers to instances where neural models generate texts that are incoherent or unrelated to the input of a D2T system.	2
40	Lee et al. [71]	2022	Large Language Model	"AI hallucination" occurs when a model is making factual errors, generating a named-entity that does not appear in the ground-truth knowledge source.	30
41	Cabezudo et al. [72]	2022	Natural Language Generation	"AI hallucination" refers to text generated by pre-trained models that is irrelevant to or contradicted with the input.	0
42	Van et al. [73]	2022	Text Summarization	"AI hallucination" refers to the occurrence of adding information to the output that was not present in the original text.	7
43	Dziri et al. [74]	2022	Dialogue	"AI hallucination" occurs when large pre-trained language models generate factually incorrect statements.	51
44	Dziri et al. [75]	2022	Dialogue	"AI hallucination" refers to the phenomenon where dialogue systems often produce unsupported utterances.	24
45	Koto et al. [76]	2022	Text Summarization	"AI hallucination" occurs where information is generated that does not exist in the source document, also called "factual inconsistencies."	21
46	Goodman et al. [77]	2022	Email Writing	"AI hallucination" refers to factually incorrect or nonexistent content generated by the LLM.	10
47	Gehrmann et al. [78]	2022	Natural Language Generation	"AI hallucination" refers to a situation where a model is not faithful as it adds information not present in the source document. "Intrinsic hallucinations" misrepresent facts in the input. "Extrinsic hallucinations" ignore the input altogether.	62
48	Erdem et al. [79]	2022	Natural Language Generation	"AI hallucination" refers to the generation of descriptions or facts that are not fully supported by the input.	20
49	Tun et al. [80]	2022	Dialogue	"AI hallucination" refers to large-scale pre-trained language models generating text that is nonsensical and struggling to remain true to the source content.	0
50	Gurrapu et al. [81]	2022	Claim Verification	"AI hallucination" refers to the phenomenon where natural language generation models introduce unintended and irrelevant text during the generation process.	1
51	Yang et al. [82]	2022	Text Summarization	"AI hallucination" occurs when models generate summaries factually inconsistent with their original document.	2

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
52	Ji et al. [10]	2023	Natural Language Generation	"AI hallucination" refers to the generated content that is nonsensical or unfaithful to the provided source content. "Intrinsic hallucinations" encompasses the generated output that contradicts the source content. "Extrinsic hallucination" involves the generated output that cannot be verified from the source content (i.e., output can neither be supported nor contradicted by the source).	482
53	Narayanan et al. [83]	2023	Natural Language Generation	"AI hallucination" refers to situations in which AI models provide responses with confidence that appear faithful but are nonsensical when evaluated against common knowledge.	2
54	Leiser et al. [84]	2023	Large Language Model	"Artificial hallucinations" are the distortion of perception in Large Language Models (LLMs) as they present incorrect information, disguising it as factual responses.	0
55	Wang et al. [85]	2023	ChatGPT	"AI hallucination" denotes the tendency to generate inaccurate outputs unfaithful to the training data.	1
56	Ma et al. [86]	2023	Software Engineering	"AI hallucination" refers to the phenomenon that can result in the fabrication of elements that do not actually exist.	8
57	Li [17]	2023	Large Language Model	"AI hallucination" occurs when LLMs generate text based on their internal logic or patterns, rather than the true context, leading to confidently but unjustified and unverified deceptive responses.	11
58	Vaghefi et al. [87]	2023	NLP and Climate Change	"AI hallucination" refers to mistakes in the generated text that are semantically incorrect or unsupported by the input text.	5
59	Daull et al. [88]	2023	Question Answering	"AI hallucination" refers to confident generated responses that contain false information not supported by the model's training data. "Extrinsic hallucination" involves a model introducing information not present in the source data. "Intrinsic hallucination" occurs when the model distorts information from the source data into a factually incorrect representation.	2
60	Zhang et al. [89]	2023	Large Language Model	"AI hallucination" commonly refers to the phenomenon where LLMs occasionally generate outputs that, although appearing plausible, deviate from the user input, previously generated context, or factual knowledge.	3
61	Romanko et al. [90]	2023	Investment Portfolio	"AI hallucination" refers to instances where the AI, although generating text based on its training, does so without a solid conceptual framework behind it.	3
62	Henderson et al. [91]	2023	AI Speech	"AI hallucination" refers to the act of generating text that includes factual claims that are untrue, and in some cases, these claims may not have been present in its training data.	2
63	Ni et al. [92]	2023	NLP and Climate Change	"AI hallucination" refers to the generation of answers in which some or all of the covered information is not adequately supported by the report, particularly when there is extrapolation or partial support involved, and it may also relate to instances where the model fails to honestly report its references while generating content that lacks fidelity to the source.	0
64	Mahmood et al. [93]	2023	Radiology Report	"AI hallucination" involves the generation of false findings in the produced reports by LLMs.	1
65	Goyal et al. [94]	2023	Question Answering	"Hallucinations" signify the production of convincing yet incorrect text outputs by LLMs, with the potential to distort scientific facts and disseminate misinformation.	0

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
66	Zhang [95]	2023	Robotic	"AI hallucination" refers to unpredictable outputs generated by LLMs.	0
67	Li et al. [96]	2023	Large Language Model	"AI hallucination" is the phenomenon where responses are confidently generated but are incorrect.	1
68	Li et al. [97]	2023	Large Language Model	"AI hallucination" occurs when AI systems generate outputs that are not aligned with the input context and encounter challenges in efficiently capturing complex dependencies.	1
69	Curran et al. [98]	2023	Legal Setting	"AI hallucination" refers to tendency for generating false information.	1
70	Feldman et al. [99]	2023	Large Language Model	"AI hallucination" in the context of LLMs refers to the phenomenon where these language models generate non-factual statements, which could have a detrimental effect on the trustworthiness of their outputs.	2
71	Mukherjee et al. [100]	2023	Generative AI	"Hallucinations," characterized by AI responses containing random inaccuracies or falsehoods, emerge when models prioritize novelty over usefulness.	0
72	Salvagno et al. [101]	2023	Large Language Model	"AI hallucination" refers to the creation of seemingly accurate bibliographic references with recognized authors and coherent titles, even though these references are entirely fictitious and have no real existence.	4
73	Beutel et al. [102]	2023	Large Language Model	"Hallucinations" are defined as the production of content that does not accurately represent the provided source and may appear nonsensical due to errors in the encoding and decoding processes between text and representations.	20
74	Azamfirei et al. [103]	2023	Text Summarization	"AI hallucination" refers to receiving a response from the model when it lacks an appropriate answer, which appears to be the most likely summary of the study, despite its potential inaccuracy.	32
75	Meyer et al. [104]	2023	Academia	"AI hallucination" can be defined as the capability of LLM-based chatbots to convey false information as though it were true.	8
76	Hernigou and Scarlat [105]	2023	Health	"AI hallucination" is a confident response that does not seem in concordance with its training data.	2
77	Patil [106]	2023	Legal Setting	"AI hallucination" is defined as a confident response by an AI system that lacks justification in its training data. These responses can appear factual but are not true, often simply being answers "made up" by the AI.	0
78	Alexander et al. [107]	2023	Academia	"AI hallucination" refers to its tendency to make up facts and references that do not exist.	1
79	Lyell [108]	2023	Academia	"AI hallucination" describes the propensity of the system to convincingly fabricate information.	0
80	Grassini [109]	2023	Academia	"AI hallucination" involves the generation of incorrect or even fabricated information.	14
81	Brodeur et al. [110]	2023	Legal Setting	"AI hallucination" refers to the situation where an AI system, due to its inability to correctly interpret data, generates inaccurate or unusual outputs.	0
82	Lee and Choi [111]	2023	Health	"AI hallucination" occurs when AI generates information about things that are untrue, presenting them as if they are true, which can diminish the reliability of the generated information.	0
83	Chatfield [112]	2023	Chatbot	"AI hallucination" represents a response that may or may not be plausible but is not rooted in reality.	0
84	McGowan et al. [113]	2023	References	"AI hallucination" refers to the occurrence of mistakes in the generated text that, while semantically or syntactically plausible, are ultimately incorrect or nonsensical upon closer examination.	4

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
85	Wu and Dang [114]	2023	Academia	"AI hallucination" refers to a scenario in which a machine generates seemingly realistic outputs without any real-world input.	6
86	Wang et al. [115]	2023	Academia	"AI hallucination" refers to the generation of content that is nonsensical or untrue in relation to certain sources.	2
87	Hou and Ji [116]	2023	Health	"AI hallucination" refers to the situation where GPT models provide confident answers that contradict the truth.	3
88	Au et al. [117]	2023	Health	"AI hallucination" refers to the phenomenon in which LLMs produce output that is nonsensical or unfaithful to the provided input or "prompt," and this problem is further amplified when the prompt contains insufficient or masked information, despite the high confidence displayed in the generated output, as observed in ChatGPT.	15
89	Hua et al. [118]	2023	Health/ Academia	"AI hallucination" is AI-generated outputs that deviate from its training data. These outputs may appear syntactically or semantically plausible, but in reality, they are incorrect or nonsensical.	1
90	Brameier et al. [119]	2023	Health/ Academia	"AI hallucination" is the production of confident responses by an NLP tool that are nonsensical or that seem realistic but are not based on any real-world data.	1
91	Lee et al. [120]	2023	Health	"AI hallucination" refers to the occurrence where GPT-4 produces false responses, which are often presented in a convincing manner, potentially leading the inquirer to believe their accuracy.	260
92	Long et al. [121]	2023	Health/ Academia	"AI hallucination" refers to the phenomenon where Language Models (LMs), including ChatGPT, produce outputs characterized by blatant factual errors, significant omissions, and erroneous information generation.	0
93	Zhang [122]	2023	Large Language Model	"AI hallucination" refers to the generation of responses by ChatGPT that frequently contain incorrect information, and it can even extend to the generation of fake scientific abstracts and research papers.	1
94	Puchert et al. [123]	2023	Health	"AI hallucination" refers to a common phenomenon where the model includes incorrect or false information in its responses, despite providing eloquent answers.	1
95	Wang et al. [124]	2023	Health	"AI hallucination" refers to misinformation, unreasonable or illogical in common-sense knowledge	0
96	Garg et al. [125]	2023	Health	"AI hallucination" refers to the production of content that sounds authoritative but can be inaccurate, incomplete, or biased in nature.	2
97	Han et al. [126]	2023	Health	"AI hallucination" occurs when AI confidently generates an impressive-sounding response that may not be justified by its training data or may even be factually incorrect.	0
98	Dolan and Freer [127]	2023	Academia	"AI hallucination" refers to the phenomenon where AI systems may create or invent text and citations that sound realistic but are not based on actual information, and in certain instances, they may even fabricate nonexistent works.	0
99	Scott-Branch et al. [128]	2023	Academia	"AI hallucination" refers to mistakes in the generated text that are semantically or syntactically plausible but are in fact incorrect or nonsensical.	0
100	Larsen-Ledet [129]	2023	Academia	"AI hallucination" refers to the generation of output, often text, containing falsities, which undermine the intended factual nature of the service, intended to be reliable and trustworthy.	0

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
101	Kim et al. [130]	2023	Health/ Academia	"AI hallucination" refers to the phenomenon in which there is no true internal understanding of the language, and LLMs generate text that sounds plausible based on the training data they have been exposed to.	9
102	Tay et al. [131]	2023	Health	"AI hallucination" refers to the situation where GPT confidently generates responses that lack any justification from training data or correspondence to real-world input.	1
103	Cai et al. [132]	2023	Health	"AI hallucination" refers to the issue of accuracy and reliability in LLMs, particularly concerning the high frequency of false information.	9
104	Jahic et al. [133]	2023	Academia	"AI hallucination" refers to the phenomenon where an AI system generates unexpected or meaningless outputs that appear coherent and plausible to human observers.	2
105	Randell and Coghlan [134]	2023	Academia	"AI hallucination" refers to the plausible yet erroneous text generated by artificial intelligence systems, resembling compression artifacts, which can only be discerned through comparison with original sources like the web or our existing knowledge.	0
106	Brender [135]	2023	Academia	"AI hallucination" refers to the unfaithful or nonsensical text occasionally generated by large language models.	1
107	Pedersen [136]	2023	Academia	"AI hallucination" refers to mistakes in the generated text that are semantically or syntactically plausible but are in fact incorrect or nonsensical.	0
108	Munoz et al. [137]	2023	Academia	"AI hallucination" which results when the system provides a response that is not factual.	1
109	Hatem et al. [138]	2023	Health	"AI hallucination" is inaccurate and stigmatizing to both AI systems and individuals who experience hallucinations. Because of this, they suggest the alternative term "AI misinformation" as they feel this is an appropriate term to describe the phenomenon at hand without attributing lifelike characteristics to AI.	0
110	Athaluri et al. [139]	2023	Health/ Academia	"AI hallucination" is a phenomenon where AI generates a convincing but completely made-up answer.	16
111	Gorichanaz [140]	2023	Academia	"AI hallucination" refers to generating misinformation, including making false statements, citing sources that do not exist and creating code that doesn't work.	1
112	Cox et al. [141]	2023	Health	"AI hallucination" occurs when Language Models (LLMs) misunderstand medical vocabulary or provide advice that does not align with established medical guidelines.	0
113	Yadava [142]	2023	Health/ Academia	"AI hallucination" is interestingly labeled as a phenomenon where responses from ChatGPT are sometimes ambiguous, nonsensical, and undesirable.	11
114	Dai et al. [143]	2023	Academia	"AI hallucination" is a phenomenon in which it sometimes produces confident but irrelevant or inaccurate responses.	11
115	Lingard [144]	2023	Academia	"AI hallucination" refers to the phenomenon where AI generates content that confidently presents as legitimate-sounding material but is not real.	1
116	Woodland [145]	2023	Health/ Academia	"AI hallucination" refers to a situation where the generated text is not derived from a factual source but instead arises from statistical predictions of words that are likely to follow the given input.	0
117	Kashangura [146]	2023	Health	"AI hallucination" is the generation or production of a factually invalid statement, characterized by confident output from AI that is not correct or not real.	0
118	Walker et al. [147]	2023	Health	"AI hallucinations" known as wrong or out of context answers generated by ChatGPT AI.	6

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
119	Tenge Hansen and Røsand Valø [148]	2023	Academia	"AI hallucination" is defined as a phenomenon where AI generates a convincing but completely made-up answer, often incorporating fake references for added persuasiveness. "Intrinsic hallucination" refers to the LLM generation that contradicts the source/input. "Extrinsic hallucination" refers to the LLM generations that cannot be verified from the source/input content, i.e., output that can neither be supported nor contradicted by the source.	0
120	Tupper et al. [149]	2023	Academia	"AI hallucination" refers to a phenomenon wherein incorrect or nonsensical responses are generated in response to prompts.	0
121	Borkowski et al. [150]	2023	Health	"AI Hallucination" refers to the phenomenon often observed in language models like ChatGPT, characterized by the generation of inaccurate information, which result from the model's inability to differentiate between real and fake information sources.	0
122	Subramanya and Furlong [151]	2023	Legal Setting	"AI hallucination" occurs when AI generates content that appears accurate but includes false references.	0
123	Tenzer [152]	2023	Legal Setting	"AI hallucination" occurs when a chatbot produces a confident but inaccurate response to a question, contributing to the erratic and unreliable behavior of A.I. large language models (L.L.M.s), which may include providing false information and acting strangely toward users.	0
124	Alarie and McCreight [153]	2023	Legal Setting	"AI hallucination" occurs when AI generates untrue information that is not backed up by real-world data.	0
125	Han et al. [154]	2023	Academia	"AI hallucination" refers to the situation where the AI, while not presenting fictitious information like a hallucination, draws from sources other than the requested one and fails to indicate this, potentially leading to incorrect conclusions when users rely on the AI-generated text for understanding a specific paper.	0
126	Picht [155]	2023	Legal Setting	"AI Hallucination" refers to the phenomenon in which ChatGPT generates responses that sound confident and compelling, despite its inability to genuinely answer the posed question.	0
127	Treleaven et al. [156]	2023	Legal Setting	"AI hallucination" is defined as a confident response that is biased, too specialized, or even entirely incorrect, with the model fabricating a seemingly plausible but factually inaccurate answer.	0
128	Ariyaratne [157]	2023	Legal Setting	"AI Hallucination" refers to the term used to describe the phenomena of generative algorithms making up facts.	0
129	Wu et al. [158]	2023	Large Language Model	"AI hallucination" occurs when the replies generated by ChatGPT frequently contain factual errors.	40
130	Wang et al. [159]	2023	Large Language Model	"AI hallucination" occurs when ChatGPT produces responses that, despite sounding plausible, are ultimately incorrect or nonsensical.	7
131	Amaro et al. [160]	2023	Large Language Model	"AI hallucination" occurs when ChatGPT generates outputs that invent facts and concepts, as it lacks objective training to assess the factual correctness of its responses.	0
132	Skrodelis et al. [161]	2023	Natural Language Generation	"AI hallucination" occurs when NLGs frequently produce text that is either nonsensical or not true to the original input.	0
133	Gupta et al. [162]	2023	Cybersecurity	"AI hallucination" refers to a phenomenon in which the AI model generates inaccurate or outright false information.	7

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
134	Bahrini et al. [163]	2023	Chatbot	"AI hallucination" refers to ChatGPT's limitations and challenges, which encompass bias and the occasional generation of nonsensical output	28
135	Hamid [164]	2023	Chatbot	"AI hallucination" refers to the 'hallucinations,' or incoherent responses, observed in Language Models (LLMs), which are essentially errors in the model's output as it is designed to prioritize coherence rather than truth.	1
136	De Silva et al. [165]	2023	Chatbot	"AI hallucination" occurs when ChatGPT exhibits anomalous behavior, producing factual inaccuracies, logical fallacies, bias, and plagiarism in its responses, a phenomenon popularly referred to as 'AI hallucinations' or 'stochastic parroting.'	3
137	Atallah [166]	2023	Health	"AI hallucination" occurs when modern LLMs produce fluent and grammatically correct textual outputs that are categorically false, and in some instances, fictitious.	1
138	Byrne [167]	2023	Large Language Model	"AI hallucination" refers to outputs that reflect misinterpretations and falsehoods.	0
139	Thirunavukarasu et al. [21]	2023	Health	"AI hallucination" refers to the occurrence where inaccurate information is invented, not represented in the training dataset, and is presented lucidly, with an alternative term like 'fact fabrication' being preferred.	30
140	Ting et al. [22]	2023	Health	AI Hallucination refers to the phenomenon where, as a generative AI, ChatGPT generates outputs based on statistical prediction of the text without human-like reasoning, potentially resulting in plausible-sounding but inaccurate responses, also known as "fabrication."	5
141	Beam et al. [168]	2023	Health	"AI hallucination" occurs when a model generates information that isn't present in the input data, leading to outputs that seem plausible but are factually incorrect or nonsensical.	0
142	Madden et al. [16]	2023	Health	"AI hallucination" occurs when ChatGPT generates false information, also called "delusions."	2
143	Komorowski [169]	2023	Health	"AI hallucination" refers to the phenomenon where ChatGPT has the ability to confidently produce answers that appear believable but may be incorrect or nonsensical.	4
144	Sallam [170]	2023	Health/ Academia	"AI hallucination" refers to concerns arising from possible bias in ChatGPT's training datasets, limiting its capabilities and potentially causing factual inaccuracies that, surprisingly, seem scientifically plausible.	334
145	Rothschild [171]	2023	Health/ Academia	"AI hallucination" refers to false responses provided by ChatGPT3.	0
146	Im [172]	2023	Health	"AI hallucination" occurs when ChatGPT generates responses that are unfaithful to the provided training data.	0
147	Hulman et al. [173]	2023	Health	"AI hallucination" refers to the occurrence of making up completely false information supported by fictitious citations.	5
148	Cheung et al. [174]	2023	Health	"AI hallucination" occurs as a phenomenon where the AI generates responses that are nonsensical or unfaithful to the provided source input.	0
149	Moskatel and Zhang [175]	2023	Academia	"AI hallucination" is an inaccurate response by an AI that is not justified by its training data.	0
150	Javid et al. [176]	2023	Health/ Academia	"AI hallucination" refers to the phenomenon where the model produces text that is either factually incorrect or nonsensical.	0

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
151	Buholayka et al. [177]	2023	Health/ Academia	"AI hallucination" occurs when AI-generated content is nonsensical or unfaithful to the provided source content.	4
152	Alkaissi and McFarlane [178]	2023	Health/ Academia	"AI hallucination" occurs when ChatGPT provides confident responses that appear faithful but are nonsensical when evaluated against common knowledge.	230
153	Eysenbach [179]	2023	Health/ Academia	"AI hallucination" is a confident response by an artificial intelligence system that does not seem to be justified by its training data.	128
154	Østergaard and Nielbo [11]	2023	Health	"Non sequitur", Latin for "it does not follow," is a term commonly used in philosophy and rhetoric to describe inferences not following from the premises. They acknowledged that "non sequitur" does not cover all false responses generated by AI models. Indeed, AI models can also make "hasty generalizations," i.e., the fallacy of making (too) strong claims based on (too) limited data. They realized that many other terms from philosophy and rhetoric describing logical fallacies of different types (e.g., "false analogy," "appeal to authority" and "false dilemma") could also be used to label false responses from AI models.	0
155	Thorne [180]	2023	Academia	"AI hallucination" refers to occurrences where ChatGPT produces statements that are difficult to verify, appear plausible, but do not withstand scrutiny, often being arbitrary in nature.	0
156	Huang et al. [181]	2023	Health/ Academia	"AI hallucination" refers to responses generated by LLMs in a convincing appearance but are actually incorrect statements.	5
157	Wecel et al. [182]	2023	Large Language Model	"AI hallucination" refers to the occurrence where AI produces texts that are not consistent with reality and contain confused facts.	0
158	Ge and Lai [18]	2023	Health/ Academia	"AI hallucination" occurs where the LLM model generates confident, specific, and fluent answers that are factually completely wrong. Also called "stochastic parroting."	21
159	Bhatti [183]	2023	Health/ Academia	"AI hallucination" occurs when the information generated by ChatGPT may not always be correct.	0
160	Bryant [25]	2023	Academia	"AI hallucination" refers to answers that are fabricated when data are insufficient for an accurate response.	2
161	Mahyoob et al. [184]	2023	Academia	"AI hallucination" occurs when it generates ideas with human-like fluency and persuasiveness but without truth or factual accuracy.	0
162	Lee [185]	2023	Mathematical Setting	"AI hallucination" occurs when GPT models generate outputs that are contextually implausible or inconsistent with the real world.	14
163	Piñeiro-Martín et al. [186]	2023	Ethical Setting	"AI hallucination" occurs when the LLM generates text that goes beyond the scope of the provided input or fabricates information that is factually incorrect.	0
164	Alhaidry et al. [187]	2023	Health	"AI hallucination" occurs when ChatGPT can produce answers that appear reliable but are incorrect.	10
165	Khurana and Vaddi [188]	2023	Health/ Academia	"AI hallucination" occurs when AI generates sentences with the intent to convince the reader, which can be misleading, particularly to inexpert readers.	1
166	Li et al. [189]	2023	Health	"AI hallucination" occurs when LLMs occasionally generate fallacious and harmful assertions beyond their knowledge expertise.	11

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
167	Walczak and Cellary [190]	2023	Academia	"AI hallucination" occurs as a primary limitation associated with LLMs, manifesting as the tendency to generate errors, including mathematical, computational, and conceptual inaccuracies, without any prior indication, often characterized by their deceptive plausibility, alignment with truthful information, and conveyed in a persuasive and self-assured manner, making their detection challenging without careful scrutiny and rigorous fact verification.	0
168	Marinaccio et al. [191]	2023	Legal Setting	"AI hallucination" occurs when the output is delivered in an authoritative and convincing manner, potentially leading uninformed users to blindly accept it as truth, despite its inaccuracies or falsehoods.	0
169	Gebrael et al. [192]	2023	Health	"AI hallucination" occurs when AI models, particularly in cases like ChatGPT, generate outputs that appear plausible but are factually incorrect or unrelated to the input context.	3
170	Tiwari et al. [193]	2023	Health	"AI hallucination" occurs as false knowledge that appears convincing from a scientific perspective.	5
171	Bhattacharyya [194]	2023	Health/ Academia	"AI hallucination" occurs as a phenomenon where nonsensical or inaccurate content is generated.	10
172	Sriwastwa et al. [23]	2023	Academia	"AI hallucination" refers to the phenomenon in ChatGPT output where the text is cogent but not necessarily true, often presenting as a complete fabrication.	0
173	Dossantos et al. [195]	2023	Health	"AI hallucination" occurs when ChatGPT produces inaccurate results, particularly in specialized topics, due to a lack of depth and inaccurate details retrieved from the LLM's database, with no assurance that ChatGPT's suggestions adhere to evidence-based guidelines or best practices.	1
174	Loos et al. [196]	2023	Academia	"AI hallucination" occurs as a phenomenon characterized by the generation of incorrect or outdated information, accompanied by the failure to provide reliable sources to evaluate the generated content.	0
175	Koga [197]	2023	Academia	"AI hallucination" occurs when LLMs generate seemingly credible but fabricated information, particularly concerning when they create fictitious citations.	0
176	Birenbaum [198]	2023	Academia	"AI hallucination" occurs when the model fabricates information and provides untraceable references to support its claims.	3
177	Cascella et al. [199]	2023	Health	"AI hallucination" refers to the ability of ChatGPT to produce answers that sound believable but may be incorrect or nonsensical.	182
178	Aronson [200]	2023	Health/ Academia	"AI hallucination" refers to all the pieces of misinformation generated.	0
179	Kumar et al. [201]	2023	Academia	"AI hallucination" refers to instances where an AI chatbot generates fictional, erroneous, or unsubstantiated information in response to queries.	0
180	Gravel et al. [24]	2023	Academia	ChatGPT fabricated a convincing response that contained several factual errors.	19
181	Ferres et al. [202]	2023	Health	"AI hallucination" occurs when a model generates content that has no basis in reality, creating entirely made-up stories or facts.	8
182	Varghese and Chapiro [203]	2023	Health	"AI hallucination" occurs when faced with topics that the model has not received adequate training or supervision for, resulting in fabricated output delivered with a strong sense of certainty.	1
183	Dunn and Cianflone [204]	2023	Health	"AI hallucination" occurs as a limitation, characterized by the lack of transparency in how the output is generated.	1

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
184	Theodosiou and Read [205]	2023	Health	"AI hallucinations" is the phenomenon in which a generative AI tool confidently asserts a factual inaccuracy.	5
185	Fesenmaier and Wober [206]	2023	Academia	"AI hallucination" occurs when ChatGPT appears to create fake or inaccurate findings that seem plausible.	0
186	Sohail et al. [207]	2023	ChatGPT	"AI hallucination" occurs when ChatGPT generates new data or information that does not exist.	5
187	Karakas et al. [14]	2023	Health	"AI hallucination" occurs when the AI sometimes generates responses that sound plausible but are, in fact, incorrect. Also called "confabulation."	0
188	Lyons et al. [208]	2023	Health	"AI hallucination" occurs when LLMs create responses to prompts by sampling from the language distribution within their dataset, potentially leading to the generation of incorrect information and the propagation of biases.	0
189	Opdahl et al. [209]	2023	Journalism	"AI hallucination" occurs when the model generates plausible-sounding nonsense, including texts that contain elements not found in the input data.	1
190	Blanchard et al. [210]	2023	Health/ Academia	"AI hallucination" occurs when probabilities from the transformer model architecture can generate made-up answers, taking numerous forms, from false references in scientific reports to misinformation in journalistic articles or incorrect formulas when debugging programming code.	2
191	Lim et al. [211]	2023	Health	"AI hallucination" occurs when LLMs lack domain-specific capabilities, rendering them susceptible to generating convincing yet potentially inaccurate responses.	1
192	Kim et al. [212]	2023	Health	"AI hallucination" occurs as the possibility of providing incorrect information or exhibiting errors in the inference process.	0
193	Alqahtani et al. [213]	2023	Academia	"AI hallucination" refers to generating non-existent or incorrect content and other related concerns associated with limited contexts, reliability, and the lack of learning from experience.	15
194	Jairoun et al. [214]	2023	Health	"AI hallucination" occurs as statements that are factually inaccurate yet appear plausible to the layman.	3
195	Šlapeta [215]	2023	Health	"AI hallucination" occurs when highly confident answers are returned by AI that cannot be explained by the training data alone.	12
196	Dillion et al. [216]	2023	Science	"AI hallucination" occurs as outputs that appear sensical but are inaccurate.	28
197	Thomson et al. [217]	2023	Text Summarization	"AI hallucination" occurs when the output text includes an attribute that was not present in the input data.	3
198	Balas and Ing [218]	2023	Health	"AI hallucination" refers to producing confident responses that sound plausible yet are factually incorrect.	19
199	Salah et al. [219]	2023	Science	"AI hallucination" refers to information that sounds plausible but is entirely fabricated or not supported by the input data.	2
200	Ilicki et al. [220]	2023	Health	"AI hallucination" refers to text responses that are either nonsensical or unfaithful to the content they should use.	3
201	Jansen et al. [221]	2023	Survey Setting	"AI hallucination" refers to generating nonsensical or inappropriate responses to survey questions.	2
202	Casal and Kessler [222]	2023	Academia	"AI hallucination" refers to the tendency to invent content.	1
203	Qi et al. [223]	2023	Health	"AI hallucination" occurs when ChatGPT generates responses that appear plausible but require correction, including the invention of terms it is familiar with.	11

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
204	Lin [224]	2023	Academia	"AI hallucination" occurs when LLMs fabricate facts, creating confident-sounding statements and legitimate-looking citations that are false.	22
205	Janssen et al. [225]	2023	Health	"AI hallucination" occurs where the model generates text that is factually incorrect or nonsensical, despite appearing confident in its ability.	16
206	Shen et al. [226]	2023	Health	"AI hallucination" occurs when LLMs produce seemingly credible but incorrect responses, including the invention of terms they should be familiar with.	235
207	Thirunavukarasu [227]	2023	Health	"AI hallucination" occurs when chatGPT describes inaccurate information as lucidly as it does with correct facts.	23
208	Qadir [228]	2023	Academia	"AI hallucination" occurs when ChatGPT generates nonsensical or false information (misinformation).	157
209	Frieder et al. [229]	2023	Mathematical Setting	"AI hallucination" occurs when GPT, after answering correctly or incorrectly, tells the user unrelated information.	126
210	Borji [20]	2023	ChatGPT	"AI hallucination" refers to inaccuracies in information or statements that are not in accordance with reality or the truth, often unintentional but resulting in incorrect or misleading information, particularly in the context of chatbots. Also called "Factual errors."	158
211	OpenAI [230]	2023	ChatGPT	"AI hallucination" occurs when ChatGPT produces content that is nonsensical or untruthful in relation to certain sources.	0
212	Manakul et al. [231]	2023	Large Language Model	"AI hallucination" refers to the tendency of LLMs to hallucinate facts and fabricate information.	48
213	Nori et al. [232]	2023	Health	"AI hallucination" refers to erroneous generations by ChatGPT.	127
214	Li et al. [233]	2023	Large Language Model	"AI hallucination" occurs when LLMs generate content that conflicts with the source or cannot be verified by factual knowledge.	4
215	Zhao et al. [234]	2023	Large Language Model	"Intrinsic hallucination" refers to the generated information in conflict with the existing source. "Extrinsic hallucination" refers to the generated information that cannot be verified by the available source.	188
216	Adlakha et al. [235]	2023	Question Answering	"AI hallucination" occurs when conversational models produce factually incorrect or unsupported statements.	5
217	Athavale et al. [236]	2023	Health	"AI hallucination" refers to generating syntactically correct but factually incorrect responses that seem plausible.	0
218	Chen et al. [237]	2023	Health	"AI hallucination" occurs when ChatGPT makes assumptions on details not provided in the input data.	1
219	Stahl and Eke [238]	2023	ChatGPT	"AI hallucination" refers to mistakes that ChatGPT makes when generating text that is semantically correct but factually incorrect or even nonsensical.	1
220	Walters and Wilder [239]	2023	Academia	"AI hallucination" occurs when ChatGPT provides factually incorrect responses.	0
221	Munro and Hope [240]	2023	Health/ Academia	"AI hallucination" occurs when ChatGPT provided confident responses that seemed faithful and nonsensical.	0
222	Hashimoto and Johnson [241]	2023	Health/ Academia	"AI hallucination" occurs when LLM-generated text is based on the statistical associations of patterns of words to those seen in training data and prompts, resulting in elements such as unnecessary and unnatural repetition, lack of clarity or specificity, out-of-context content, inconsistent phrasing and arguments (particularly across long essays), incorrect word selection, and occasional incorrect statements presented as facts.	1

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
223	Jain [242]	2023	ChatGPT	"AI hallucination" occurs when ChatGPT occasionally generates responses that appear correct but are nonsensical and discordant from real-world data, resulting in data inaccuracy.	0
224	Ho et al. [243]	2023	ChatGPT	"AI hallucination" occurs when ChatGPT "hallucinates" false citations that appear convincingly real but cannot be located in any medical database.	0
225	Currie [244]	2023	ChatGPT	"AI hallucination" refer to false or misleading information produced by AI. A hallucination is a plausible response that is incorrect (it seems correct to ChatGPT but is not).	21
226	Liu et al. [245]	2023	Question Answering	"AI hallucination" is LLM-generated answer which is not factual-grounded and sometimes severely wrong.	1
227	Campbell and Jovanović [246]	2023	Generative AI	"AI hallucination" refers to the propensity to generate assertions with no factual data, often deceives users into believing they are accurate.	0
228	Kshetri [247]	2023	Generative AI	"AI hallucination" refers to ChatGPT's results that are incomplete or misleading.	0
229	Ali et al. [248]	2023	Health	"AI hallucination" occurs when AI creates ungrounded, subtly incorrect information without self-awareness.	1
230	Tay [249]	2023	Health/ Academia	"AI hallucination" refers to the phenomenon by which ChatGPT could convincingly produce factually inaccurate statements	0
231	Xu and Cohen [250]	2023	Text Summarization	"AI hallucination" occurs when the model confidently generates false information.	1
232	Kernan Ferier et al. [251]	2023	Large Language Model	"AI hallucination" commonly refers to the phenomenon where LLMs occasionally generate outputs that, although appearing plausible, deviate from the user input, previously generated context, or factual knowledge. "Intrinsic hallucination" refers to contradictions between source material (e.g., training data and prompts), where generated content contradicts the information present in the source material. "Extrinsic hallucination" refers to information generated that cannot be verified by the source material, and may include content that lacks support or contradiction within the provided source data.	0
233	Tsai et al. [252]	2023	Legal Setting	"AI hallucination" occurs when LLMs generate misleading text and information.	3
234	Xin et al. [253]	2023	Dialogue	"AI hallucination" refers to the occurrence where pre-trained text generation models occasionally generate text that is nonsensical or unfaithful to the provided source input.	0
235	Murgia et al. [254]	2023	ChatGPT	"AI hallucination" occurs when generative LLMs generate unintended text, leading to degraded system performance and unmet user expectations in real-world scenarios.	1
236	Kunze et al. [255]	2023	Large Language Model	"AI hallucination" refers to the phenomenon where the GPT model appears to be designed to provide incorrect answers rather than admit this to its users.	0
237	Cascella [256]	2023	Academia	"AI hallucination" refers to ChatGPT generating answers that sound credible but may be incorrect or nonsensical.	0
238	Watanabe [257]	2023	Government	"AI hallucination" refers to the phenomenon where AI text generators produce statements that are consistent with the system's internal logic but are not based on any true context or source.	0

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
239	Nakaura et al. [258]	2023	Health	"AI hallucination" occurs when LLMs like GPT series have the potential to generate inaccurate content.	0
240	Feuerriegel et al. [259]	2023	Generative AI	"AI hallucination" refers to mistakes in the generated text that are semantically or syntactically plausible but are actually nonsensical or incorrect. In other words, the generative AI model produces content that is not based on any facts or evidence, but rather on its own assumptions or biases. Moreover, the output of generative AI, especially that of LLMs, is typically not easily verifiable.	0
241	Hryciw et al. [260]	2023	Health	"AI hallucination" occurs when AI algorithms, which are not infallible, produce false or misleading information.	0
242	Bran et al. [261]	2023	Generative AI	"AI hallucination" refers to nonsense expressed in an authoritative tone and filling knowledge gaps with falsehoods.	5
243	Sovrano et al. [262]	2023	ChatGPT	"AI hallucination" occurs when content is generated that does not maintain fidelity to a given context or source.	0
244	Mishra et al. [263]	2023	ChatGPT	"AI hallucination" occurs when models produce outputs that are not grounded in their training data.	0
245	Lam et al. [264]	2023	Legal Setting	"AI hallucination" refers to LLMs output results are not realistic, do not follow user given context or match any data patterns that it has been trained on.	1
246	Wan et al. [265]	2023	Text Summarization	"AI hallucination" occurs when the generated summary contains facts or entities not present in the original document.	4
247	Houston and Corrado [266]	2023	Academia	"AI hallucination" commonly refers to the phenomenon where LLMs occasionally generate outputs that, although appearing plausible, deviate from the user input, previously generated context, or factual knowledge.	2
248	González-Mora et al. [267]	2023	Natural Language Generation	"AI hallucination" refers to the generation of texts that are apparently well-written but unsubstantiated and not faithful to the provided data.	2
249	Lim et al. [268]	2023	Dialogue	"AI hallucination" refers to situations where the generated output contradicts the reference knowledge and includes instances when the generated output cannot be confirmed from the knowledge source.	0
250	Alowais et al. [269]	2023	Health	"AI hallucination" refers to the tendency of AI-generated data and/or analysis to fabricate and create false information that cannot be supported by existing evidence, even though it may appear realistic and convincing.	2
251	Liu et al. [270]	2023	Health	"AI hallucination" can be defined as the capability of LLM-based chatbots to convey false information as though it were true.	33
252	Xie et al. [271]	2023	Health	"AI hallucination" refers to when a medical AI system is considered unfaithful or to have a factual inconsistency issue, as it generates content that is not supported by existing knowledge, reference, or data. Intrinsic Error: The generated output contradicts with existing knowledge, reference, or data. Extrinsic Error: The generated output cannot be confirmed (either supported or contradicted) by existing knowledge, reference, or data.	0
253	Bernstein et al. [272]	2023	Health	"AI hallucination" refers to chatbot outputs that sound convincingly plausible yet are factually inaccurate.	0

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
254	Heck [273]	2023	Health	"AI hallucination" refers to ChatGPT providing text that combines real and fabricated evidence, resulting in plausible answers, occasionally appearing as nonsense when assessed according to the common knowledge of experts in different areas, also called "mistakes." "Intrinsic hallucinations" occur when the generation output contradicts the source content. "Extrinsic hallucinations" occur when the output can neither be supported nor contradicted by the source.	3
255	Ray and Majumder [274]	2023	Health	"AI hallucination" refers to the generation of inaccurate or false information by AI-based systems such as ChatGPT.	0
256	Ang et al. [275]	2023	Health/ Academia	"AI hallucination" refers to the phenomenon where text generated by ChatGPT may appear credible but can be pure confabulation, containing a combination of both facts and fabricated information, or entirely fictitious pseudoscientific material.	5
257	Talyshinskii et al. [276]	2023	Health/ Academia	"AI hallucination" refers to a phenomenon in writing influenced more by learned patterns than scientific facts, which leads to mistakes.	0
258	Kung et al. [277]	2023	Health	"AI hallucination" refers to the phenomenon where ChatGPT, while citing a verifiable source, may draw information that is outdated or entirely incorrect, despite providing logical justifications for its answer choices, thus leading to logical errors and assertions of false facts.	0
259	Friederichs et al. [278]	2023	Health/ Academia	"AI hallucination" refers to the behavior in which wrong answers are just as convincingly justified as correct ones, a phenomenon not uncommon in large language models.	5
260	Ghorashi et al. [279]	2023	Health/ Academia	"AI hallucination" occurs when chatbots have the potential to falsify and create references.	0
261	McBee et al. [280]	2023	Sport	"AI hallucination" occurs as the generation of unsupported or false information, which is a prevalent issue with LLM-based chatbots.	0
262	Reis [281]	2023	Health	"AI hallucination" may generate plausible sounding but incorrect or nonsensical answers that does not seem to be justified by its training data, such as claim to be human.	1
263	Joachimiak et al. [282]	2023	Genetic	"AI hallucination" ocuurse when LLM model fabricated a term for a gene set.	0
264	Chen [283]	2023	Academia	"AI hallucination" refers to false or nonsense information presented as fact by an LLM.	0
265	Delsoz et al. [284]	2023	Health/ ChatGPT	"AI hallucination" occurs when ChatGPT generates responses that appear fluent and believable but may contain factual inaccuracies.	0
266	Takagi et al. [285]	2023	Health/ ChatGPT	"AI hallucination" is defined as producing nonsensical or untruthful content concerning certain sources.	11
267	Kaneda et al. [286]	2023	Health/ ChatGPT	"AI hallucination" occurs when ChatGPT provides erroneous information in a naturalistic manner.	3
268	Ahn [287]	2023	Health	"AI hallucination" refers to the generation of false or logically incorrect text that appears plausible and grammatically correct.	12
269	Jin et al. [288]	2023	Health	"AI hallucination" refers to the occurrence where LLMs usually produce plausible-sounding but incorrect outputs.	0

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
270	Ahn [289]	2023	Health/ ChatGPT	"AI hallucination" refers to a phenomenon where outputs may deviate from factual accuracy or provided context.	0
271	Liu et al. [290]	2023	Health/ ChatGPT	"AI hallucination" refers to the fact that the content generated by the model is not based on reality, creating a completely fabricated story or fact.	20
272	Doorn [291]	2023	Water Domain	"AI hallucination" refers to LLMs' tendency to generate factually nonsensical or incorrect text.	0
273	Hiesinger et al. [292]	2023	Health	"AI hallucination" refers to the propensity of LLMs to generate factually incorrect statements.	3
274	Palal et al. [293]	2023	Health/ ChatGPT	"AI hallucination" occurs when ChatGPT generates inaccurate or contradictory information.	0
275	Coskun et al. [294]	2023	Question Answering/ ChatGPT	"AI hallucination" refers to an instance where the model generates information that is not supported by existing evidence or factual data.	0
276	Baldassarre et al. [295]	2023	ChatGPT	"AI hallucination" refers to the phenomenon of inaccurate information.	0
277	Boujemaa et al. [296]	2023	Retail	"AI hallucination" occures where the model produces untruthful or misleading information.	0
278	Li et al. [297]	2023	Dialogue	"AI hallucination" are NLP generated content that appear to be relevant bbut are not actually faithful to the underlting text.	0
279	Urban et al. [298]	2023	Database	"AI hallucination" is a phenomenon where LLMs generate non-factual statements.	1
280	Mahon et al. [299]	2023	Academia	"AI hallucination" is the generation of output that appears convincing but is factually untrue or unrelated to the current context.	0
281	August et al. [300]	2023	Health	"AI hallucination" refers to the limitation in current text generation capabilities, which carries the risk of generating factually incorrect or inconsistent text.	20
282	Fischer [301]	2023	Legal Setting	"AI hallucination" refers to a phenomenon where an LM, or Language Model, is a system for haphazardly stitching together sequences of linguistic forms it has observed in its vast training data, according to probabilistic information about how they combine, but without any reference to meaning. Also called "Stochastic Parrots."	0
283	Lee et al. [302]	2023	Dialogue	"AI hallucination" are LLMs generated information that is non-factual or nonsensical.	3
284	Scotti et al. [303]	2023	Chatbot	"AI hallucination" refers to the case where agents generate responses without actually knowing the information it is talking about or without referring to some knowledge base, leading to possibly wrong/misleading information.	0
285	Zhan et al. [304]	2023	ChatGPT	"AI hallucination" refers to the phenomenon, also known as the "hallucination effect," where chatbots like ChatGPT may generate misleading and deceptive information that can have adverse impacts on users who may struggle to distinguish fact from fiction.	0
286	Vargas-Murillo et al. [305]	2023	Academia	"AI hallucination" refers to the phenomenon, also known as the "hallucination effect," which causes an AI to invent familiar terms.	0
287	White et al. [306]	2023	Text Summarization	"AI hallucination" refers to the tendency to produce content that is nonsensical or untruthful in relation to certain sources.	0
288	Cusumano [307]	2023	Technology	"AI hallucination" occurs when LLMs, in the absence of an answer to a query, use predictive analytics to make up reasonable but sometimes incorrect responses.	0

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
289	Nashid et al. [308]	2023	Large Language Model	"AI hallucination" refers to ChatGPT's outputs that are incorrect or nonsensical.	12
290	Pataranutaporn et al. [309]	2023	Technology	"AI hallucination" refers to the phenomenon where ChatGPT generates content that is nonsensical or unfaithful to the provided source content.	2
291	Faggioli et al. [310]	2023	Large Language Model	"AI hallucination" occurs when LLMs generate text that contains inaccurate or false information, often in an affirmative manner that makes it difficult for humans to even suspect errors.	0
292	Marczak-Czajka and Cleland-Huang [311]	2023	Human-Value Story	"AI hallucination" occurs when generative models produce output that is nonsensical with incorrect, bizarre logic.	0
293	Stefińska et al. [312]	2023	Quantum Physics	"AI hallucination" occurs when LLMs generate data not observed in the training dataset.	0
294	Ezenkwu [313]	2023	Customer Service	"AI hallucination" occurs when ChatGPT generates responses that are irrelevant or incorrect.	0
295	Fayyad [314]	2023	ChatGPT	"AI hallucination" occurs when generative AI models lose track of the source of information, lacking reasoning capability or semantic understanding, and instead resort to autocompletion through pattern matching, resulting in making things up.	0
296	Mrabet and Studholme [315]	2023	ChatGPT	"AI hallucination" refers to The AI's inability to understand what it has written is clear.	2
297	Pitt [316]	2023	Academia	"AI hallucination" occurs when the AI/LLM produces a plausible output that, however, does not seem to be warranted by the training data.	0
298	Crosthwaite and Baisa [317]	2023	ChatGPT	"AI hallucination" occurs when ChatGPT invents terms that lie outside of its training data.	1
299	Vinny [318]	2023	Health	"AI hallucination" occurs when LLMs generate erroneous medical information to support their opinions	0
300	Solyman et al. [319]	2023	Grammar	"AI hallucination" occurs when the system produces translations that are completely inadequate due to an overreliance on the target context in NMT.	4
301	Waqas et al. [320]	2023	Health	"AI hallucination" refers to a known limitation of generative AI, encompassing mistakes in the generated text or images that are semantically, syntactically, or visually plausible but are, in fact, incorrect, nonsensical, and do not refer to any real-world concepts.	0
302	Stephens et al. [321]	2023	Health	"AI hallucination" occurs when chatbots may propagate erroneous information or even make up information.	0
303	Dien [322]	2023	Health/ Academia	"AI hallucination" refers to the highly susceptible nature of ChatGPT to produce erroneous outputs.	5
304	Abu-Farha et al. [323]	2023	Health	"AI hallucination" refers to the generation of scientifically false content that might seem convincing to nonexperts.	0
305	Tippareddy et al. [324]	2023	Health/ Academia	"AI hallucination" refers to confident responses generated by ChatGPT without being justified by training data.	0
306	Oviedo-Trespalacios et al. [325]	2023	ChatGPT	"AI hallucination" refers to the occurrence where ChatGPT can produce answers that appear credible but may be incorrect or nonsensical.	18
307	Sarraju et al. [326]	2023	Health/ ChatGPT	"AI hallucination" refers to Inaccurate information may be presented in a confident manner, including nonexistent references to scientific literature. Also called "confabulation."	0

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
308	Coskun et al. [327]	2023	Health/ ChatGPT	"AI hallucination" refers to situations where the AI generates responses that are not derived from its training data, potentially leading to inaccuracies or misunderstandings.	0
309	Lam [328]	2023	Health	"AI hallucination" refers to occurrences where AI models output plausible but factually incorrect information.	0
310	Pantanowitz and Pantanowitz [329]	2023	Health/ ChatGPT	"AI hallucination" occurs when the AI imagines facts that are not real as being real or makes reasoning errors that it should not be making.	0
311	Boussen et al. [330]	2023	Health/ Academia	"AI hallucination" refers to the occurrence where a model might 'invent' information that seems plausible based on the patterns and structures it has learned.	0
312	Schlam et al. [331]	2023	Health/ Academia	"AI hallucination" refers to the phenomenon where NLPs often make subtle mistakes.	0
313	Lyu and Wu [332]	2023	Academia	"AI hallucination" occurs when LLM responses can sometimes be repetitive or contain non-factual information.	0
314	Sparkes [333]	2023	Chatbot	"AI hallucination" refers to the phenomenon where an AI, in response to prompts, will produce convincing statements that are actually inaccurate or totally false	0
315	Bhatia and Kulkarni [334]	2023	Academia	"AI hallucination" occurs when the AI sometimes writes plausible-sounding but incorrect, nonsensical answers.	0
316	Chatelan et al. [335]	2023	Health	"AI hallucination" occurs when ChatGPT makes up or distorts facts, including the creation of made-up references.	0
317	Lareyre et al. [336]	2023	Health	"AI hallucination" refers to the phenomenon where the model can generate content output that is incorrect or nonsensical, despite appearing reliable.	0
318	Wilkins [337]	2023	Health	"AI hallucination" occurs when the system erroneously generates "fantastical, unfaithful, or nonsensical outputs."	0
319	Piazza et al. [338]	2023	Academia	"AI hallucination" refers to the phenomenon where ChatGPT may generate output that is grammatically correct and coherent but may not be appropriate for the intended audience or purpose.	0
320	Scanlon et al. [339]	2023	Academia	"AI hallucination" refers to the phenomenon where ChatGPT prioritize generating humanlike text in response to a prompt, often leading them to focus on providing an answer rather than the correct one, resulting in inaccurate or incorrect responses presented to users with an unfounded confidence, and even generating fake bibliographic information when asked for references, which could be misleading if not critically examined.	2
321	Kaneda [340]	2023	Health	"AI hallucination" refers to a phenomenon where ChatGPT generates plausible but untrue responses.	1
322	Tan et al. [341]	2023	Health	"AI hallucination" refers to the occurrence where invented, inaccurate statements are presented as lucidly as accurate information. Also called "fact fabrication."	0
323	Ai et al. [342]	2023	Information Retrieval	"AI hallucination" occurs when LLMs may occasionally generate erroneous or nonsensical responses.	3
324	Ruma et al. [343]	2023	Natural Language Generation	"AI hallucination" occurs when an NLG system generates unfaithful or nonfactual content.	0
325	Mao et al. [344]	2023	Large Language Model	"AI hallucination" refers to a scenario where the generated content by LLMs appears plausible but is, in fact, entirely fictional.	0

Continued on next page

TABLE V – continued from previous page

Num.	Author(s)	Year	Application	Definition	Citation
326	Chaiken [345]	2023	Health	"Hallucination in AI" refers to the generation of outputs that may sound plausible but are either factually incorrect or unrelated to the given context. These outputs often emerge from the AI model's inherent biases, lack of real-world understanding, or training data limitations. In other words, the AI system "hallucinates" information that it has not been explicitly trained on, leading to unreliable or misleading responses.	0
327	Otaki [346]	2023	Academia	"AI hallucination" refers to the production of inaccurate or logically incorrect text that appears believable and grammatically sound.	0
328	Triguero et al. [347]	2023	Academia	"AI hallucination" occurs when in LLMs, the system may output untrue statements with high confidence.	0
329	Huang [348]	2023	Health/ Academia	"AI hallucination" occurs when ChatGPT produces content that may cause users to depend on measures making it challenging to determine the accuracy of specific information.	0
330	Muranga et al. [349]	2023	Academia	"AI hallucination" refers to the phenomenon where AI often produces completely false information conveyed in a convincing manner, including the invention of items such as references and citations.	0
331	Polverini and Gregorcic [350]	2023	Physics	"AI hallucination" refers to the occurrence where LLM-generated output contains factually incorrect statements.	0
332	Rajendran et al. [351]	2023	Hardware Tech	"AI hallucination" refers to a phenomenon where AI models create outputs that appear seemingly correct and confident but are, in fact, factually wrong.	0
333	Lobentanzer and Saez-Rodriguez [352]	2023	Biomedicine	"AI hallucination" refers to the phenomenon where LLMs make up facts as they go along, and, to make matters worse, are convinced - and convincing - regarding the truth of their hallucinations.	0

REFERENCES

[1] S. Baker and T. Kanade, "Hallucinating faces," in *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*, 2000, pp. 83–88.

[2] H. Xiang, Q. Zou, M. A. Nawaz, X. Huang, F. Zhang, and H. Yu, "Deep learning for image inpainting: A survey," *Pattern Recognition*, vol. 134, p. 109046, 2023.

[3] A. Pumarola, A. Agudo, A. Sanfeliu, and F. Moreno-Noguer, "Unsupervised person image synthesis in arbitrary poses," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8620–8628.

[4] A. F. Biten, L. Gómez, and D. Karatzas, "Let there be a clock on the beach: Reducing object hallucination in image captioning," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2022, pp. 1381–1390.

[5] A. Braunegg, A. Chakraborty, M. Krumdick, N. Lape, S. Leary, K. Manville, E. Merkhofer, L. Strickhart, and M. Walmer, "Apricot: A dataset of physical adversarial attacks on object detection," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXI 16*. Springer, 2020, pp. 35–50.

[6] P. Koehn and R. Knowles, "Six challenges for neural machine translation," 2017.

[7] S. Wiseman, S. M. Shieber, and A. M. Rush, "Challenges in data-to-document generation," 2017.

[8] K. Filippova, "Controlled hallucinations: Learning to generate faithfully from noisy data," *arXiv preprint arXiv:2010.05873*, 2020.

[9] T. R. Insel, "Rethinking schizophrenia," *Nature*, vol. 468, no. 7321, pp. 187–193, 2010.

[10] Z. Ji, N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, E. Ishii, Y. J. Bang, A. Madotto, and P. Fung, "Survey of hallucination in natural language generation," *ACM Computing Surveys*, vol. 55, no. 12, pp. 1–38, 2023.

[11] S. D. Østergaard and K. L. Nielbo, "False responses from artificial intelligence models are not hallucinations," *Schizophrenia Bulletin*, p. sbad068, 2023.

[12] R. Emsley, "Chatgpt: these are not hallucinations–they're fabrications and falsifications," *Schizophrenia*, vol. 9, no. 1, p. 52, 2023.

[13] H. Ye, T. Liu, A. Zhang, W. Hua, and W. Jia, "Cognitive mirage: A review of hallucinations in large language models," 2023.

[14] C. Karakas, D. Brock, and A. Lakhotia, "Leveraging chatgpt in the pediatric neurology clinic: Practical considerations for use to improve efficiency and outcomes," *Available at SSRN 4475000*, 2023.

[15] D. L. Rodgers, M. Needler, A. Robinson, R. Barnes, T. Brosche, J. Hernandez, J. Poore, P. VandeKoppel, and R. Ahmed, "Artificial intelligence and the simulationists," *Simulation in Healthcare: Journal of the Society for Simulation in Healthcare*, 2023.

[16] M. G. Madden, B. A. McNicholas, and J. G. Laffey, "Assessing the usefulness of a large language model to query and summarize unstructured medical notes in intensive care," *Intensive Care Medicine*, pp. 1–3, 2023.

[17] Z. Li, "The dark side of chatgpt: Legal and ethical challenges from stochastic parrots and hallucination," *arXiv preprint arXiv:2304.14347*, 2023.

[18] J. Ge and J. Lai, "Artificial intelligence-based text generators in hepatology: Chatgpt is just the beginning," *Hepatology Communications*, vol. 7, no. 4, 2023.

[19] N. Curtis *et al.*, "To chatgpt or not to chatgpt? the impact of artificial intelligence on academic publishing," *The Pediatric Infectious Disease Journal*, vol. 42, no. 4, p. 275, 2023.

[20] A. Borji, "A categorical archive of chatgpt failures," *arXiv preprint arXiv:2302.03494*, 2023.

[21] A. J. Thirunavukarasu, D. S. J. Ting, K. Elangovan, L. Gutierrez, T. F. Tan, and D. S. W. Ting, "Large language models in medicine," *Nature medicine*, pp. 1–11, 2023.

[22] D. S. J. Ting, T. F. Tan, and D. S. W. Ting, "Chatgpt in ophthalmology: the dawn of a new era?" *Eye*, pp. 1–4, 2023.

[23] A. Sriwastwa, P. Ravi, A. Emmert, S. Chokshi, S. Kondor, K. Dhal, P. Patel, L. L. Chepelev, F. J. Rybicki, and R. Gupta, "Generative ai for medical 3d printing: a comparison of chatgpt outputs to reference standard education," *3D Printing in Medicine*, vol. 9, no. 1, p. 21, 2023.

[24] J. Gravel, M. D'Amours-Gravel, and E. Osmanlliu, "Learning to fake it: limited responses and fabricated references provided by chatgpt for medical questions," *Mayo Clinic Proceedings: Digital Health*, vol. 1, no. 3, pp. 226–234, 2023.

[25] A. Bryant, "Ai chatbots: Threat or opportunity?" *Informatics*, vol. 10, no. 2, 2023.

[26] M. Miao, F. Meng, Y. Liu, X.-H. Zhou, and J. Zhou, "Prevent the language model from being overconfident in neural machine translation," *arXiv preprint arXiv:2105.11098*, 2021.

[27] K. Lee, O. Firat, A. Agarwal, C. Fannjiang, and D. Sussillo, "Hallucinations in neural machine translation," 2019. [Online]. Available: https://openreview.net/forum?id=SkxJ-309FQ

[28] Z. Zhao, S. B. Cohen, and B. Webber, "Reducing quantity hallucinations in abstractive summarization," *arXiv preprint arXiv:2009.13312*, 2020.

[29] Y. Dong, S. Wang, Z. Gan, Y. Cheng, J. C. K. Cheung, and J. Liu, "Multi-fact correction in abstractive text summarization," *arXiv preprint arXiv:2010.02443*, 2020.

[30] J. Maynez, S. Narayan, B. Bohnet, and R. McDonald, "On faithfulness and factuality in abstractive summarization," *arXiv preprint arXiv:2005.00661*, 2020.

[31] Y. Lyu, C. Zhu, T. Xu, Z. Yin, and E. Chen, "Faithful abstractive summarization via fact-aware consistency-constrained transformer," in *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, 2022, pp. 1410–1419.

[32] OpenAI. (2022) Chatgpt 3.5. https://www.openai.com/.

[33] C. DeVon. (2023) Ai chatbots can 'hallucinate' and make things up—why it happens and how to spot it. Accessed on Jan 03, 2024. [Online]. Available: https://www.cnbc.com/2023/12/22/why-ai-chatbots-hallucinate.html

[34] K. Weise and C. Metz. (2023) When a.i. chatbots hallucinate. Accessed on Jan 03, 2024. [Online]. Available: https://www.nytimes.com/2023/05/01/business/ai-chatbots-hallucination.html

[35] S. Diamond. (2023) A.i. chatbots, hens and humans can all 'hallucinate'. Accessed on Jan 03, 2024. [Online]. Available: https://www.nytimes.com/2023/12/17/insider/ai-chatbots-humans-hallucinate.html

[36] C. Thorbecke. (2023) Ai tools make things up a lot, and that's a huge problem. Accessed on Jan 03, 2024. [Online]. Available: https://www.cnn.com/2023/08/29/tech/ai-chatbot-hallucinations/index.html

[37] G. Press. (2023) Celebrating 80 years of hallucinating about artificial intelligence. Accessed on Jan 03, 2024. [Online]. Available: https://www.forbes.com/sites/gilpress/2023/12/30/celebrating-80-years-of-hallucinatin

[38] C. Dictionary. (2023) The cambridge dictionary word of the year 2023 is.... Accessed on Jan 03, 2024. [Online]. Available: https://dictionary.cambridge.org/editorial/woty

[39] Dictionary.com. (2023) Word of the year. Accessed on Jan 03, 2024. [Online]. Available: https://content.dictionary.com/word-of-the-year-2023/

[40] K. Lee, O. Firat, A. Agarwal, C. Fannjiang, and D. Sussillo, "Hallucinations in neural machine translation," 2018.

[41] F. Nie, J.-G. Yao, J. Wang, R. Pan, and C.-Y. Lin, "A simple recipe towards reducing hallucination in neural surface realisation," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 2019, pp. 2673–2679.

[42] R. Tian, S. Narayan, T. Sellam, and A. P. Parikh, "Sticking to the facts: Confident decoding for faithful data-to-text generation," *arXiv preprint arXiv:1910.08684*, 2019.

[43] O. Dušek, D. M. Howcroft, and V. Rieser, "Semantic noise matters for neural natural language generation," *arXiv preprint arXiv:1911.03905*, 2019.

[44] T. C. Ferreira, C. van der Lee, E. Van Miltenburg, and E. Krahmer, "Neural data-to-text generation: A comparison between pipeline and end-to-end architectures," *arXiv preprint arXiv:1908.09022*, 2019.

[45] M. Martindale, M. Carpuat, K. Duh, and P. McNamee, "Identifying fluently inadequate output in neural and statistical machine translation," in *Proceedings of Machine Translation Summit XVII: Research Track*, 2019, pp. 233–243.

[46] O. Dušek and Z. Kasner, "Evaluating semantic accuracy of data-to-text generation with natural language inference," *arXiv preprint arXiv:2011.10819*, 2020.

[47] D. Kang and T. Hashimoto, "Improved natural language generation via loss truncation," *arXiv preprint arXiv:2004.14589*, 2020.

[48] A. P. Parikh, X. Wang, S. Gehrmann, M. Faruqui, B. Dhingra, D. Yang, and D. Das, "Totto: A controlled table-to-text generation dataset," *arXiv preprint arXiv:2004.14373*, 2020.

[49] E. Durmus, H. He, and M. Diab, "Feqa: A question answering evaluation framework for faithfulness assessment in abstractive summarization," *arXiv preprint arXiv:2005.03754*, 2020.

[50] C. Zhou, G. Neubig, J. Gu, M. Diab, P. Guzman, L. Zettlemoyer, and M. Ghazvininejad, "Detecting hallucinated content in conditional neural sequence generation," *arXiv preprint arXiv:2011.02593*, 2020.

[51] H. Elsahar, M. Coavoux, M. Gallé, and J. Rozen, "Self-supervised and controlled multi-document opinion summarization," *arXiv preprint arXiv:2004.14754*, 2020.

[52] X. Tang, A. Nair, B. Wang, B. Wang, J. Desai, A. Wade, H. Li, A. Celikyilmaz, Y. Mehdad, and D. Radev, "Confit: Toward faithful dialogue summarization with linguistically-informed contrastive fine-tuning," *arXiv preprint arXiv:2112.08713*, 2021.

[53] M. Cao, Y. Dong, and J. C. K. Cheung, "Hallucinated but factual! inspecting the factuality of hallucinations in abstractive summarization," *arXiv preprint arXiv:2109.09784*, 2021.

[54] S. Chen, F. Zhang, K. Sone, and D. Roth, "Improving faithfulness in abstractive summarization with contrast candidate generation and selection," *arXiv preprint arXiv:2104.09061*, 2021.

[55] H. Wang, Y. Gao, Y. Bai, M. Lapata, and H. Huang, "Exploring explainable selection to control abstractive summarization," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 15, 2021, pp. 13 933–13 941.

[56] Y. Xiao and W. Y. Wang, "On hallucination and predictive uncertainty in conditional language generation," *arXiv preprint arXiv:2103.15025*, 2021.

[57] N. Dziri, A. Madotto, O. Zaiane, and A. J. Bose, "Neural path hunter: Reducing hallucination in dialogue systems via path grounding," *arXiv preprint arXiv:2104.08455*, 2021.

[58] T. Liu, Y. Zhang, C. Brockett, Y. Mao, Z. Sui, W. Chen, and B. Dolan, "A token-level reference-free hallucination detection benchmark for free-form text generation," *arXiv preprint arXiv:2104.08704*, 2021.

[59] Y. Huang, X. Feng, X. Feng, and B. Qin, "The factual inconsistency problem in abstractive text summarization: A survey," *arXiv preprint arXiv:2104.14839*, 2021.

[60] S. Lin, J. Hilton, and O. Evans, "Truthfulqa: Measuring how models mimic human falsehoods," *arXiv preprint arXiv:2109.07958*, 2021.

[61] K. Shuster, S. Poff, M. Chen, D. Kiela, and J. Weston, "Retrieval augmentation reduces hallucination in conversation," *arXiv preprint arXiv:2104.07567*, 2021.

[62] I. Sekulić, M. Aliannejadi, and F. Crestani, "Towards facet-driven generation of clarifying questions for conversational search," in *Proceedings of the 2021 ACM SIGIR international conference on theory of information retrieval*, 2021, pp. 167–175.

[63] S. Ghosh, Z. Qi, S. Chaturvedi, and S. Srivastava, "How helpful is inverse reinforcement learning for table-to-text generation?" in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, 2021, pp. 71–79.

[64] L. Perez-Beltrachini and M. Lapata, "Multi-document summarization with determinantal point process attention," *Journal of Artificial Intelligence Research*, vol. 71, pp. 371–399, 2021.

[65] R. Ali, O. Y. Tang, I. D. Connolly, J. S. Fridley, J. H. Shin, P. L. Z. Sullivan, D. Cielo, A. A. Oyelese, C. E. Doberstein, A. E. Telfeian *et al.*, "Performance of chatgpt, gpt-4, and google bard on a neurosurgery oral boards preparation question bank," *Neurosurgery*, pp. 10–1227, 2022.

[66] K. Ando, T. Okumura, M. Komachi, H. Horiguchi, and Y. Matsumoto, "Is artificial intelligence capable of generating hospital discharge summaries from inpatient records?" *PLOS Digital Health*, vol. 1, no. 12, p. e0000158, 2022.

[67] ——, "Exploring optimal granularity for extractive summarization of unstructured health records: Analysis of the largest multi-institutional archive of health records in japan," *PLOS Digital Health*, vol. 1, no. 9, p. e0000099, 2022.

[68] C. Rebuffel, M. Roberti, L. Soulier, G. Scoutheeten, R. Cancelliere, and P. Gallinari, "Controlling hallucinations at word level in data-to-text generation," *Data Mining and Knowledge Discovery*, pp. 1–37, 2022.

[69] D. Wan and M. Bansal, "Factpegasus: Factuality-aware pre-training and fine-tuning for abstractive summarization," *arXiv preprint arXiv:2205.07830*, 2022.

[70] J. G. Corbelle, A. B. Diz, J. Alonso-Moral, and J. Taboada, "Dealing with hallucination and omission in neural natural language generation: A use case on meteorology." in *Proceedings of the 15th International Conference on Natural Language Generation*, 2022, pp. 121–130.

[71] N. Lee, W. Ping, P. Xu, M. Patwary, P. N. Fung, M. Shoeybi, and B. Catanzaro, "Factuality enhanced language models for open-ended text generation," *Advances in Neural Information Processing Systems*, vol. 35, pp. 34 586–34 599, 2022.

[72] M. A. S. Cabezudo and T. A. S. Pardo, "Exploring a pos-based two-stage approach for improving low-resource amr-to-text generation," in *Proceedings*, 2022.

[73] L. van der Poel, R. Cotterell, and C. Meister, "Mutual information alleviates hallucinations in abstractive summarization," *arXiv preprint arXiv:2210.13210*, 2022.

[74] N. Dziri, S. Milton, M. Yu, O. Zaiane, and S. Reddy, "On the origin of hallucinations in conversational models: Is it the datasets or the models?" *arXiv preprint arXiv:2204.07931*, 2022.

[75] N. Dziri, E. Kamalloo, S. Milton, O. Zaiane, M. Yu, E. M. Ponti, and S. Reddy, "Faithdial: A faithful benchmark for information-seeking dialogue," *Transactions of the Association for Computational Linguistics*, vol. 10, pp. 1473–1490, 2022.

[76] F. Koto, T. Baldwin, and J. H. Lau, "Ffci: A framework for interpretable automatic evaluation of summarization," *Journal of Artificial Intelligence Research*, vol. 73, pp. 1553–1607, 2022.

[77] S. M. Goodman, E. Buehler, P. Clary, A. Coenen, A. Donsbach, T. N. Horne, M. Lahav, R. MacDonald, R. B. Michaels, A. Narayanan *et al.*, "Lampost: Design and evaluation of an ai-assisted email writing prototype for adults with dyslexia," in *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility*, 2022, pp. 1–18.

[78] S. Gehrmann, E. Clark, and T. Sellam, "Repairing the cracked foundation: A survey of obstacles in evaluation practices for generated text," *Journal of Artificial Intelligence Research*, vol. 77, pp. 103–166, 2023.

[79] E. Erdem, M. Kuyu, S. Yagcioglu, A. Frank, L. Parcalabescu, B. Plank, A. Babii, O. Turuta, A. Erdem, I. Calixto *et al.*, "Neural natural language generation: A survey on multilinguality, multimodality, controllability and learning," *Journal of Artificial Intelligence Research*, vol. 73, pp. 1131–1207, 2022.

[80] Z. Y. Tun, A. Speggiorin, J. Dalton, and M. Stamper, "Comex: A multi-task benchmark for knowledge-grounded conversational media exploration," in *Proceedings of the 4th Conference on Conversational User Interfaces*, 2022, pp. 1–11.

[81] S. Gurrapu, L. Huang, and F. A. Batarseh, "Exclaim: Explainable neural claim verification using rationalization," in *2022 IEEE 29th Annual Software Technology Conference (STC)*. IEEE, 2022, pp. 19–26.

[82] J. Yang, D. Vega-Oliveros, T. Seibt, and A. Rocha, "Explainable fact-checking through question answering," in *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022, pp. 8952–8956.

[83] P. Narayanan Venkit, S. Gautam, R. Panchanadikar, T.-H. Huang, and S. Wilson, "Unmasking nationality bias: A study of human perception of nationalities in ai-generated articles," in *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, 2023, pp. 554–565.

[84] F. Leiser, S. Eckhardt, M. Knaeble, A. Maedche, G. Schwabe, and A. Sunyaev, "From chatgpt to factgpt: A participatory design study to mitigate the effects of large language model hallucinations on users," in *Mensch und Computer 2023*, 2023, pp. 81–90.

[85] S. Wang, N. Cooper, M. Eby, and E. S. Jo, "From human-centered to social-centered artificial intelligence: Assessing chatgpt's impact through disruptive events," *arXiv preprint arXiv:2306.00227*, 2023.

[86] W. Ma, S. Liu, W. Wang, Q. Hu, Y. Liu, C. Zhang, L. Nie, and Y. Liu, "The scope of chatgpt in software engineering: A thorough investigation," *arXiv preprint arXiv:2305.12138*, 2023.

[87] S. A. Vaghefi, Q. Wang, V. Muccione, J. Ni, M. Kraus, J. Bingler, T. Schimanski, C. Colesanti-Senni, N. Webersinke, C. Huggel, and M. Leippold, "chatclimate: Grounding conversational ai in climate science," 2023.

[88] X. Daull, P. Bellot, E. Bruno, V. Martin, and E. Murisasco, "Complex qa and language models hybrid architectures, survey," *arXiv preprint arXiv:2302.09051*, 2023.

[89] Y. Zhang, Y. Li, L. Cui, D. Cai, L. Liu, T. Fu, X. Huang, E. Zhao, Y. Zhang, Y. Chen *et al.*, "Siren's song in the ai ocean: A survey on hallucination in large language models," *arXiv preprint arXiv:2309.01219*, 2023.

[90] O. Romanko, A. Narayan, and R. H. Kwon, "Chatgpt-based investment portfolio selection," *arXiv preprint arXiv:2308.06260*, 2023.

[91] P. Henderson, T. Hashimoto, and M. Lemley, "Where's the liability in harmful ai speech?" *arXiv preprint arXiv:2308.04635*, 2023.

[92] J. Ni, J. Bingler, C. Colesanti-Senni, M. Kraus, G. Gostlow, T. Schimanski, D. Stammbach, S. A. Vaghefi, Q. Wang, N. Webersinke *et al.*, "Chatreport: Democratizing sustainability disclosure analysis through llm-based tools," *arXiv preprint arXiv:2307.15770*, 2023.

[93] R. Mahmood, G. Wang, M. Kalra, and P. Yan, "Fact-checking of ai-generated reports," *arXiv preprint arXiv:2307.14634*, 2023.

[94] A. Goyal, M. Siddique, N. Parekh, Z. Schwitzky, C. Broekaert, C. Michelotti, A. Wong, L. Y. Cheung, R. O. Hanlon, M. De Choudhury *et al.*, "Chatgpt and bard responses to polarizing questions," *arXiv preprint arXiv:2307.12402*, 2023.

[95] S. Zhang, "Bridging intelligence and instinct: A new control paradigm for autonomous robots," *arXiv preprint arXiv:2307.10690*, 2023.

[96] S. Li, S. Park, I. Lee, and O. Bastani, "Trac: Trustworthy retrieval augmented chatbot," *arXiv preprint arXiv:2307.04642*, 2023.

[97] Z. Li, S. Zhang, H. Zhao, Y. Yang, and D. Yang, "Batgpt: A bidirectional autoregessive talker from generative pre-trained transformer," *arXiv preprint arXiv:2307.00360*, 2023.

[98] S. Curran, S. Lansley, and O. Bethell, "Hallucination is the last thing you need," *arXiv preprint arXiv:2306.11520*, 2023.

[99] P. Feldman, J. R. Foulds, and S. Pan, "Trapping llm hallucinations using tagged context prompts," *arXiv preprint arXiv:2306.06085*, 2023.

[100] A. Mukherjee and H. Chang, "The creative frontier of generative ai: Managing the novelty-usefulness tradeoff," *arXiv preprint arXiv:2306.03601*, 2023.

[101] M. Salvagno, F. S. Taccone, and A. G. Gerli, "Artificial intelligence hallucinations," *Critical Care*, vol. 27, no. 1, pp. 1–2, 2023.

[102] G. Beutel, E. Geerits, and J. T. Kielstein, "Artificial hallucination: Gpt on lsd?" *Critical Care*, vol. 27, no. 1, p. 148, 2023.

[103] R. Azamfirei, S. R. Kudchadkar, and J. Fackler, "Large language models and the perils of their hallucinations," *Critical Care*, vol. 27, no. 1, pp. 1–2, 2023.

[104] J. G. Meyer, R. J. Urbanowicz, P. C. Martin, K. O'Connor, R. Li, P.-C. Peng, T. J. Bright, N. Tatonetti, K. J. Won, G. Gonzalez-Hernandez *et al.*, "Chatgpt and large language models in academia: opportunities and challenges," *Biodata Mining*, vol. 16, no. 1, p. 20, 2023.

[105] P. Hernigou and M. M. Scarlat, "Two minutes of orthopaedics with chatgpt: it is just the beginning; it's going to be hot, hot, hot!" *International Orthopaedics*, pp. 1–7, 2023.

[106] H. R. Patil. (2023) Don't be a victim of ai hallucinations. Accessed on Sep 10, 2023. [Online]. Available: https://www.accountingtoday.com/opinion/dont-be-a-victim-of-ai-hallucinations

[107] K. Alexander, C. Savvidou, and C. Alexander, "Who wrote this essay? detecting ai-generated writing in second language education in higher education." *Teaching English with Technology*, vol. 23, no. 2, 2023.

[108] I. Lyell, "What history teachers need to know about chatgpt," *agora*, vol. 58, no. 2, pp. 3–7, 2023.

[109] S. Grassini, "Shaping the future of education: exploring the potential and consequences of ai and chatgpt in educational settings," *Education Sciences*, vol. 13, no. 7, p. 692, 2023.

[110] G. L. Brodeur, G. Hall, and E. Tynch, "Chatgpt for legal and tax professionals," *The CPA Journal*, vol. 93, no. 7/8, pp. 68–71, 2023.

[111] S.-W. Lee and W.-J. Choi, "Utilizing chatgpt in clinical research related to anesthesiology: a comprehensive review of opportunities and limitations," *Anesthesia and Pain Medicine*, vol. 18, no. 3, pp. 244–251, 2023.

[112] T. Chatfield, "Ai hallucination," *New Philosopher*, p. 76, June 2023, accessed 10 Sept. 2023.

[113] A. McGowan, Y. Gui, M. Dobbs, S. Shuster, M. Cotter, A. Selloni, M. Goodman, A. Srivastava, G. A. Cecchi, and C. M. Corcoran, "Chatgpt and bard exhibit spontaneous citation fabrication during psychiatry literature search," *Psychiatry Research*, vol. 326, p. 115334, 2023.

[114] R. T. Wu and R. R. Dang, "Chatgpt in head and neck scientific writing: A precautionary anecdote," *American Journal of Otolaryngology*, vol. 44, no. 6, p. 103980, 2023.

[115] H. Wang, W. Wu, Z. Dou, L. He, and L. Yang, "Performance and exploration of chatgpt in medical examination, records and education in chinese: Pave the way for medical ai," *International Journal of Medical Informatics*, vol. 177, p. 105173, 2023.

[116] W. Hou and Z. Ji, "Geneturing tests gpt models in genomics," *bioRxiv*, pp. 2023–03, 2023.

[117] J. Au Yeung, Z. Kraljevic, A. Luintel, A. Balston, E. Idowu, R. J. Dobson, and J. T. Teo, "Ai chatbots not yet ready for clinical use," *Frontiers in Digital Health*, vol. 5, p. 60, 2023.

[118] H.-U. Hua, A.-H. Kaakour, A. Rachitskaya, S. Srivastava, S. Sharma, and D. A. Mammo, "Evaluation and comparison of ophthalmic scientific abstracts and references by current artificial intelligence chatbots," *JAMA ophthalmology*, 2023.

[119] D. T. Brameier, A. A. Alnasser, J. M. Carnino, A. R. Bhashyam, A. G. von Keudell, and M. J. Weaver, "Artificial intelligence in orthopaedic surgery: Can a large language model "write" a believable orthopaedic journal article?" *JBJS*, vol. 105, no. 17, pp. 1388–1392, 2023.

[120] P. Lee, S. Bubeck, and J. Petro, "Benefits, limits, and risks of gpt-4 as an ai chatbot for medicine," *New England Journal of Medicine*, vol. 388, no. 13, pp. 1233–1239, 2023.

[121] C. Long, K. Lowe, A. dos Santos, J. Zhang, A. Alanazi, D. O'Brien, E. Wright, and D. Cote, "Evaluating chatgpt-4 in otolaryngology–head and neck surgery board examination using the cvsa model," *medRxiv*, 2023. [Online]. Available: https://www.medrxiv.org/content/early/2023/06/01/2023.05.30.23290758

[122] D. Zhang, "Should chatgpt and bard share revenue with their data providers? a new business model for the ai era," *arXiv preprint arXiv:2305.02555*, 2023.

[123] P. Puchert, P. Poonam, C. van Onzenoodt, and T. Ropinski, "Llmmaps–a visual metaphor for stratified evaluation of large language models," *arXiv preprint arXiv:2304.00457*, 2023.

[124] Y. Wang, S. Visweswaran, S. Kappor, S. Kooragayalu, and X. Wu, "Chatgpt, enhanced with clinical practice guidelines, is a superior decision support tool," *medRxiv*, pp. 2023–08, 2023.

[125] R. K. Garg, V. L. Urs, A. A. Agrawal, S. K. Chaudhary, V. Paliwal, and S. K. Kar, "Exploring the role of chat gpt in patient care (diagnosis and treatment) and medical research: A systematic review," *medRxiv*, pp. 2023–06, 2023.

[126] C. Han, D. W. Kim, S. Kim, S. C. You, S. Bae, and D. Yoon, "Large-language-model-based 10-year risk prediction of cardiovascular disease: insight from the uk biobank data," *medRxiv*, pp. 2023–05, 2023.

[127] N. T. Dolan and J. Freer, "5 things to know about generative text ai tools... that might be outdated or upgraded by the time of publication," 2023.

[128] J. Scott-Branch, R. Laws, and P. Terzi, "The intersection of ai, information and digital literacy: Harnessing chatgpt and other generative tools to enhance teaching and learning," 2023.

[129] I. Larsen-Ledet, "Embracing data noise," 2023.

[130] J. K. Kim, M. Chua, M. Rickard, and A. Lorenzo, "Chatgpt and large language model (llm) chatbots: the current state of acceptability and a proposal for guidelines on utilization in academic medicine," *Journal of Pediatric Urology*, 2023.

[131] J. R. H. Tay, N. Ethan, D. Y. Chow, and C. P. C. Sim, "The use of artificial intelligence to aid in oral hygiene education: A scoping review," *Journal of Dentistry*, p. 104564, 2023.

[132] L. Z. Cai, A. Shaheen, A. Jin, R. Fukui, S. Y. Jonathan, N. Yannuzzi, and C. Alabiad, "Performance of generative large language models on ophthalmology board style questions," *American Journal of Ophthalmology*, 2023.

[133] I. Jahic, M. Ebner, and S. Schön, "Harnessing the power of artificial intelligence and chatgpt in education–a first rapid literature review," *Proceedings of EdMedia+ Innovate Learning*, vol. 2023, pp. 1462–1470, 2023.

[134] B. Randell and B. Coghlan, "Chatgpt's astonishing fabrications about percy ludgate," *IEEE Annals of the History of Computing*, vol. 45, no. 2, pp. 71–72, 2023.

[135] T. D. Brender, "Chatbot confabulations are not hallucinations—reply," *JAMA Internal Medicine*, 2023.

[136] I. Pedersen, "Adapting to ai writing," 2023.

[137] A. Munoz, A. Wilson, B. Pereira Nunes, C. Del Medico, C. Slade, D. Bennett, D. Tyler, E. Seymour, G. Hepplewhite, H. Randell-Moon *et al.*, "Aain generative artificial intelligence guidelines," 2023.

[138] R. Hatem, B. Simmons, and J. E. Thornton, "A call to address ai "hallucinations" and how healthcare professionals can mitigate their risks," *Cureus*, vol. 15, no. 9, 2023.

[139] S. A. Athaluri, S. V. Manthena, V. K. M. Kesapragada, V. Yarlagadda, T. Dave, and R. T. S. Duddumpudi, "Exploring the boundaries of reality: investigating the phenomenon of artificial intelligence hallucination in scientific writing through chatgpt references," *Cureus*, vol. 15, no. 4, 2023.

[140] T. Gorichanaz, "Accused: How students respond to allegations of using chatgpt on assessments," *Learning: Research and Practice*, pp. 1–14, 2023.

[141] S. R. Cox, A. Abdul, and W. T. Ooi, "Prompting a large language model to generate diverse motivational messages: A comparison with human-written messages," *arXiv preprint arXiv:2308.13479*, 2023.

[142] O. P. Yadava, "Chatgpt—a foe or an ally?" *Indian Journal of Thoracic and Cardiovascular Surgery*, vol. 39, no. 3, pp. 217–221, 2023.

[143] Y. Dai, A. Liu, and C. P. Lim, "Reconceptualizing chatgpt and generative ai as a student-driven innovation in higher education," Apr 2023. [Online]. Available: edarxiv.org/nwqju

[144] L. Lingard, "Writing with chatgpt: An illustration of its capacity, limitations & implications for academic writers," *Perspectives on Medical Education*, vol. 12, no. 1, p. 261, 2023.

[145] T. Woodland, "Chatgpt for improving medical education: Proceed with caution," *Mayo Clinic Proceedings: Digital Health*, vol. 1, no. 3, pp. 294–295, 2023.

[146] C. Kashangura, "The human-machine evo-ecosystem; the emergence of a non-life species that adapts through the algorithm and not dna or rna: Is our human dna ready?" 2023.

[147] H. L. Walker, S. Ghani, C. Kuemmerli, C. A. Nebiker, B. P. Müller, D. A. Raptis, and S. M. Staubli, "Reliability of medical information provided by chatgpt: Assessment against clinical guidelines and patient information quality instrument," *Journal of Medical Internet Research*, vol. 25, p. e47479, 2023.

[148] H. Tenge Hansen and T. Røsand Valø, "Cybersecurity mindfulness in the age of mindless ais: Investigating ai assistants impact in high-reliability organizations," Master's thesis, University of Agder, 2023.

[149] M. Tupper, I. W. Hendy, and J. R. Shipway, "Field courses for dummies: can chatgpt design a higher education field course?" Aug 2023.

[150] A. A. Borkowski, C. E. Jakey, S. M. Mastorides, A. L. Kraus, G. Vidyarthi, N. Viswanadhan, and J. L. Lezama, "Applications of chatgpt and large language models in medicine and health care: Benefits and pitfalls." *Federal Practitioner*, vol. 40, no. 6, 2023.

[151] R. Subramanya and P. Furlong, "Revolutionising rulemaking: How digitised rules can accelerate digital transformation," 2023.

[152] L. Y. G. Tenzer, "Defamation in the age of artificial intelligence," *Available at SSRN 4545070*, 2023.

[153] B. Alarie and R. McCreight, "The ethics of generative ai in tax practice," *Tax Notes Federal*, pp. 785–793, 2023.

[154] J. Han, C. Gong, W. Qiu, and E. Lichtfouse, "What does ai think of my paper?" *Available at SSRN 4525950*, 2023.

[155] P. G. Picht, "Chatgpt, microsoft and competition law–nemesis or fresh chance for digital markets enforcement?" *Microsoft and Competition Law–Nemesis or Fresh Chance for Digital Markets Enforcement*, 2023.

[156] P. Treleaven, J. Barnett, D. Brown, A. Bud, E. Fenoglio, C. Kerrigan, A. Koshiyama, S. Sfeir-Tait, and M. Schoernig. (2023, July) The future of cybercrime: Ai and emerging technologies are creating a cybercrime tsunami. [Online]. Available: https://ssrn.com/abstract=4507244

[157] H. Ariyaratne. (2023, May) Chatgpt and intermediary liability: Why section 230 does not and should not protect generative algorithms. [Online]. Available: https://ssrn.com/abstract=4422583

[158] T. Wu, S. He, J. Liu, S. Sun, K. Liu, Q.-L. Han, and Y. Tang, "A brief overview of chatgpt: The history, status quo and potential future development," *IEEE/CAA Journal of Automatica Sinica*, vol. 10, no. 5, pp. 1122–1136, 2023.

[159] Y. Wang, Y. Pan, M. Yan, Z. Su, and T. H. Luan, "A survey on chatgpt: Ai-generated contents, challenges, and solutions," *arXiv preprint arXiv:2305.18339*, 2023.

[160] I. Amaro, P. Barra, A. Della Greca, R. Francese, and C. Tucci, "Believe in artificial intelligence? a user study on the chatgpt's fake information impact," *IEEE Transactions on Computational Social Systems*, 2023.

[161] H. K. Skrodelis, A. Romanovs, N. Zenina, and H. Gorskis, "The latest in natural language generation: Trends, tools and applications in industry," in *2023 IEEE 10th Jubilee Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE)*. IEEE, 2023, pp. 1–5.

[162] M. Gupta, C. Akiri, K. Aryal, E. Parker, and L. Praharaj, "From chatgpt to threatgpt: Impact of generative ai in cybersecurity and privacy," *IEEE Access*, 2023.

[163] A. Bahrini, M. Khamoshifar, H. Abbasimehr, R. J. Riggs, M. Esmaeili, R. M. Majdabadkohne, and M. Pasehvar, "Chatgpt: Applications, opportunities, and threats," in *2023 Systems and Information Engineering Design Symposium (SIEDS)*. IEEE, 2023, pp. 274–279.

[164] O. H. Hamid, "Chatgpt and the chinese room argument: An eloquent ai conversationalist lacking true understanding and consciousness," in *2023 9th International Conference on Information Technology Trends (ITT)*. IEEE, 2023, pp. 238–241.

[165] D. De Silva, N. Mills, M. El-Ayoubi, M. Manic, and D. Alahakoon, "Chatgpt and generative ai guidelines for addressing academic integrity and augmenting pre-existing chatbots," in *2023 IEEE International Conference on Industrial Technology (ICIT)*. IEEE, 2023, pp. 1–6.

[166] S. Atallah, N. Banda, A. Banda, and N. Roeck, "How large language models including generative pre-trained transformer (gpt) 3 and 4 will impact medicine and surgery," *Techniques in Coloproctology*, pp. 1–6, 2023.

[167] M. Byrne, "The disruptive impacts of next generation generative artificial intelligence," *CIN: Computers, Informatics, Nursing*, vol. 41, no. 7, pp. 479–481, 2023.

[168] K. Beam, P. Sharma, P. Levy, and A. L. Beam, "Artificial intelligence in the neonatal intensive care unit: the time is now," *Journal of Perinatology*, pp. 1–5, 2023.

[169] M. Komorowski, M. del Pilar Arias López, and A. C. Chang, "How could chatgpt impact my practice as an intensivist? an overview of potential applications, risks and limitations," *Intensive Care Medicine*, pp. 1–4, 2023.

[170] M. Sallam, "Chatgpt utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns," in *Healthcare*, vol. 11, no. 6. MDPI, 2023, p. 887.

[171] A. J. Rothschild, "Artificial intelligence and the journal of clinical psychopharmacology," pp. 397–398, 2023.

[172] E.-O. Editor(s): Im, "Ai and nursing knowledge generation," *Advances in Nursing Science*, August 2023.

[173] A. Hulman, O. L. Dollerup, J. F. Mortensen, M. E. Fenech, K. Norman, H. Støvring, and T. K. Hansen, "Chatgpt-versus human-generated answers to frequently asked questions about diabetes: A turing test-inspired survey among employees of a danish diabetes center," *Plos one*, vol. 18, no. 8, p. e0290773, 2023.

[174] B. H. H. Cheung, G. K. K. Lau, G. T. C. Wong, E. Y. P. Lee, D. Kulkarni, C. S. Seow, R. Wong, and M. T.-H. Co, "Chatgpt versus human in generating medical graduate exam multiple choice questions—a multinational prospective study (hong kong sar, singapore, ireland, and the united kingdom)," *PloS one*, vol. 18, no. 8, p. e0290691, 2023.

[175] L. S. Moskatel and N. Zhang, "The utility of chatgpt in the assessment of literature on the prevention of migraine: an observational, qualitative study," *Frontiers in Neurology*, vol. 14, 2023.

[176] M. Javid, M. Reddiboina, and M. Bhandari, "Emergence of artificial generative intelligence and its potential impact on urology," *The Canadian Journal of Urology*, vol. 30, no. 4, pp. 11 588–11 598, 2023.

[177] M. Buholayka, R. Zouabi, and A. Tadinada, "The readiness of chatgpt to write scientific case reports independently: A comparative evaluation between human and artificial intelligence," *Cureus*, vol. 15, no. 5, 2023.

[178] H. Alkaissi and S. I. McFarlane, "Artificial hallucinations in chatgpt: implications in scientific writing," *Cureus*, vol. 15, no. 2, 2023.

[179] G. Eysenbach *et al.*, "The role of chatgpt, generative language models, and artificial intelligence in medical education: a conversation with chatgpt and a call for papers," *JMIR Medical Education*, vol. 9, no. 1, p. e46885, 2023.

[180] S. Thorne, "Experimenting with chatgpt for spreadsheet formula generation: Evidence of risk in ai generated spreadsheets," *arXiv preprint arXiv:2309.00095*, 2023.

[181] Y. Huang, A. Gomaa, T. Weissmann, J. Grigo, H. B. Tkhayat, B. Frey, U. S. Gaipl, L. V. Distel, A. Maier, R. Fietkau *et al.*, "Benchmarking chatgpt-4 on acr radiation oncology in-training exam (txit): Potentials and challenges for ai-assisted medical education and decision making in radiation oncology," *arXiv preprint arXiv:2304.11957*, 2023.

[182] K. Węcel, M. Sawiński, M. Stróżyna, W. Lewoniewski, P. Stolarski, E. Księżniak, and W. Abramowicz, "Artificial intelligence - friend or foe in fake news campaigns," *Economics and Business Review*, vol. 9, no. 2, pp. 41–70, 2023.

[183] A. M. Bhatti, "Ai in medical research – chatgpt and beyond," *Pakistan Armed Forces Medical Journal*, vol. 73, no. 4, pp. 954–954, 2023.

[184] M. Mahyoob, J. Al-Garaady, and A. Alblwi, "A proposed framework for human-like language processing of chatgpt in academic writing," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 18, no. 14, 2023.

[185] M. Lee, "A mathematical investigation of hallucination and creativity in gpt models," *Mathematics*, vol. 11, no. 10, p. 2320, 2023.

[186] A. Piñeiro-Martín, C. García-Mateo, L. Docío-Fernández, and M. d. C. López-Pérez, "Ethical challenges in the development of virtual assistants powered by large language models," *Electronics*, vol. 12, no. 14, p. 3170, 2023.

[187] H. M. Alhaidry, B. Fatani, J. O. Alrayes, A. M. Almana, N. K. Alfhaed, H. Alhaidry, J. Alrayes, A. Almana, and N. K. Alfhaed Sr, "Chatgpt in dentistry: A comprehensive review," *Cureus*, vol. 15, no. 4, 2023.

[188] S. Khurana and A. Vaddi, "Chatgpt from the perspective of an academic oral and maxillofacial radiologist," *Cureus*, vol. 15, no. 6, 2023.

[189] Y. Li, Z. Li, K. Zhang, R. Dan, S. Jiang, and Y. Zhang, "Chatdoctor: A medical chat model fine-tuned on a large language model meta-ai (llama) using medical domain knowledge," *Cureus*, vol. 15, no. 6, 2023.

[190] K. Walczak and W. Cellary, "Challenges for higher education in the era of widespread access to generative ai," *Economics and Business Review*, vol. 9, no. 2, pp. 71–100, 2023.

[191] R. Marinaccio, A. M. Clark, and J. O'Connor, "Harnessing the power of ai: Legal considerations for employers," *Rochester Business Journal*, vol. 39, no. 10, p. 12, Aug 04 2023. [Online]. Available: https://www.proquest.com/trade-journals/harnessing-power-ai-legal-considerations/docview/2856491812/se-2

[192] G. Gebrael, K. K. Sahu, B. Chigarira, N. Tripathi, V. Mathew Thomas, N. Sayegh, B. L. Maughan, N. Agarwal, U. Swami, and H. Li, "Enhancing triage efficiency and accuracy in emergency rooms for patients with metastatic prostate cancer: a retrospective analysis of artificial intelligence-assisted triage using chatgpt 4.0," *Cancers*, vol. 15, no. 14, p. 3717, 2023.

[193] A. Tiwari, A. Kumar, S. Jain, K. S. Dhull, A. Sajjanar, R. Puthenkandathil, K. Paiwal, R. Singh, and A. Sajjanar, "Implications of chatgpt in public health dentistry: A systematic review," *Cureus*, vol. 15, no. 6, 2023.

[194] M. Bhattacharyya, V. M. Miller, D. Bhattacharyya, L. E. Miller, and V. Miller, "High rates of fabricated and inaccurate references in chatgpt-generated medical content," *Cureus*, vol. 15, no. 5, 2023.

[195] J. Dossantos, J. An, and R. Javan, "Eyes on ai: Chatgpt's transformative potential impact on ophthalmology," *Cureus*, vol. 15, no. 6, 2023.

[196] E. Loos, J. Gröpler, and M.-L. S. Goudeau, "Using chatgpt in education: Human reflection on chatgpt's self-reflection," *Societies*, vol. 13, no. 8, p. 196, 2023.

[197] S. Koga, "The integration of large language models such as chatgpt in scientific writing: Harnessing potential and addressing pitfalls," *Korean Journal of Radiology*, vol. 24, no. 9, p. 924, 2023.

[198] M. Birenbaum, "The chatbots' challenge to education: Disruption or destruction?" *Education Sciences*, vol. 13, no. 7, p. 711, 2023.

[199] M. Cascella, J. Montomoli, V. Bellini, and E. Bignami, "Evaluating the feasibility of chatgpt in healthcare: an analysis of multiple clinical and research scenarios," *Journal of Medical Systems*, vol. 47, no. 1, p. 33, 2023.

[200] J. K. Aronson, "When i use a word... chatgpt: a differential diagnosis," 2023.

[201] M. Kumar, U. A. Mani, P. Tripathi, M. Saalim, S. Roy, and S. Roy Sr, "Artificial hallucinations by google bard: Think before you leap," *Cureus*, vol. 15, no. 8, 2023.

[202] J. M. L. Ferres, W. B. Weeks, L. C. Chu, S. P. Rowe, and E. K. Fishman, "Beyond chatting: The opportunities and challenges of chatgpt in medicine and radiology," *Diagnostic and Interventional Imaging*, pp. S2211–5684, 2023.

[203] J. Varghese and J. Chapiro, "Chatgpt: The transformative influence of generative ai on science and healthcare," *Journal of Hepatology*, 2023.

[204] P. Dunn and D. Cianflone, "Artificial intelligence in cardiology: Exciting but handle with caution," *International Journal of Cardiology*, 2023.

[205] A. A. Theodosiou and R. C. Read, "Artificial intelligence, machine learning and deep learning: Potential resources for the infection clinician," *Journal of Infection*, 2023.

[206] D. R. Fesenmaier and K. Wöber, "Ai, chatgpt and the university," *Annals of Tourism Research*, vol. 101, p. 103578, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0160738323000518

[207] S. S. Sohail, F. Farhat, Y. Himeur, M. Nadeem, D. Ø. Madsen, Y. Singh, S. Atalla, and W. Mansoor, "Decoding chatgpt: A taxonomy of existing research, current challenges, and possible future directions," *Journal of King Saud University-Computer and Information Sciences*, p. 101675, 2023.

[208] R. J. Lyons, S. R. Arepalli, O. Fromal, J. D. Choi, and N. Jain, "Artificial intelligence chatbot performance in triage of ophthalmic conditions," *Canadian Journal of Ophthalmology*, 2023.

[209] A. L. Opdahl, B. Tessem, D.-T. Dang-Nguyen, E. Motta, V. Setty, E. Throndsen, A. Tverberg, and C. Trattner, "Trustworthy journalism through ai," *Data & Knowledge Engineering*, vol. 146, p. 102182, 2023.

[210] F. Blanchard, M. Assefi, N. Gatulle, and J.-M. Constantin, "Chatgpt in the world of medical research: From how it works to how to use it." *Anaesthesia, Critical Care & Pain Medicine*, pp. 101 231–101 231, 2023.

[211] Z. W. Lim, K. Pushpanathan, S. M. E. Yew, Y. Lai, C.-H. Sun, J. S. H. Lam, D. Z. Chen, J. H. L. Goh, M. C. J. Tan, B. Sheng *et al.*, "Benchmarking large language models' performances for myopia care: a comparative analysis of chatgpt-3.5, chatgpt-4.0, and google bard," *Ebiomedicine*, vol. 95, 2023.

[212] T.-H. Kim, J. W. Kang, and M. S. Lee, "Ai chat bot-chatgpt-4: A new opportunity and challenges in complementary and alternative medicine (cam)," *Integrative Medicine Research*, vol. 12, no. 3, p. 100977, 2023.

[213] T. Alqahtani, H. A. Badreldin, M. Alrashed, A. I. Alshaya, S. S. Alghamdi, K. bin Saleh, S. A. Alowais, O. A. Alshaya, I. Rahman, M. S. Al Yami *et al.*, "The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research," *Research in Social and Administrative Pharmacy*, 2023.

[214] A. A. Jairoun, S. S. Al-Hemyari, M. Shahwan, G. R. H. Alnuaimi, H. Z. Sa'ed, and M. Jairoun, "Chatgpt: Threat or boon to the future of pharmacy practice?" pp. S1551–7411, 2023.

[215] J. Šlapeta, "Are chatgpt and other pretrained language models good parasitologists?" *Trends in Parasitology*, 2023.

[216] D. Dillion, N. Tandon, Y. Gu, and K. Gray, "Can ai language models replace human participants?" *Trends in Cognitive Sciences*, 2023.

[217] C. Thomson, E. Reiter, and B. Sundararajan, "Evaluating factual accuracy in complex data-to-text," *Computer Speech & Language*, vol. 80, p. 101482, 2023.

[218] M. Balas and E. B. Ing, "Conversational ai models for ophthalmic diagnosis: Comparison of chatgpt and the isabel pro differential diagnosis generator," *JFO Open Ophthalmology*, vol. 1, p. 100005, 2023.

[219] M. Salah, H. Al Halbusi, and F. Abdelfattah, "May the force of text data analysis be with you: Unleashing the power of generative ai for social psychology research," *Computers in Human Behavior: Artificial Humans*, p. 100006, 2023.

[220] J. Ilicki, "A framework for critically assessing chatgpt and other large language artificial intelligence model applications in health care," *Mayo Clinic Proceedings: Digital Health*, vol. 1, no. 2, pp. 185–188, 2023.

[221] B. J. Jansen, S.-g. Jung, and J. Salminen, "Employing large language models in survey research," *Natural Language Processing Journal*, vol. 4, p. 100020, 2023.

[222] J. E. Casal and M. Kessler, "Can linguists distinguish between chatgpt/ai and human writing?: A study of research ethics and academic publishing," *Research Methods in Applied Linguistics*, vol. 2, no. 3, p. 100068, 2023.

[223] X. Qi, Z. Zhu, and B. Wu, "The promise and peril of chatgpt in geriatric nursing education: What we know and do not know," p. 100136, 2023.

[224] Z. Lin, "Why and how to embrace ai such as chatgpt in your academic life," Feb 2023. [Online]. Available: psyarxiv.com/sdx3j

[225] B. V. Janssen, G. Kazemier, and M. G. Besselink, "The use of chatgpt and other large language models in surgical science," p. zrad032, 2023.

[226] Y. Shen, L. Heacock, J. Elias, K. D. Hentel, B. Reig, G. Shih, and L. Moy, "Chatgpt and other large language models are double-edged swords," p. e230163, 2023.

[227] A. J. Thirunavukarasu, R. Hassan, S. Mahmood, R. Sanghera, K. Barzangi, M. El Mukashfi, and S. Shah, "Trialling a large language model (chatgpt) in general practice with the applied knowledge test: observational study demonstrating opportunities and limitations in primary care," *JMIR Medical Education*, vol. 9, no. 1, p. e46599, 2023.

[228] J. Qadir, "Engineering education in the era of chatgpt: Promise and pitfalls of generative ai for education," in *2023 IEEE Global Engineering Education Conference (EDUCON)*. IEEE, 2023, pp. 1–9.

[229] S. Frieder, L. Pinchetti, R.-R. Griffiths, T. Salvatori, T. Lukasiewicz, P. C. Petersen, A. Chevalier, and J. Berner, "Mathematical capabilities of chatgpt," *arXiv preprint arXiv:2301.13867*, 2023.

[230] OpenAI, "Gpt-4 technical report," 2023.

[231] P. Manakul, A. Liusie, and M. J. Gales, "Selfcheckgpt: Zero-resource black-box hallucination detection for generative large language models," *arXiv preprint arXiv:2303.08896*, 2023.

[232] H. Nori, N. King, S. M. McKinney, D. Carignan, and E. Horvitz, "Capabilities of gpt-4 on medical challenge problems," *arXiv preprint arXiv:2303.13375*, 2023.

[233] J. Li, X. Cheng, W. X. Zhao, J.-Y. Nie, and J.-R. Wen, "Halueval: A large-scale hallucination evaluation benchmark for large language models," *arXiv e-prints*, pp. arXiv–2305, 2023.

[234] W. X. Zhao, K. Zhou, J. Li, T. Tang, X. Wang, Y. Hou, Y. Min, B. Zhang, J. Zhang, Z. Dong *et al.*, "A survey of large language models," *arXiv preprint arXiv:2303.18223*, 2023.

[235] V. Adlakha, P. BehnamGhader, X. H. Lu, N. Meade, and S. Reddy, "Evaluating correctness and faithfulness of instruction-following models for question answering," *arXiv preprint arXiv:2307.16877*, 2023.

[236] A. Athavale, J. Baier, E. Ross, and E. Fukaya, "The potential of chatbots in chronic venous disease patient management," *JVS-Vascular Insights*, p. 100019, 2023.

[237] T. C. Chen, E. Kaminski, L. Koduri, A. Singer, J. Singer, M. Couldwell, J. Delashaw, A. Dumont, and A. Wang, "Chat gpt as a neuro-score calculator: Analysis of a large language model's performance on various neurological exam grading scales," *World Neurosurgery*, 2023.

[238] B. C. Stahl and D. Eke, "The ethics of chatgpt–exploring the ethical issues of an emerging technology," *International Journal of Information Management*, vol. 74, p. 102700, 2024.

[239] W. H. Walters and E. I. Wilder, "Fabrication and errors in the bibliographic citations generated by chatgpt," *Scientific Reports*, vol. 13, no. 1, p. 14045, 2023.

[240] C. L. Munro and A. A. Hope, "Artificial intelligence in critical care practice and research," *American Journal of Critical Care*, vol. 32, no. 5, pp. 321–323, 2023.

[241] D. A. Hashimoto and K. B. Johnson, "The use of artificial intelligence tools to prepare medical school applications," *Academic Medicine*, vol. 98, no. 9, pp. 978–982, 2023.

[242] A. Jain, "Chatgpt for scientific community: Boon or bane?" *Medical journal, Armed Forces India*, vol. 79, no. 5, pp. 498–499, 2023.

[243] W. L. J. Ho, B. Koussayer, and J. Sujka, "Chatgpt: Friend or foe in medical writing? an example of how chatgpt can be utilized in writing case reports," *Surgery in Practice and Science*, p. 100185, 2023.

[244] G. M. Currie, "Academic integrity and artificial intelligence: is chatgpt hype, hero or heresy?" in *Seminars in Nuclear Medicine*. Elsevier, 2023.

[245] X. Liu, H. Lai, H. Yu, Y. Xu, A. Zeng, Z. Du, P. Zhang, Y. Dong, and J. Tang, "Webglm: Towards an efficient web-enhanced question answering system with human preferences," *arXiv preprint arXiv:2306.07906*, 2023.

[246] M. Campbell and M. Jovanović, "Detecting artificial intelligence: A new cyberarms race begins," *Computer*, vol. 56, no. 8, pp. 100–105, 2023.

[247] N. Kshetri, "The economics of generative artificial intelligence in the academic industry," *Computer*, vol. 56, no. 8, pp. 77–83, 2023.

[248] M. R. Ali, C. A. Lawson, A. M. Wood, and K. Khunti, "Addressing ethnic and global health inequalities in the era of artificial intelligence healthcare models: a call for responsible implementation," *Journal of the Royal Society of Medicine*, vol. 116, no. 8, pp. 260–262, 2023.

[249] J. Q. Tay, "Chatgpt and the future of plastic surgery research: evolutionary tool or revolutionary force in academic publishing?" *European Journal of Plastic Surgery*, pp. 1–2, 2023.

[250] Z. Xu and D. Cohen, "A lightweight constrained generation alternative for query-focused summarization," *arXiv preprint arXiv:2304.11721*, 2023.

[251] S. Kernan Freire, M. Foosherian, C. Wang, and E. Niforatos, "Harnessing large language models for cognitive assistants in factories," in *Proceedings of the 5th International Conference on Conversational User Interfaces*, 2023, pp. 1–6.

[252] M.-L. Tsai, C. W. Ong, and C.-L. Chen, "Exploring the use of large language models (llms) in chemical engineering education: Building core course problem models with chat-gpt," *Education for Chemical Engineers*, vol. 44, pp. 71–95, 2023.

[253] C. Xin, H. Lin, S. Wu, X. Han, B. Chen, W. Dai, S. Chen, B. Wang, and L. Sun, "Dialogue rewriting via skeleton-guided generation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 11, 2023, pp. 13 825–13 833.

[254] E. Murgia, Z. Abbasiantaeb, M. Aliannejadi, T. Huibers, M. Landoni, and M. S. Pera, "Chatgpt in the classroom: A preliminary exploration on the feasibility of adapting chatgpt to support children's information discovery," in *Adjunct Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization*, 2023, pp. 22–27.

[255] K. N. Kunze, S. J. Jang, M. A. Fullerton, J. M. Vigdorchik, and F. S. Haddad, "What's all the chatter about?: current applications and ethical considerations of artificial intelligence language models," *The Bone & Joint Journal*, vol. 105, no. 6, pp. 587–589, 2023.

[256] M. Cascella, J. Montomoli, V. Bellini, A. Ottaiano, M. Santorsola, F. Perri, F. Sabbatino, A. Vittori, and E. G. Bignami, "Writing the paper "unveiling artificial intelligence: an insight into ethics and applications in anesthesia" implementing the large language model chatgpt: a qualitative study," *Journal of Medical Artificial Intelligence*, vol. 6, 2023.

[257] A. Watanabe, "Exploring totalitarian elements of artificial intelligence in higher education with hannah arendt," *International Journal of Technoethics (IJT)*, vol. 14, no. 1, pp. 1–15, 2023.

[258] T. Nakaura, N. Yoshida, N. Kobayashi, K. Shiraishi, Y. Nagayama, H. Uetani, M. Kidoh, M. Hokamura, Y. Funama, and T. Hirai, "Preliminary assessment of automated radiology report generation with generative pre-trained transformers: comparing results to radiologist-generated reports," *Japanese Journal of Radiology*, pp. 1–11, 2023.

[259] S. Feuerriegel, J. Hartmann, C. Janiesch, and et al., "Generative ai," *Business & Information Systems Engineering*, 2023.

[260] B. N. Hryciw, Z. Fortin, J. Ghossein, and K. Kyeremanteng, "Doctor-patient interactions in the age of ai: navigating innovation and expertise," *Frontiers in Medicine*, vol. 10, 2023.

[261] E. Bran, C. Rughiniş, G. Nadoleanu, and M. G. Flaherty, "The emerging social status of generative ai: Vocabularies of ai competence in public discourse," in *2023 24th International Conference on Control Systems and Computer Science (CSCS)*. IEEE, 2023, pp. 391–398.

[262] F. Sovrano, K. Ashley, and A. Bacchelli, "Toward eliminating hallucinations: Gpt-based explanatory ai for intelligent textbooks and documentation," 2023.

[263] P. Mishra, M. Warr, and R. Islam, "Tpack in the age of chatgpt and generative ai," *Journal of Digital Learning in Teacher Education*, pp. 1–17, 2023.

[264] K.-Y. Lam, V. C. Cheng, and Z. K. Yeong, "Applying large language models for enhancing contract drafting," 2023.

[265] D. Wan, M. Liu, K. McKeown, M. Dreyer, and M. Bansal, "Faithfulness-aware decoding strategies for abstractive summarization," *arXiv preprint arXiv:2303.03278*, 2023.

[266] A. B. Houston and E. M. Corrado, "Embracing chatgpt: Implications of emergent language models for academia and libraries," *Technical Services Quarterly*, vol. 40, no. 2, pp. 76–91, 2023.

[267] C. González-Mora, C. Barros, I. Garrigós, J. Zubcoff, E. Lloret, and J.-N. Mazón, "Improving open data web api documentation through interactivity and natural language generation," *Computer Standards & Interfaces*, vol. 83, p. 103657, 2023.

[268] J. Lim, M. Kang, Y. Hur, S. Jung, J. Kim, Y. Jang, D. Lee, H. Ji, D. Shin, S. Kim *et al.*, "You truly understand what i need: Intellectual and friendly dialogue agents grounding knowledge and persona," *arXiv preprint arXiv:2301.02401*, 2023.

[269] S. A. Alowais, S. S. Alghamdi, N. Alsuhebany, T. Alqahtani, A. I. Alshaya, S. N. Almohareb, A. Aldairem, M. Alrashed, K. Bin Saleh, H. A. Badreldin *et al.*, "Revolutionizing healthcare: the role of artificial intelligence in clinical practice," *BMC Medical Education*, vol. 23, no. 1, pp. 1–15, 2023.

[270] S. Liu, A. P. Wright, B. L. Patterson, J. P. Wanderer, R. W. Turer, S. D. Nelson, A. B. McCoy, D. F. Sittig, and A. Wright, "Using ai-generated suggestions from chatgpt to optimize clinical decision support," *Journal of the American Medical Informatics Association*, vol. 30, no. 7, pp. 1237–1245, 2023.

[271] Q. Xie, E. J. Schenck, H. S. Yang, Y. Chen, Y. Peng, and F. Wang, "Faithful ai in medicine: A systematic review with large language models and beyond," *Medrxiv: the Preprint Server for Health Sciences*, 2023.

[272] I. A. Bernstein, Y. V. Zhang, D. Govil, I. Majid, R. T. Chang, Y. Sun, A. Shue, J. C. Chou, E. Schehlein, K. L. Christopher *et al.*, "Comparison of ophthalmologist and large language model chatbot responses to online patient eye care questions," *JAMA Network Open*, vol. 6, no. 8, pp. e2 330 320–e2 330 320, 2023.

[273] T. G. Heck, "What artificial intelligence knows about 70 kda heat shock proteins, and how we will face this chatgpt era," *Cell Stress and Chaperones*, vol. 28, no. 3, pp. 225–229, 2023.

[274] P. P. Ray and P. Majumder, "The potential of chatgpt to transform healthcare and address ethical challenges in artificial intelligence-driven medicine," *Journal of Clinical Neurology (Seoul, Korea)*, vol. 19, no. 5, p. 509, 2023.

[275] T. L. Ang, M. Choolani, K. C. See, and K. K. Poh, "The rise of artificial intelligence: addressing the impact of large language models such as chatgpt on scientific publications," *Singapore Medical Journal*, vol. 64, no. 4, p. 219, 2023.

[276] A. Talyshinskii, N. Naik, B. Z. Hameed, U. Zhanbyrbekuly, G. Khairli, B. Guliev, P. Juilebø-Jones, L. Tzelves, and B. K. Somani, "Expanding horizons and navigating challenges for enhanced clinical workflows: Chatgpt in urology," *Frontiers in Surgery*, vol. 10, 2023.

[277] J. E. Kung, C. Marshall, C. Gauthier, T. A. Gonzalez, and J. B. Jackson III, "Evaluating chatgpt performance on the orthopaedic in-training examination," *JBJS Open Access*, vol. 8, no. 3, p. e23, 2023.

[278] H. Friederichs, W. J. Friederichs, and M. März, "Chatgpt in medical school: how successful is ai in progress testing?" *Medical Education Online*, vol. 28, no. 1, p. 2220920, 2023.

[279] N. Ghorashi, A. Ismail, P. Ghosh, A. Sidawy, R. Javan, and N. S. Ghorashi, "Ai-powered chatbots in medical education: Potential applications and implications," *Cureus*, vol. 15, no. 8, 2023.

[280] J. C. McBee, D. Y. Han, L. Liu, L. Ma, D. A. Adjeroh, D. Xu, and G. Hu, "Interdisciplinary inquiry via panelgpt: Application to explore chatbot application in sports rehabilitation," *medRxiv*, pp. 2023–07, 2023.

[281] L. O. Reis, "Chatgpt for medical applications and urological science," *International braz j urol*, vol. 49, pp. 652–656, 2023.

[282] M. P. Joachimiak, J. H. Caufield, N. Harris, and C. J. Mungall, "Gene set summarization using large language models," *arXiv preprint arXiv:2305.13338*, 2023.

[283] A. R. Chen, "Research training in an ai world," 2023.

[284] M. Delsoz, Y. Madadi, W. M. Munir, B. Tamm, S. Mehravaran, M. Soleimani, A. Djalilian, and S. Yousefi, "Performance of chatgpt in diagnosis of corneal eye diseases," *medRxiv*, 2023.

[285] S. Takagi, T. Watari, A. Erabi, K. Sakaguchi *et al.*, "Performance of gpt-3.5 and gpt-4 on the japanese medical licensing examination: comparison study," *JMIR Medical Education*, vol. 9, no. 1, p. e48002, 2023.

[286] Y. Kaneda, R. Takahashi, U. Kaneda, S. Akashima, H. Okita, S. Misaki, A. Yamashiro, A. Ozaki, and T. Tanimoto, "Assessing the performance of gpt-3.5 and gpt-4 on the 2023 japanese nursing examination," *Cureus*, vol. 15, no. 8, 2023.

[287] S. Ahn, "The impending impacts of large language models on medical education," *Korean Journal of Medical Education*, vol. 35, no. 1, p. 103, 2023.

[288] Q. Jin, Z. Wang, C. S. Floudas, J. Sun, and Z. Lu, "Matching patients to clinical trials with large language models," *arXiv preprint arXiv:2307.15051*, 2023.

[289] S. Ahn, "A use case of chatgpt in a flipped medical terminology course," *Korean Journal of Medical Education*, vol. 35, no. 3, p. 303, 2023.

[290] J. Liu, C. Wang, and S. Liu, "Utility of chatgpt in clinical practice," *Journal of Medical Internet Research*, vol. 25, p. e48568, 2023.

[291] N. Doorn, "On the use of large language models in the water domain: Navigating the scylla of naïve techno-optimism and the charybdis of technology denial," *The Science of the Total Environment*, 2023.

[292] W. Hiesinger, C. Zakka, A. Chaurasia, R. Shad, A. Dalal, J. Kim, M. Moor, K. Alexander, E. Ashley, J. Boyd *et al.*, "Almanac: Retrieval-augmented language models for clinical medicine," 2023.

[293] D. Palal, S. Ghonge, V. Jadav, and H. Rathod, "Chatgpt: A double-edged sword?" *Health Services Insights*, vol. 16, p. 11786329231174338, 2023.

[294] B. N. Coskun, B. Yagiz, G. Ocakoglu, E. Dalkilic, and Y. Pehlivan, "Assessing the accuracy and completeness of artificial intelligence language models in providing information on methotrexate use," *Rheumatology International*, pp. 1–7, 2023.

[295] M. T. Baldassarre, D. Caivano, B. Fernandez Nieto, D. Gigante, and A. Ragone, "The social impact of generative ai: An analysis on chatgpt," in *Proceedings of the 2023 ACM Conference on Information Technology for Social Good*, 2023, pp. 363–373.

[296] N. Boujemaa, A. Hassan, G. Kokaia, and P. K. Sinha, "How responsible llms are beneficial to search and exploration in retail industry," in *Proceedings of the 2023 ACM International Conference on Multimedia Retrieval*, 2023, pp. 669–669.

[297] D. Li, T. Chen, A. Zadikian, A. Tung, and L. B. Chilton, "Improving automatic summarization for browsing longform spoken dialog," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023, pp. 1–20.

[298] M. Urban, D. D. Nguyen, and C. Binnig, "Omniscientdb: A large language model-augmented dbms that knows what other dbmss do not know," in *Proceedings of the Sixth International Workshop on Exploiting Artificial Intelligence Techniques for Data Management*, 2023, pp. 1–7.

[299] J. Mahon, B. Mac Namee, and B. A. Becker, "No more pencils no more books: Capabilities of generative ai on irish and uk computer science school leaving examinations," in *Proceedings of the 2023 Conference on United Kingdom & Ireland Computing Education Research*, ser. UKICER '23. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: https://doi.org/10.1145/3610969.3610982

[300] T. August, L. L. Wang, J. Bragg, M. A. Hearst, A. Head, and K. Lo, "Paper plain: Making medical research papers approachable to healthcare consumers with natural language processing," *ACM Transactions on Computer-Human Interaction*, 2022.

[301] J. E. Fischer, "Generative ai considered harmful," 2023.

[302] Y. Lee, T. S. Kim, S. Kim, Y. Yun, and J. Kim, "Dapie: Interactive step-by-step explanatory dialogues to answer children's why and how questions," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023, pp. 1–22.

[303] V. Scotti, L. Sbattella, and R. Tedesco, "A primer on seq2seq models for generative chatbots," *ACM Computing Surveys*, 2023.

[304] X. Zhan, Y. Xu, and S. Sarkadi, "Deceptive ai ecosystems: The case of chatgpt," *arXiv preprint arXiv:2306.13671*, 2023.

[305] A. R. Vargas-Murillo, I. N. M. d. l. A. Pari-Bedoya, and F. d. J. Guevara-Soto, "The ethics of ai assisted learning: A systematic literature review on the impacts of chatgpt usage in education," in *Proceedings of the 2023 8th International Conference on Distance Education and Learning*, 2023, pp. 8–13.

[306] R. White, T. Peng, P. Sripitak, A. Rosenberg Johansen, and M. Snyder, "Clinidigest: A case study in large language model based large-scale summarization of clinical trial descriptions," in *Proceedings of the 2023 ACM Conference on Information Technology for Social Good*, 2023, pp. 396–402.

[307] M. A. Cusumano, "Generative ai as a new innovation platform," *Communications of the ACM*, vol. 66, no. 10, pp. 18–21, 2023.

[308] N. Nashid, M. Sintaha, and A. Mesbah, "Retrieval-based prompt selection for code-related few-shot learning," in *Proceedings of the 45th International Conference on Software Engineering (ICSE'23)*, 2023.

[309] P. Pataranutaporn, V. Danry, L. Blanchard, L. Thakral, N. Ohsugi, P. Maes, and M. Sra, "Living memories: Ai-generated characters as digital mementos," in *Proceedings of the 28th International Conference on Intelligent User Interfaces*, 2023, pp. 889–901.

[310] G. Faggioli, L. Dietz, C. L. Clarke, G. Demartini, M. Hagen, C. Hauff, N. Kando, E. Kanoulas, M. Potthast, B. Stein *et al.*, "Perspectives on large language models for relevance judgment," in *Proceedings of the 2023 ACM SIGIR International Conference on Theory of Information Retrieval*, 2023, pp. 39–50.

[311] A. Marczak-Czajka and J. Cleland-Huang, "Using chatgpt to generate human-value user stories as inspirational triggers," in *2023 IEEE 31st International Requirements Engineering Conference Workshops (REW)*, 2023, pp. 52–61.

[312] A. Stefińska, T. P. Stefański, and M. Czubenko, "Evaluation of chatgpt applicability to learning quantum physics," in *2023 16th International Conference on Signal Processing and Communication System (ICSPCS)*, 2023, pp. 1–10.

[313] C. P. Ezenkwu, "Towards expert systems for improved customer services using chatgpt as an inference engine," in *2023 International Conference on Digital Applications, Transformation & Economy (ICDATE)*, 2023, pp. 1–5.

[314] U. M. Fayyad, "From stochastic parrots to intelligent assistants—the secrets of data and human interventions," *IEEE Intelligent Systems*, vol. 38, no. 3, pp. 63–67, 2023.

[315] J. Mrabet and R. Studholme, "Chatgpt: A friend or a foe?" in *2023 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)*. IEEE, 2023, pp. 269–274.

[316] J. Pitt, "Chatsh* t and other conversations (that we should be having, but mostly are not)," *IEEE Technology and Society Magazine*, vol. 42, no. 3, pp. 7–13, 2023.

[317] P. Crosthwaite and V. Baisa, "Generative ai and the end of corpus-assisted data-driven learning? not so fast!" *Applied Corpus Linguistics*, vol. 3, no. 3, p. 100066, 2023.

[318] P. W. Vinny, "Invoking ai for diagnosis: Art at the cutting edge of science," *Journal of the Neurological Sciences*, 2023.

[319] A. Solyman, M. Zappatore, W. Zhenyu, Z. Mahmoud, A. Alfatemi, A. O. Ibrahim, and L. A. Gabralla, "Optimizing the impact of data augmentation for low-resource grammatical error correction," *Journal of King Saud University-Computer and Information Sciences*, vol. 35, no. 6, p. 101572, 2023.

[320] A. Waqas, M. M. Bui, E. F. Glassy, I. El Naqa, P. Borkowski, A. A. Borkowski, and G. Rasool, "Revolutionizing digital pathology with the power of generative artificial intelligence and foundation models," *Laboratory Investigation*, p. 100255, 2023.

[321] L. D. Stephens, J. W. Jacobs, B. D. Adkins, and G. S. Booth, "Battle of the (chat) bots: comparing large language models to practice guidelines for transfusion-associated graft-versus-host disease prevention," *Transfusion Medicine Reviews*, p. 150753, 2023.

[322] J. Dien, "Generative artificial intelligence as a plagiarism problem," p. 108621, 2023.

[323] R. Abu-Farha, L. Fino, F. Y. Al-Ashwal, M. Zawiah, L. Gharaibeh, M. H. Mea'ad, and F. D. Elhajji, "Evaluation of community pharmacists' perceptions and willingness to integrate chatgpt into their pharmacy practice: A study from jordan," *Journal of the American Pharmacists Association*, 2023.

[324] C. Tippareddy, S. Jiang, K. Bera, and N. Ramaiya, "Radiology reading room for the future: Harnessing the power of large language models like chatgpt," *Current Problems in Diagnostic Radiology*, 2023.

[325] O. Oviedo-Trespalacios, A. E. Peden, T. Cole-Hunter, A. Costantini, M. Haghani, J. Rod, S. Kelly, H. Torkamaan, A. Tariq, J. D. A. Newton *et al.*, "The risks of using chatgpt to obtain common safety-related information and advice," *Safety Science*, vol. 167, p. 106244, 2023.

[326] A. Sarraju, D. Ouyang, and D. Itchhaporia, "The opportunities and challenges of large language models in cardiology," *JACC: Advances*, vol. 2, no. 7, p. 100438, 2023.

[327] B. Coskun, G. Ocakoglu, M. Yetemen, and O. Kaygisiz, "Author reply," *Urology*, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0090429523005721

[328] K. Lam, "Chatgpt for low-and middle-income countries: a greek gift?" *The Lancet Regional Health–Western Pacific*, vol. 41, 2023.

[329] J. Pantanowitz and L. Pantanowitz, "Implications of chatgpt for cytopathology and recommendations for updating jasc guidelines on the responsible use of artificial intelligence." *Journal of the American Society of Cytopathology*, pp. S2213–2945, 2023.

[330] S. Boussen, J.-B. Denis, P. Simeone, D. Lagier, N. Bruder, and L. Velly, "Chatgpt and the stochastic parrot: artificial intelligence in medical research," *British Journal of Anaesthesia*, vol. 131, no. 4, pp. e120–e121, 2023.

[331] I. Schlam, M. S. Menezes, C. Corti, A. Tan, I. Abuali, and S. Tolaney, "Artificial intelligence as an adjunct tool for breast oncologists–are we there yet?" *ESMO open*, vol. 8, no. 5, 2023.

[332] Y.-G. Lyu and F. Wu, "Toward a more general empowering artificial intelligence," *Engineering*, vol. 25, pp. 1–2, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2095809923002096

[333] M. Sparkes, "Are chatbots really able to think like people?" 2023.

[334] G. Bhatia and A. Kulkarni, "Chatgpt as co-author: Are researchers impressed or distressed?" *Asian Journal of Psychiatry*, vol. 84, p. 103564, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1876201823001193

[335] A. Chatelan, A. Clerc, and P.-A. Fonta, "Chatgpt and future artificial intelligence chatbots: What may be the influence on credentialed nutrition and dietetics practitioners?" *Journal of the Academy of Nutrition and Dietetics*, 2023.

[336] F. Lareyre, B. Nasr, A. Chaudhuri, G. Di Lorenzo, M. Carlier, and J. Raffort, "Comprehensive review of natural language processing (nlp) in vascular surgery," in *EJVES Vascular Forum*. Elsevier, 2023.

[337] A. Wilkins, "The robot doctor will see you soon," *New Scientist*, vol. 257, no. 3423, p. 28, 2023.

[338] P. Piazza, E. Checcucci, S. Puliatti, I. R. Belenchòn, A. Veccia, J. G. Rivas, M. Taratkin, K.-F. Kowalewski, S. Rodler, and G. E. Cacciamani, "The long but necessary journey towards optimization of the cause-effect relationship between input and output for accountable use of chatgpt for academic purposes," *European Urology Focus*, 2023.

[339] M. Scanlon, B. Nikkel, and Z. Geradts, "Digital forensic investigation in the age of chatgpt," *Forensic Science International: Digital Investigation*, vol. 44, 2023.

[340] Y. Kaneda, "Chatgpt in infectious diseases: A practical evaluation and future considerations," *New Microbes and New Infections*, vol. 54, 2023.

[341] T. F. Tan, A. J. Thirunavukarasu, J. P. Campbell, P. A. Keane, L. R. Pasquale, M. D. Abramoff, J. Kalpathy-Cramer, F. Lum, J. E. Kim, S. L. Baxter *et al.*, "Generative artificial intelligence through chatgpt and other large language models in ophthalmology: Clinical applications and challenges," *Ophthalmology Science*, p. 100394, 2023.

[342] Q. Ai, T. Bai, Z. Cao, Y. Chang, J. Chen, Z. Chen, Z. Cheng, S. Dong, Z. Dou, F. Feng *et al.*, "Information retrieval meets large language models: A strategic report from chinese ir community," *AI Open*, vol. 4, pp. 80–90, 2023.

[343] J. F. Ruma, T. T. Mayeesha, and R. M. Rahman, "Transformer based answer-aware bengali question generation," *International Journal of Cognitive Computing in Engineering*, 2023.

[344] R. Mao, K. He, X. Zhang, G. Chen, J. Ni, Z. Yang, and E. Cambria, "A survey on semantic processing techniques," *Information Fusion*, p. 101988, 2023.

[345] B. P. Chaiken, "Unleashing precision medicine to deliver personalized care," 2023.

[346] B. Otaki, "Feedback in the era of generative ai," 2023.

[347] I. Triguero, D. Molina, J. Poyatos, J. Del Ser, and F. Herrera, "General purpose artificial intelligence systems (gpais): Properties, definition, taxonomy, open challenges and implications," *arXiv preprint arXiv:2307.14283*, 2023.

[348] H. Huang, "Performance of chatgpt on registered nurse license exam in taiwan: A descriptive study," 2023.

[349] K. Muranga, I. S. Muse, E. N. Köroğlu, and Y. Yildirim, "Artificial intelligence and underfunded education," *London Journal of Social Sciences*, no. 6, pp. 56–68, 2023.

[350] G. Polverini and B. Gregorcic, "How understanding large language models can inform their use in physics education," *arXiv preprint arXiv:2309.12074*, 2023.

[351] B. Rajendran, O. Simeone, and B. M. Al-Hashimi, "Towards efficient and trustworthy ai through hardware-algorithm-communication co-design," 2023.

[352] S. Lobentanzer and J. Saez-Rodriguez, "A platform for the biomedical application of large language models," 2023.