



Information Research – Vol. 31 No. 2 (2026)

# The credibility of information generated by AI chatbots: an analysis of online discussions in Reddit

Reijo Savolainen

DOI: <https://doi.org/10.47989/ir31262998>

## Abstract

**Introduction.** The present investigation elaborates how participants of online discussion in Reddit assess the credibility of information generated by AI chatbots.

**Method.** The empirical findings draw on the descriptive quantitative analysis and qualitative content analysis of 1464 posts submitted by 919 individual participants to 30 Reddit discussion threads during the period of October 2023 – September 2025.

**Analysis.** The analysis focuses on the criteria by which the online participants assessed the credibility of information generated by AI chatbots. Second, it was examined how the credibility assessments were grounded by making use of various answer types.

**Results.** The most frequently used credibility criteria were correctness, trustworthiness and usefulness of information, as well as chatbot's capability to generate relevant information. Accuracy, coverage, currency and verifiability of information were employed less frequently. Overall, the findings indicate that online participants adopted a critical stance on the credibility of information generated by AI chatbots. In particular, the occurrence of hallucinations undermined the beliefs that chatbots offer correct and trustworthy information. The credibility assessments were mainly grounded by drawing on personal opinions and experiences obtained from the use of chatbots. To a lesser extent, the assessments were also grounded by drawing on comparison and explanation.

**Conclusion.** Although the credibility of information generated chatbots is evaluated critically, people are not expecting that AI tools would offer fully credible answers.

## Introduction

The emergence of generative Artificial Intelligence (AI) has made Large Language Model (LLM) powered chatbots such as ChatGPT, Claude, Gemini, and Perplexity a viable alternative for information seeking. AI chatbots do not simply retrieve existing content (like that which could be done with Google, for example) but also generate new content on a topic (Schuetzler et al., 2024, p. 7504). With its remarkable natural language understanding and generation capabilities, AI chatbots complement search engines by organising answers in a natural and easily digestible format (Liu et al., 2024, pp. 103-104). The extent to which people trust in AI-generated content is dependent on how they evaluate its perceived accuracy, bias, personalisation and content explainability (e.g. Zhong and Fang, 2026). A recent study conducted by Yun and Bickmore (2025) found that while search engines and institutional health websites continue to be the dominant sources of online health information, 21% of the respondents had used LLM-based chatbots for medical purposes. In a national survey conducted in 2023, Mendel et al. (2025) found that 32% of U.S. adults reported using AI chatbots for health-related queries, including interpreting symptoms, understanding treatment options and reviewing test results.

On the other hand, the use of generative AI tools is not without problems because chatbots may expose users to undesirable consequences such as *hallucinations*. They occur when chatbots generate responses that appear plausible and factually grounded but are in fact misleading, inaccurate or entirely fabricated (Lim et al., 2025). Hallucinations occur because the contemporary generative AI models are probabilistic in nature. They make a prediction about what a meaningful answer might look like without applying any forms of logical or contextual reasoning (Grimes et al., 2023, pp. 1619-1620). Different from traditional search engines that present a variety of information resources for users to choose from, AI chatbots offer a single response based on what LLMs deem most relevant to the user's query (Ray, 2023). This means that the success or failure of information seeking hinges on the AI model's ability to understand and accurately gather information, thereby limiting user autonomy in evaluating diverse information sources (Lund et al., 2025). AI chatbots can perform well in expert evaluations for accuracy, clarity and clinical usefulness but the responses often lack proper source citations and sometimes include outdated information (Sezgin et al., 2025). Although LLM-powered chatbots can access a vast repository of information on the open internet, their capacity to generate accurate and contextually appropriate responses may be limited (Kim et al., 2025).

Given the above limitations of current LLMs, the credibility of information generated by AI chatbots has become a significant issue in diverse fields such as information research (Choi, Bak and An, 2025), medicine (Wang et al., 2025) and health communication (Lim and Hong, 2025). So far, investigations on this topic have resulted in mixed findings. On the one hand, as discussed in more detail in the literature review below, there are observations suggesting that AI chatbots can offer trustworthy and accurate responses. On the other hand, there is opposite evidence about incorrect and misleading information generated by generative AI tools. The present study contributes to information behaviour research by elaborating the picture of the pros and cons of AI-generated information by exploring how a relatively large number of participants in online discussion assess the credibility of responses offered by AI chatbots. To this end, an empirical study was made by analysing 1464 posts on the above topic, submitted by 919 individual participants to 30 discussion threads in Reddit - a major social media forum. The findings elaborate the picture of credibility assessment in two ways. First, the results deepen our understanding about the significance of diverse criteria used in the credibility assessments focusing on chatbots as information sources. Second, the findings specify the ways in which credibility assessments are grounded in online discussion.

## Background

### Observations about the credibility of information generated by AI chatbots

*Credibility* is a multidimensional concept that lacks a generally accepted definition. Researchers have approached credibility in diverse terms such as believability, trustworthiness, reliability, accuracy and objectivity (Hilligoss and Rieh, 2008; Metzger et al., 2003). Most researchers agree, however, that the key dimensions of credibility are trustworthiness and expertise. Information is trustworthy when it appears to be reliable and unbiased. Expertise indicates an individual's ability to provide information that is both accurate and valid (Hilligoss and Rieh, 2008, p. 1469). Researchers have also distinguished between message credibility and source credibility. *Message credibility* indicates how message characteristics impact perceptions of believability (Metzger et al., 2003). Of such characteristics, message content indicates the extent to which information available in the message is correct and accurate. *Source credibility* is indicative of the believability of the author of the message in terms of his or her perceived reputation, expertise and honesty. For the present study, message credibility is more significant because the information content generated by AI chatbots is the crux issue. Source credibility is less important because information generated by chatbots cannot primarily be seen as a contribution of individual authors whose trustworthiness depends on their perceived expertise, honesty and reputation.

The evaluation of the credibility of information generated by chatbots is rendered more difficult due to the opaque nature of AI algorithms and the absence of traditional cues signifying information trustworthiness. For example, the occurrence of hallucinations can complicate the evaluation process (Schuetzler et al., 2024). Nevertheless, there is a growing body of empirical studies examining the credibility of information generated by chatbots in diverse domains. Many of these investigations approach hallucinations as an indicator of the provision of incorrect information. Shen et al. (2023) identified several cases in which ChatGPT generated misleading information. For example, when asked 'Who composed the tune of Twinkle, Twinkle, Little Star?' where the composer is still a mystery in history, ChatGPT incorrectly responded with Wolfgang Amadeus Mozart as the composer.

The credibility of information generated by AI chatbots has also been examined in interview studies. Zhang et al. (2025) explored how parents and their children use ChatGPT. The findings indicate that more than half of the interviewed parents considered search engines such as Google to be more accurate than ChatGPT. Moreover, the study participants had concerns with ChatGPT's lack of transparency over how it sourced information and the potential for ChatGPT to generate misinformation. Choi, Bak and An (2025) examined college students' credibility assessments of AI-generated information for academic tasks. The findings revealed that the students conceptualised credibility using correctness-related terms, such as accurate and factual, especially in the context of erroneous outputs from AI tools, such as fabricated references for academic writing. The students also mentioned *refinedness* or related terms, such as *concise* and *sophisticated* as criteria. They expected chatbots to integrate information from relevant sources and process it to answer their questions directly. The study participants also considered currency as criterion of credibility. Finally, although many participants considered unbiasedness an important construct reflecting credibility, their views on the potential biases in AI-generated information varied.

The findings of many empirical investigations suggest that AI chatbots perform best when the prompts formulated by users focus on a specific issue. For example, ChatGPT demonstrated a 97 % accuracy in answering questions from the 'Common Cancer Myths and Misconceptions' page (Johnson et al., 2023). More recently, Hanss et al. (2025) assessed the accuracy and reliability of GPT-3.5, GPT-4 and GPT-5o in psychiatry using standardised multiple-choice questions. The results indicate the accuracy displayed by GPT-4 and GPT-5o (84% and 87%, respectively) is comparable to previously reported GPT-4 accuracy rates in other specialties, such as

ophthalmology (82%) and neurosurgery (83%). On the other hand, all studies on the accuracy and reliability of AI chatbots have not produced favourable results. Kim et al. (2025) tested the capability of ChatGPT for parent and caregiver information seeking. Despite being generally correct, the AI-generated responses frequently missed context-specific details, particularly regarding child developmental stages and emotional nuances. Furthermore, the absence of reliable citations and occasional inaccuracies in references were notable drawbacks. Moreover, the coverage of information was limited because ChatGPT did not provide an exhaustive list of the documents, databases, or websites included in the training process. This means that it was not possible to verify or trace the provenance of the information presented by ChatGPT.

Compared to fields of medicine and health communication, there is a paucity of studies critically examining the credibility of information generated by AI chatbots in other domains. A rare example is offered by Murashko (2025), examining how users evaluate ChatGPT 3.5's responses on the Russian-Ukrainian war. The study was limited in that ChatGPT's responses were restricted by its data cut-off in January 2022, that is, one month before Russia's invasion to Ukraine. Consequently, this limitation led to incomplete or outdated information. The study participants expressed scepticism toward ChatGPT's credibility. Moreover, ChatGPT's inability to provide links or concrete references significantly undermined participants' evaluation of its credibility.

Many of the studies assessing the credibility of information generated by AI chatbots report mixed results. Behesti et al. (2025) reviewed 128 studies examining the reliability of ChatGPT for health-related questions. The findings indicate that in the above investigations, the overall mean accuracy of ChatGPT was 73%. On the other hand, the results revealed a wide variation in performance depending on the medical specialty. Psychiatry and mental health, along with dermatology and skin conditions, emerged as the areas where ChatGPT exhibits the highest accuracy, with relatively consistent results. In contrast, specialties such as cardiovascular diseases and oncology presented more variability in accuracy. Similarly, Barbosa-Silva et al. (2024) found that ChatGPT's accuracy and completeness vary in addressing urinary incontinence queries. Most of the answers were classified as adequate, as they provided the minimum information expected to be classified as correct. Barbosa-Silva et al. (2024) concluded that ChatGPT's ability to generate responses is restricted by its dependence on a diverse range of sources, including both high-quality peer-reviewed scientific studies and less reliable websites or blogs. On the other hand, the credibility of information on the same topic can vary between diverse chatbots. Sidhu and Selvamogan (2025) assessed the quality of AI chatbot responses to frequently asked questions about basal cell carcinoma. The overall quality of responses varied significantly across platforms. Gemini outperformed its competitors (ChatGPT, DeepSeek and Grok). The superior performance of Gemini was attributed to the consistent inclusion of references within its responses. Although the presence of citations may enhance the perceived credibility of AI-generated content, it does not inherently ensure overall quality of information. Therefore, factors such as factual accuracy and contextual relevance remain critical in evaluating the credibility of information generated by AI chatbots.

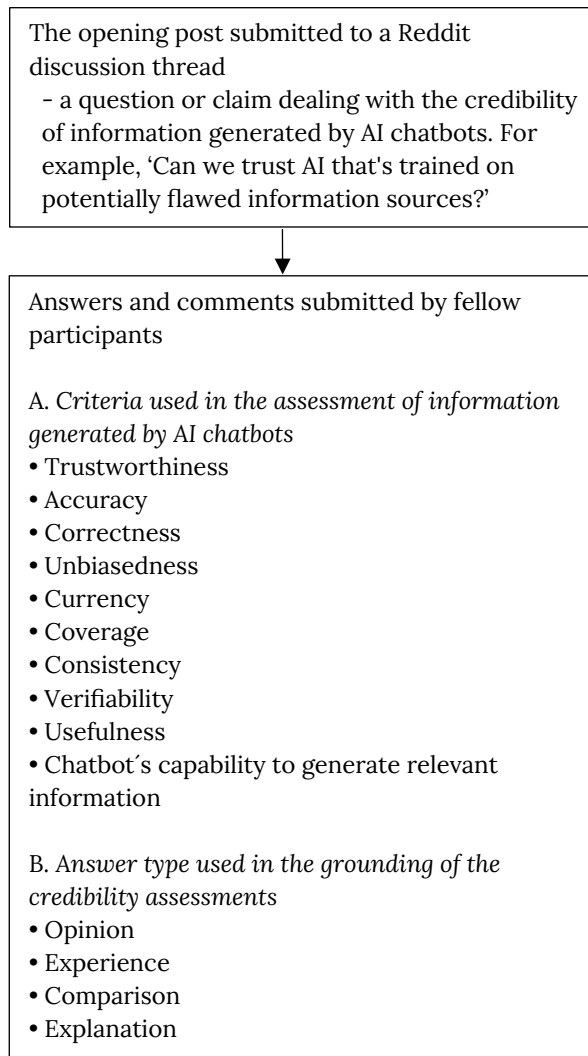
## Research framework

The literature review identified a number of criteria such as accuracy, reliability and currency by which people evaluate the credibility of responses offered by AI chatbots. Overall, the findings of prior investigations suggest that chatbots can generate relatively accurate information in certain domains of medicine. In opposite cases, information can be incorrect and misleading, particularly due to hallucinations generated by LLM powered tools. The present investigation aims at deepening our understanding about how people assess the credibility of information generated by chatbots in online discussion. To achieve this, the study was focused on discussions occurring in Reddit discussion threads. They were chosen for the study for two reasons. First, as Reddit is a global social media forum, its discussion threads attract a wide variety of participants' views on

the research topic. Second, as a publicly open platform, Reddit offers free access to discussion threads containing rich empirical data.

To specify the research framework further, it is assumed that online participants employ diverse criteria while making credibility assessments. Moreover, it is assumed that the participants ground their assessments by making use of various answer strategies manifesting themselves in diverse answer types. In the present study, *credibility* is approached as a subordinate category generally denoting the quality of being trusted and believed in. The nature of credibility can be specified by identifying a set of subcategories indicative of criteria by which the credibility of information generated by chatbots is assessed. To this end, the findings of prior studies reviewed above were used. The criteria thus identified include the following: *accuracy of information*, *currency of information*, *unbiasedness of information* (Choi, Bak and An, 2025), *correctness of information* (Zhang et al., 2025), and *coverage of information* (Kim et al., 2025). The list of criteria was substantiated by inductively identifying categories from the preliminary analysis of the empirical data used in the present investigation. These criteria include *trustworthiness of information*, *consistency of information*, *verifiability of information*, *usefulness of information* and *chatbot's capability to generate relevant information*. The ways in which the above 10 criteria were defined for empirical analysis will be specified below in the methodology section.

Second, in the identification of relevant answer types indicative of answer strategies, the findings of Savolainen (2025) and Westbrook (2015) were used. Westbrook (2015) specified the category of *explanation answer*, while Savolainen (2025) identified *opinion answers*. The list of relevant answer types was substantiated by inductively identifying two categories from the empirical data of the present investigation, that is, *experience answers* and *comparison answers*. The empirical definitions of the above four answer types will be specified in the methodology section. The research framework is presented in Figure 1.



**Figure 1.** The research framework

Figure 1 suggests that online participants assess the credibility of information generated by AI chatbots by answering the opening question or claim submitted by the initiator of the discussion thread. While assessing the credibility of information, multiple criteria can be used, ranging from trustworthiness and accuracy to currency and usefulness of information. While grounding their assessment, the participants may use diverse answer strategies, by drawing on their opinions about the trustworthiness of such information. Alternatively, the participants may describe their experiences about the usefulness of answers offered by chatbots. The contributors can also compare the credibility of information obtained from chatbots and other sources, as well as explain why chatbots tend to offer incorrect information.

## Research questions

Drawing on the above framework, the present study seeks answers to the following research questions.

- RQ1. Drawing on diverse criteria, how do the contributors to Reddit discussion threads assess the credibility of information generated by AI chatbots?
- RQ2. What types of answers do the Reddit contributors use while grounding their credibility assessments?

## Empirical data and analysis

The empirical data were gathered from Reddit – a major social media platform headquartered in San Francisco. Reddit consists of individual forums called *subreddits*, which are mostly user-created and organised by topic of discussion. Subreddits are managed by volunteer moderators who enforce rules for posting. Reddit users (*Redditors*) can join a forum to participate in posting and commenting on others' posts (Eldridge, 2025). As of May 2025, Reddit had an estimated 91 million daily active users all over the world (Kumar, 2025).

The empirical data were downloaded in the beginning of October 2025 from nine subreddits focusing on the discussion about the trust in AI chatbots and credibility of information generated by them. The subreddits of this kind included the following:

- r/ChatGPT
- r/OpenAI
- r/artificial
- r/ArtificialIntelligence
- r/futorology
- r/science
- r/psychology
- r/technology
- r/changemyview

These subreddits were chosen because they offer the most relevant repertoire of discussion threads for the topic of the present investigation. To obtain an overall picture of the posts submitted to the above subreddits, relevant discussion threads were tentatively identified using the search terms *AI chatbot*, *trust in chatbot*, and *credibility of information*. At the beginning of October 2025, the search resulted in the identification of 132 potentially relevant discussion threads from the above subreddits. The preliminary reading indicated that many of the discussion threads appeared to be of low informational value because they contained only the opening post or a small number (1-9) of posts submitted by fellow contributors. In addition, there were discussion threads focusing on a narrow topic, for example, trust in chatbots in digital marketing. After having excluded 92 irrelevant threads of this type, the remaining sample of 30 discussion threads was chosen for closer analysis. Examples of the opening posts indicative of the topics of the discussion threads include 'Can we trust AI that's trained on potentially flawed information sources?', 'How can we trust that any specific thing an AI says is accurate?' and 'How do you all trust ChatGPT'?

Altogether 919 individual participants submitted 1464 posts to the 30 threads during the period of October 2023 – September 2025. Of the 1464 posts, 30 opened the discussion, while 1434 posts were answers and comments to the opening question or claim presented by the initiator of the thread. The number of posts per discussion thread varied from 10 to 368. On average, a thread contained 48 posts. Of the contributors, the most active wrote no less than 64 posts, while the majority of the participants, that is, 66% submitted 1-2 posts. Although the discussion topics attracted a relatively high number of participants, most of them were occasional contributors who did not delve deeper into the debate about the credibility of information generated by AI chatbots. Importantly, however, the sample of 30 discussion threads appeared to be sufficient for the needs of qualitative analysis because the data became saturated. Data saturation occurs at the point at which collecting any further data will not produce value-added insights about the matter under investigation (Saunders et al., 2018, p. 1895). Therefore, it became evident that the analysis of additional discussion threads initiated after the beginning of October 2025 would not essentially change the qualitative picture of how the Redditors assess the credibility of information generated by AI chatbots. The sample of 30 discussion threads was downloaded in a separate file and coded by making use of the categories specified in Table 1.

| Category                                              | Definition                                                                                            | Illustrative example taken from the empirical data                                                                                          |
|-------------------------------------------------------|-------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|
| <i>A. Criteria used in the credibility assessment</i> |                                                                                                       |                                                                                                                                             |
| Trustworthiness                                       | The extent to which information provides reliable description of reality                              | ‘I trust Gemini when it has access to current information in RAG or web searches’. (Thread 10)                                              |
| Accuracy                                              | The extent to which information provides an exact description of reality                              | ‘I have found ChatGPT4o to be pretty accurate’. (Thread 19)                                                                                 |
| Correctness                                           | The extent to which information provides true description of reality, free of error                   | ‘I have seen OpenEvidence incorrectly give confident answers that are not supported by the cited literature’. (Thread 2).                   |
| Unbiasedness                                          | The extent to which information provides an impartial description of reality                          | ‘Chatbots can indeed be biased. Specifically, they reflect the biases of their developers’. (Thread 18)                                     |
| Currency                                              | The extent to which information provides an up-to-date description of reality                         | ‘Due to absence of the most recent studies, I have stopped using OpenEvidence altogether’. (Thread 2)                                       |
| Coverage                                              | The extent to which information provides a comprehensive description of reality                       | ‘AI chatbots are very bad at saying, "I don't know," or openly admitting where there are gaps in their results’. (Thread 18)                |
| Consistency                                           | The extent to which information provides a compatible description of reality, free from contradiction | ‘In the instances where I have read the summary and wanted to know more, the summary has been consistent with the full article’. (Thread 7) |
| Verifiability                                         | The extent to which the veracity of information can be confirmed                                      | ‘If it (chatbot) is not giving you linked sources, then the answer is not verifiable’. (Thread 15)                                          |
| Usefulness                                            | The extent to which information is considered as helpful to meet the need of a person                 | ‘I use OpenEvidence to help me find a source to help me make a decision or learn something and it is very useful for this’. (Thread 2)      |
| Chatbot’s capability to generate relevant information | The extent to which chatbot is considered to be capable of providing appropriate information          | ‘AI also has the capability of recognizing patterns, which can help find the correct sources of information’. (Thread 9)                    |
| <i>B. Answer type</i>                                 |                                                                                                       |                                                                                                                                             |
| Opinion                                               | A personal view on an object, issue or event                                                          | ‘LLMs are currently designed to prioritize satisfying answers rather than correct answers.’ (Thread 27)                                     |
| Experience                                            | A subjective account of an event or observation in a particular situation                             | ‘I have not yet seen OpenEvidence cite a non-existent study. But I have seen the summary contain blatantly wrong information’. (Thread 2)   |
| Comparison                                            | A consideration or estimate of the similarities or dissimilarities between two or more things         | ‘I trust it (chatbot) far more than any random Redditor’. (Thread 29)                                                                       |
| Explanation                                           | A statement offering                                                                                  | ‘The simple explanation is the models try to predict                                                                                        |

|                                                     |                                                                               |
|-----------------------------------------------------|-------------------------------------------------------------------------------|
| information about why and how certain things happen | the best answer. When they don't know, they guess (hallucinate)'. (Thread 16) |
|-----------------------------------------------------|-------------------------------------------------------------------------------|

**Table 1.** The coding categories

The coding was an iterative process in which the data were scrutinised several times by the author. The pre-defined categories specified in Table 1 were used to code all the data - while still allowing new codes to emerge. However, the content of all posts submitted by the Redditors fit into the existing categories defined in Table 1 and no new categories were needed to cover the data. The 1464 posts were assigned with 2111 codes. Of them, 935 dealt with the criteria used in the credibility assessment, while 1176 identified the answer type. In the coding, a post was coded only once for a criterion category, once it was identified for the first time in the post. In long posts in particular, it was not unusual that the same criterion was identified in several segments of the same post. In these cases, a post was coded for a criterion category, for example, accuracy (credibility criterion) and opinion (answer type) when the category appeared for the first time in the text; other instances of the same category were simply ignored. However, a post could be assigned with several criteria indicative of credibility assessment and answer type, for example, accuracy and usefulness of information, as well as opinion and explanation. To strengthen the reliability of the coding, the initial coding was refined by repeated reading of the data. Miles and Huberman (1994, p. 65) recommended that check-coding the same data is useful for the lone researcher and that code-recode consistencies should be at least 90%. To achieve this, several iterations were executed to ensure that the codes appropriately describe the data and that there are no anomalies.

In order to examine the relative share of the categories specified in Table 1, percentage distribution was calculated for individual items indicative of the criteria used in the credibility assessment and diverse answer types. To this end, the number of codes assigned to a category, for example, correctness of information ( $n = 305$ ) was divided by the total number of the codes dealing with credibility criteria, 935. Second and more importantly, the data were scrutinised by means of qualitative content analysis. To achieve this, the constant comparative method was used to capture the variety of the participants' articulations (Lincoln and Guba, 1985, pp. 339-344). More specifically, the qualitative analysis concentrated on the participants' articulations dealing with the assessment of the credibility of information generated by chatbots. The participants' responses concerning these assessments were systematically compared per individual instances of credibility assessments of diverse kind. In this way, it was possible to identify similarities and differences in the ways in which the Redditors, for example, assessed the trustworthiness of information generated by chatbots and how they grounded their assessments by depicting their opinions.

Although the posts submitted to Reddit are freely accessible to all readers and can be used as research material, the anonymity of the Redditors was protected while presenting illustrative extracts from their posts. This was done by deleting all information identifying the usernames of individual contributors, as well as the dates for their posts. Instead, individual posts were identified using technical codes. For example, in R628-T4, R628 refers to the Redditor who appears in the 628<sup>th</sup> place in the alphabetical list of the 919 participants, while T4 identifies discussion thread 4 where his or her post appeared.

## Findings

To provide background for the qualitative findings, the quantitative overview of the criteria and answer types used in the credibility assessments is presented first. Table 2 specifies how the online participants preferred diverse criteria while judging the credibility of information generated by AI chatbots.

| Credibility criterion                       | %     |
|---------------------------------------------|-------|
| Correctness                                 | 32.6  |
| Trustworthiness                             | 26.5  |
| Capability to generate relevant information | 11.7  |
| Usefulness                                  | 10.6  |
| Verifiability                               | 7.6   |
| Accuracy                                    | 3.9   |
| Consistency                                 | 2.2   |
| Coverage                                    | 2.0   |
| Unbiasedness                                | 2.0   |
| Currency                                    | 0.9   |
| In total                                    | 100.0 |

**Table 2.** The percentage distribution of codes assigned to the criteria of credibility assessments (n = 935)

Table 2 indicates that the participants most frequently devoted attention to the correctness of information while assessing the credibility of answers offered by chatbots. This emphasis is mainly due to the frequently expressed doubts that hallucinations generated by chatbots can result in the provision of incorrect information. Trustworthiness of information was another criterion frequently used by the Redditors. In their assessments, they also often drew on the chatbot's capability to generate relevant information. In addition, the usefulness of information was a significant criterion for many contributors. Other criteria such as verifiability, accuracy, consistency, coverage and currency of information were used less frequently.

A more detailed analysis of the frequencies presented in Table 2 revealed that the majority of the 935 codes assigned to the credibility assessments, that is, 72.8% reflected a critical approach to information generated by AI chatbots, while only 27.2% of the codes were positive in this regard. For example, of the assessments dealing with the correctness of information, no less than 92.5% were negative, that is, indicative of the judgment that information generated by chatbots tend to be incorrect rather than correct. Similarly, it was believed that the information generated by chatbots tends to be untrustworthy (71.1%) rather than trustworthy (28.9% of the codes assigned to this criterion). The only exceptions were the criteria of capability to generate relevant information and usefulness of information. For example, of the codes assigned to the latter criterion, 77.8% were positive suggesting that chatbots offer useful information, despite the fact that it may be not always be fully correct.

The ways in which the participants grounded their credibility assessments by drawing on various types of answers are specified in Table 3.

| Answer type | %     |
|-------------|-------|
| Opinion     | 56.3  |
| Experience  | 30.4  |
| Comparison  | 8.2   |
| Explanation | 5.1   |
| In total    | 100.0 |

**Table 3.** The percentage distribution of codes assigned to answer types (n = 1176)

Table 3 demonstrates that while grounding the credibility assessments, the Redditors mainly depicted their personal opinions about the credibility of chatbots as sources of information. The assessments were also often based on the description of concrete experiences obtained from the chatbot use. In contrast, analytical approaches drawing on comparison and explanation were employed quite seldom. All in all, due to the popularity of opinion and experience answers, the

discussions were characterised by the dominance of a descriptive approach to the topic. This means that the online discussions were broad but not particularly deep in the analytical sense.

## Qualitative features of the credibility assessments

### Correctness of information

The quantitative overview indicated that the Redditors most frequently devoted attention to the correctness of information while assessing the credibility of information generated by AI chatbots. Reflecting the critical approach to credibility assessments more generally, the answers suggesting that chatbots offer correct information were relatively rare.

*I have asked medical advice regarding exercise injuries and fact checked everything to find that GPT was quite correct at identifying the issue and coming up with a solution. (R718-T29)*

Critical assessments were more common, mainly due to the doubts that chatbots can generate hallucinated data. It was claimed, for example, that ‘ChatGPT is known to fabricate references and information’ (R24-T4). The participants also expressed their concerns about the provision of incorrect information by describing negative use experiences. As sources of information used in the training of chatbots, social media forums in particular were put in a dubious light.

*I looked up something today and the answer was completely wrong. The only source was a Reddit post where someone claimed it as fact. (R89-T3)*

As hallucinations tend to be quite common, chatbot users may face difficulties in trying to find out the extent to which the answers offered by chatbots are correct. This is due to how chatbots often present such answers with full confidence. Reflecting this issue, one of the contributors (R681-T5) claimed that ‘AI is not only confidently wrong. It is also convincingly wrong’. The incorrectness of information offered by chatbots was also made understandable by comparing them to other information sources which are traditionally deemed more credible.

*Google's "summarization" AI conflates the wrong things together or just makes shit up all the time, in a way that few human-written answers in places like Wikipedia or technical forums do. (R855-T5)*

The participants also offered explanatory answers to find reasons for incorrect information generated by chatbots. Many of the explanations were based on the characterisation of the ways in which LLM powered chatbots are designed.

*The simple explanation is the models try to predict the best answer. When they don't know, they guess (hallucinate). It is because if they guess, they still have a chance at being right, but if they don't answer, they have a 0% chance of being right. (R730-T16)*

The tendency to offer incorrect information may also be explained by the nature of the topic, as well as the degree of specificity of the prompts submitted to the chatbot.

*The further you venture to the edges, the more it (chatbot) will hallucinate. This is virtually an unsolvable problem and is likely why it seems like each new model regresses. (R455-T29)*

### **Trustworthiness of information**

Similar to correctness, trustworthiness of information was often seen as a key criterion while assessing the credibility of answers offered by AI chatbots. Again, positive evaluations were quite rare, thus reflecting the Redditors' suspicious attitude toward the chatbots. However, there were a few examples of positive opinions and experiences suggesting that AI chatbots represent trustworthy sources of information.

*I trust Gemini when it has access to current information in RAG (Retrieval-Augmented Generation) or web searches. (R92-T10)*

*I exclusively use Perplexity and have a high degree of trust in it. (R467-T29)*

Nevertheless, trust in AI chatbots was conditioned by certain factors. Chatbots may offer trustworthy information, given their limitations. Trust in chatbots was also mirrored against the trustworthiness of alternative information sources. In this regard, chatbots were preferred over material available in the social media forums in particular, as well as information offered by ordinary people.

*I can say that I trust ChatGPT, but only on certain things. For example, I trust it with basic questions, but not for serious health concerns. (R398-T29)*

*I think LLMs outperform random dude in every aspect. The thing is random dude is not necessarily great source of truth. (R779-T29)*

Nevertheless, critical assessments about the trustworthiness of information dominated among the participants. Sceptical views were grounded by claiming that chatbots are newcomers whose training data are all too heterogeneous.

*AI in its current form has not earned that kind of trust yet, nor should anything or anyone deserve blind faith that easily. (R587-T9)*

*LLMs are trained on a bunch of bad data, too. They will spit input words that sound good together, regardless of whether they are true. (R239-T5)*

Critical assessments also dominated the Redditors' experiences obtained from the chatbot use. The evaluations about the untrustworthiness of chatbots were specified by offering explanatory answers. Due to the doubts about the trustworthiness of information generated by chatbots, the Redditors also provided recommendations about how to judge the credibility of AI-based outputs.

*Never get complacent with ChatGPT. Never, ever trust it for legal, financial or tax advice. (R442-T28)*

### **Chatbot's capability to generate relevant information**

The above findings suggest that the assessments of the correctness and trustworthiness of information drawn on is people's experiences about the chatbot's capability to generate relevant material for the users. Even though most Redditors were inclined to doubt the correctness and trustworthiness of answers offered by chatbots, some participants believed in the potential of these tools, despite their current limitations.

*AI has the capability of recognizing patterns, which can help find the correct sources of information. (R706-T9)*

The potential of AI chatbots was specified further by presenting comparative notions. The major advantage of chatbots was associated with their capability to quickly generate relevant data whose coverage is broader than the result lists offered by Google search, for example.

*ChatGPT is faster and more to the point/context aware (gives you the answer for the question you ask, not the answers already available on search engines top results). (R109-T30)*

Nevertheless, the capabilities of current AI chatbots as credible sources of information were associated with many limitations. They were reflected in the Redditors' opinions, as well as their experiences obtained from the use of chatbots.

*In my experience it (chatbot) cannot do math, it cannot reliably verify data, it cannot confirm or deny queries without sycophancy biases and hallucinations, it cannot reliably cite legitimate sources. (R432-T8)*

### **Usefulness of information**

Closely related to chatbot's capability to generate relevant answers, the credibility of information was assessed by drawing attention to its perceived usefulness in work-related, as well as non-work contexts. Similar to the criterion of chatbot's capability to generate relevant information, positive assessments were dominating. Although chatbots may not always be capable of generating fully correct information, they nevertheless may offer helpful answers in certain situations.

*It (chatbot) helped me identify a patient's sickness that even 2 attending doctors missed. (R564-T1)*

*I just fixed my 60' TV with Gemini 2.5. One 0.05\$ diode was broken. Showed it the photo of PCB, it suggested to measure with multimeter couple, possible points of failure. (R340-T3)*

The usefulness of information was specified further by means of comparisons. For example, participant R24-T3 was convinced that 'Gemini certainly beats through SEO (search optimisation) websites and obscure Reddit threads'. Despite the chatbots' perceived strengths such as these, these tools do not always generate useful information.

*Every time I have tried to use it for physics (writing a PhD thesis makes you desperate), it (chatbot) was absolute garbage. (R450-T8)*

### **Verifiability of information**

Given the opaque nature of AI algorithms, the verifiability of information generated by chatbots is often a demanding task. Nevertheless, a few participants believed that the veracity of information generated by chatbots can be confirmed and thus proven credible.

*Large Language Models can reference peer-reviewed sources and academic material especially when integrated with citation tools or plugins that point to real, verifiable sources (like academic databases, news outlets, or even Wikipedia itself). (R763-T5)*

Notwithstanding, most Redditors doubted the verifiability of information because chatbots are not designed to check the veracity of material offered for the user. The participants also reported their failed attempts to verify information generated by AI chatbots.

*I used a deep research report from Google's AI for a project at work, as a basis with verification. I found its references lead to dead links basically immediately, verification failed. (R332-T8)*

### **Accuracy of information**

The quantitative analysis revealed that the participants seldom drew on the accuracy of information while assessing the credibility of answers generated by AI chatbots. This may be due to the common belief that chatbots tend to hallucinate; fabricated information may not offer an exact picture of reality. Nevertheless, the participants reported examples of cases in which chatbots had offered sufficiently accurate information.

*There is plenty of things were learning something 95% accurate is good enough. I'm using it to learn French. It (chatbot) getting a word or conjugation wrong once is not exactly a concern. I'm sure not everything it tells me is 100% correct, and I'm fine with that. (R634-T29)*

Reflecting the existence of hallucinations, many of the opinions about the accuracy of information were critical. Critical views on the insufficient accuracy were also articulated in the experiences obtained from the use of chatbots.

*I literally asked it (chatbot) how many stars are within 10 light-years of Earth. It said zero. Then it listed 2, 4 and 6 but neither is within 10 (light-years). (R905-T3)*

The provision of inaccurate information may also have practical consequences. In the worst cases, they may be serious in health contexts if the user blindly relies on a chatbot.

*I have had things like outdated models presented as accurate (e.g., that excessive oxygenation is bad in COPD patients because it suppresses their respiratory drive when the answer has more to do with how it affects dead space). (R53-T2)*

### **Consistency of information**

One of the problems undermining the credibility of answers generated by AI chatbots is that they may offer conflicting information. There were only a couple of examples in which the Redditors believed that chatbots can provide consistent information.

*It (chatbot) does a fine job summarizing news for me. In the instances where I have read the summary and wanted to know more, the summary has been consistent with the full article. (R24-T7)*

However, critical assessments about the provision of inconsistent information were more common.

*I have had plenty of examples where it cites a credible source, only for the source to contradict with what the AI answer spat out. (R124-T30)*

### Unbiasedness of information

As LLMs are trained by a large body of heterogeneous information, it may contain bias which is difficult to identify from the answers generated by chatbots. In general, the Redditors' experiences about the provision of balanced information were rare.

*I will use Grok as my example. I find this one to be quite unbiased and very helpful because of how it (Grok) presents its answers. (R173-T18).*

The participants' experiences reporting about the provision of partial or unbalanced information were more common particularly in cases in which the answers generated by chatbots dealt with politically sensitive issues.

*Ask any AI chatbot about the situation in Gaza. Almost every single one will give you a final answer that the best answer is that Gaza should have been made its own country decades ago. Which is nothing but an open-ended answer that reinforces anyone's particular point of view of the subject. (R24-T27)*

### Coverage of information

The provision of biased information is also reflected in the coverage of answers generated by chatbots. All participants assessing the coverage of information were critical in this regard. It was argued that chatbots draw on a limited set of information sources, thus omitting important material. One of the major problems faced in the assessment of the coverage of information is that chatbots are incapable of identifying their limitations regarding the gaps and specificity of information.

*AI chatbots are very bad at saying, "I don't know," or openly admitting where there are gaps in their results. (R510-T18)*

Notwithstanding, individual chatbots can offer cues about the insufficient coverage of information, thus helping to evaluate its credibility. Such cues should be seen as a signal encouraging the users to check the coverage of information.

*Grok specifically will often say when something cannot be known for certain from its searches and will say something like "further inquiry is required" when there are inconclusive results. (R173-T18)*

### Currency of information

Finally, the credibility of information was assessed by the criterion of currency. Similar to the coverage of information reviewed above, the Redditors felt that chatbots tend to be wanting in this regard. The critical views were reflected in the participants' opinions, as well as their experiences obtained from the use of chatbots.

*LLMs have whatever data they were trained on like several or many months ago. If you are reading news articles today about a thing, you can be 100% sure that the model has no clue what you are talking about. (R79-T27)*

*Due to absence of the most recent studies, I have stopped using it (OpenEvidence) altogether. (R98-T2).*

Given the problems originating from the lack of up-to-date information, the participants were active to provide tips about how to ensure that information seeking would result in the identification of current material.

*In fact, a normal search engine will probably provide more recent and relevant articles if you know the right keywords. (R804-T2)*

## Discussion

The main contribution of the present study is the elaboration of the criteria by which people assess the credibility of information content generated by AI chatbots and how such assessments are grounded by making use of diverse answer strategies. The findings draw on the quantitative and qualitative analysis of credibility assessments made by 900+ participants contributing to Reddit discussion threads. Considering that the online participants represent a wide range of views on AI chatbots and experiences obtained about their use, the findings offer a realistic picture of how people assess the credibility of information generated by current LLM applications.

In sum, the findings suggest that correctness, trustworthiness and usefulness of information, as well as the chatbot's capability to generate relevant information are particularly important credibility criteria. Moreover, criteria such as accuracy, verifiability, coverage and currency of information are used to assess the information content of answers offered by chatbots. The importance of correctness and trustworthiness of information may be due to how AI chatbots are seen as newcomers which have not yet established their position as credible sources of information, as compared to Google search engine and Wikipedia. The findings suggest that overall, people tend to take a critical view on AI chatbots, as far as their credibility as information sources is concerned. The critical stance is mainly due to the risk of the provision of hallucinated data generated by chatbots. Even though chatbots are seen capable of offering useful information in certain situations, it is doubted that such information may not be fully trustworthy and correct. Reflecting the limitations of current LLMs, critical assessments also dominate while weighing the credibility of information by means of other criteria such as accuracy, coverage, currency, consistency and verifiability of information. Overall, the main recommendation arising from the Redditors' assessments of the credibility of AI-generated information can be crystallised as 'trust with reservations and always verify'.

While assessing the credibility of information, the participants mainly grounded their judgments in personal opinions. Assessments expressed in opinions deal with *presumed credibility* of information, that is, they pertain to the degree of belief a perceiver holds due to general assumptions in their mind (Tseng and Fogg, 1999). Redditors also often grounded their assessments in their use experiences. Assessments of this kind are indicative of *experienced credibility*: the degree of belief formed through direct, firsthand experience about an object (Tseng and Fogg, 1999). The popularity of experience answers is understandable because they enable people to express what is most familiar and concrete to them when talking about chatbots. It is easier to report one's recent experience about a useless answer offered by ChatGPT than make use of answer strategies that are cognitively more demanding. Such strategies include the comparison of accuracy of answers offered by diverse chatbots and the explanation of why information generated by chatbots is difficult to verify.

The reflection of the novelty value of the above findings is rendered more difficult, due to the paucity of similar studies. However, a few comparative notions can be made. First, the observations of the present investigation support the findings of prior studies in that information generated by chatbots is deficient in some regard. Such deficiencies manifest themselves in

incorrect or erroneous answers (Choi et al., 2025; Shen et al., 2023), the provision of inaccurate data (Kim et al., 2025) and the presentation of outdated or incomplete information (Murashko, 2025). On the other hand, the results of the present study lend support to conclusions drawn by prior investigations in that chatbots can generate information whose credibility is sufficient in certain situations (Barbosa-Silva et al., 2024). Lim and Hong (2025, pp. 137-138) demonstrated that people's concern about the risk of obtaining misleading information from AI chatbots does not necessarily mean that their use is abandoned. Many people continue seeking information from chatbots, despite recognising their potential for factual errors. This counterintuitive pattern may reflect users' growing familiarity with AI chatbots' limitations and a willingness to tolerate minor inaccuracies in exchange for convenience and speed of information seeking. People do not require complete credibility from chatbots to continue using them, but rather credibility that is good enough.

The above conclusions concur with observations about *reflexive mistrust* (Citrin and Stoker, 2018). Reflexive mistrust deals with the critical engagement with information available in diverse sources. Given the occurrence of hallucinations, the Redditors exhibited *mistrust* in that they questioned, verified and cross-referenced information generated by AI chatbots in order to ensure that such information would not contain fabricated elements. Reflexive mistrust of this kind can be understood as a form of information resilience - the capacity to sustain informed decision-making under conditions of contradiction and uncertainty. Rather than undermining trust, the critical stance drawing on reflective mistrust enables individuals to maintain control over their information environment, balancing openness to new data with protective scepticism. On the other hand, the Redditors also exhibited *distrust*, that is, a settled belief that an information source is inherently untrustworthy (Citrin and Stoker, 2018). Distrust appeared most clearly in cases in which it was believed that LLM powered chatbots draw on low-level training data originating from social media forums such as blogs and online discussion threads.

While considering the possibilities to enhance the credibility of information generated by chatbots, the Redditors mainly drew attention to the need to check the correctness, accuracy, coverage and currency of answers offered by AI tools. To compare, the participants seldom reflected the potential of prompt engineering, that is, the process of designing, crafting and refining instructions (prompts) to guide generative AI models toward producing desired, accurate, and relevant outputs (Lo, 2023). One of the promising generative AI techniques is *retrieval augmented generation* (RAG). It can enhance the responses generated by LLM so that it references an authoritative knowledge base outside of its training data sources. Kelly et al. (2025) in their study on the effectiveness of AI chatbots in answering questions about type 2 diabetes found that the RAG aspect of the LLM (when a source is cited) resulted in 94% of responses being completely appropriate in one-off questions and 100% being appropriate in a test conversation. This suggests that the formulation of the prompt is critical to ensure that chatbots provide only the requested information, do not stray from the topic of interest, and do not fabricate information. On the other hand, as Magesh et al. (2025) demonstrated, RAG is not necessarily a panacea for the elimination of hallucinations. The findings indicate that while RAG appears to improve the performance of language models in answering legal queries, the hallucination problem persisted at significant levels. This is because RAG may encounter its limitations while generating more specific responses to legal issues (Magesh et al., 2025, p. 220-221). We may think that the above reservations are also relevant in other domains while making attempts to enhance the credibility of information generated by AI chatbots.

## Conclusion

The present investigation contributed to empirical research on credibility assessment by focusing on information generated by AI chatbots. The findings highlight that in this regard,

people's opinions and experiences about the correctness, trustworthiness and usefulness of information are particularly important. As the present study focused on a sample of Reddit discussion threads, the findings cannot be generalised to concern credibility assessments made in other contexts. This is because the critical tone of most assessments presented by the Redditors may be due to the particular features of online discussion culture. It may encourage the articulation of critical or negative rather than positive assessments while debating controversial issues in particular (Savolainen, 2023). Thus, there is a need to expand the research approach by making use of empirical data gathered from other groups of people. To this end, interviews with diverse groups of chatbot users, as exemplified by Choi, Bak and An (2025) can be particularly fruitful. It is evident that interviews - particularly if substantiated with the analysis of critical incidents dealing with chatbot use in information seeking - would offer a more nuanced picture of the ways in which people assess the credibility of information generated by AI chatbots.

## About the author

**Reijo Savolainen** is Professor Emeritus at the Faculty of Information Technology and Communication Sciences, Tampere University, Finland. His main research interests are in the theoretical and empirical issues of everyday information practices and information behaviour. He can be contacted at [Reijo.Savolainen@tuni.fi](mailto:Reijo.Savolainen@tuni.fi)

## References

- Barbosa-Silva, J., Driusso, P., Ferreira, E.A., & de Abreu, R.M. (2024). Exploring the efficacy of artificial intelligence: a comprehensive analysis of CHAT-GPT's accuracy and completeness in addressing urinary incontinence queries. *Neurourology and Urodynamics*, 44(1), 153-164. <https://doi.org/10.1002/nau.25603>
- Behesti, M., Toubal, I.E., Alaboud, K., Almalaysha, M., Ogundele, O.B., Turabieh, H., Abdalnabi, N., Boren, S.A., Scott, G.J., & Dahu, B.M. (2025). Evaluating the reliability of ChatGPT for health-related questions: a systematic review. *Informatics*, 12(9), 9. <https://doi.org/10.3390/informatics12010009>
- Choi, W., Bak, H., & An, J. (2025). College students' credibility assessments of GenAI-generated information for academic tasks: an interview study. *Journal of the Association for Information Science & Technology*, 76(6), 867-883. <https://doi.org/10.1002/asi.24978>
- Citrin, J., & Stoker, L. (2018). Political trust in a cynical age. *Annual Review of Political Science*, 21(1), 49-70. <https://doi.org/10.1146/annurev-polisci-050316-092550>
- Eldridge, A. (2025). Reddit. American social media forum website. *Britannica Money*. <https://archive.org/details/httpswww.britannica.commoneyreddit> (Internet Archive)
- Grimes, M., von Krogh, G., Feuerriegel, S., Rink, F., & Gruber, M. (2023). From scarcity to abundance: scholars and scholarship in an age of generative artificial intelligence. *Academy of Management Journal*, 66(6), 1617-1624. <https://doi.org/10.5465/amj.2023.4006>
- Hanss, K., Sarma, K.V., Glowinski, A.L., Krystal, A., Saunders, R., Halls, A., Gorrell, S., & Reilly, E. (2025). Assessing the accuracy and reliability of large language models in psychiatry using standardized multiple-choice questions: cross-sectional study. *Journal of Medical Internet Research*, 27, e69910. <https://doi.org/10.2196/69910>

- Hilligoss, B. & Rieh, S.Y. (2008). Developing a unifying framework of credibility assessment: construct, heuristics, and interaction in context. *Information Processing & Management*, 44(4), 1467-1484. <https://doi.org/10.1016/j.ipm.2007.10.001>
- Johnson, S.B., King, A.J., Warner, E.L., Aneja, S., Kann, B.H., & Bylund, C.L. (2023). Using ChatGPT to evaluate cancer myths and misconceptions: artificial intelligence and cancer information. *NCI Cancer Spectrum*, 7(2), pkad015. <https://doi.org/10.1093/jncics/pkad015>
- Kelly, A., Noctor, E., Ryan, L., & van de Ven, P. (2025). The effectiveness of a custom AI chatbot for type 2 diabetes mellitus health literacy: development and evaluation study. *Journal of Medical Internet Research*, 27, e70131. <https://doi.org/doi:10.2196/70131>
- Kim, Y.B., Vilches, S.L., & Shapiro, S. (2025). Testing the capability of generative artificial intelligence for parent and caregiver information seeking. *Family Relations*, 74(3), 1266-1284. <https://doi.org/10.1111/fare.13167>
- Kumar, N. (2025). *Reddit statistics of 2025: users & revenue data*. Demandsage. <https://archive.org/details/httpwww.demandsage.comreddit-statistics> (Internet Archive)
- Lim, J.S., & Hong, N. (2025). Perceived search overload, generative AI credibility, and comparative usefulness: a channel complementarity approach to health information seeking. *Health & New Media Research*, 91, 124-145. <https://doi.org/10.22720/hnmr.2025.00115>
- Lim, J.S., Shin, D., Lee, C., Kim, J., & Zhang, J. (2025). The role of user empowerment, AI hallucination, and privacy concerns in continued use and premium subscription intentions: an extended technology acceptance model for generative AI. *Journal of Broadcasting & Electronic Media*, 69(3), 183-199. <https://doi.org/10.1080/08838151.2025.2487679>
- Lincoln, Y.S., & Guba, E. (1985). *Naturalistic inquiry*. Sage.
- Liu, S., Hu, Y., Tian, Z., Jin, Z., Ruan, S., & Mao, J. (2024). Investigating users' search behavior and outcome with ChatGPT in learning-oriented search tasks. In *SIGIR-AP 2024: Proceedings of the 2024 Annual International ACM SIGIR Conference on Research and Development in Information Retrieval in the Asia Pacific Region, Tokyo, December 9-12, 2024* (pp. 103-113). Association for Computing Machinery. <https://doi.org/10.1145/3673791.3698406>
- Lo, L.S. (2023). The art and science of prompt engineering: a new literacy in the information age. *Internet Reference Services Quarterly*, 27(4), 203-210. <https://doi.org/10.1080/10875301.2023.2227621>
- Lund, B.D., Mannuru, N.R., Katta, M., Hota, S.S.L.M., Pamukuntla, A., Uppala, S., Kola, S.M., & Mannuru, A. (2025). Bringing artificial intelligence (AI) into health information seeking behavior: a study of AI and information seeking research. *Journal of Health Communication*, 30(10-12), 330-335. <https://doi.org/10.1080/10810730.2025.2533820>
- Magesh, V., Surani, F., Dahl, M., Suzgun, M., Manning, C.D., & Ho, D.E. (2025). Hallucination-free? Assessing the reliability of leading AI legal research tools. *Journal of Empirical Legal Studies*, 22(2), 216-242. <https://doi.org/10.1111/jels.12413>
- Mendel, T., Singh, N., Mann, D.M., Wiesenfeld, B., & Nov, O. (2025). Laypeople's use of and attitudes toward large language models and search engines for health queries: survey study. *Journal of Medical Internet Research*, 27, e64290. <https://doi.org/10.2196/64290>
- Metzger, M.J., Flanagin, A.J., Eyal, K., Lemus, D.R. & McCann, R.M. (2003). Credibility for the 21st century: integrating perspectives on source, message, and media credibility in the

- contemporary media environment. In P.J. Kalbfleisch (Ed.), *Communication Yearbook*, 27 (pp. 293-335). Lawrence Erlbaum Associates.
- Miles, M.B. & Huberman, A.M. (1994). *Qualitative data analysis: an expanded sourcebook* (2nd ed.). Sage.
- Murashko, N. (2025). Communicative AI, trust and the stories of war: an ethnographic exploration of user evaluation of ChatGPT 3.5's responses on the Russian-Ukrainian war. In P. Åker & A. Kaun (Eds.), *Gendering media: framing of AI, interacting with ChatGPT, and anti-fandom* (pp. 43-70). Södertörns högskola, Sweden. <https://archive.org/details/httpswww.diva-portal.orgsmashgetdiva21947890fulltext01.pdf-page43> (Internet Archive)
- Ray, P.P. (2023). ChatGPT: a comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, 3, 121-154. <https://doi.org/10.1016/j.iotcps.2023.04.003>
- Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality and Quantity*, 52(4), 1893-1907. <https://doi.org/10.1007/s11135-017-0574-8>
- Savolainen, R. (2023). Assessing the credibility of information sources in times of uncertainty: online debate about Finland's NATO membership. *Journal of Documentation*, 79(7), 30-50. <https://doi.org/10.1108/JD-08-2022-0172>
- Savolainen, R. (2025). Seeking and sharing information about the threat of nuclear war. *Journal of Librarianship and Information Science*, 57(1), 265-278. <https://doi.org/10.1177/096100062312192>
- Schuetzler, R.M., Giboney, J.S., Wells, T.M., Richardson, B., Meservy, T., & Sutton, C. (2024). Student interaction with generative AI: an exploration of an emergent information-search process. In *Proceedings of the 57th Hawaii International Conference on System Sciences* (pp. 7500-7509). <https://archive.org/details/httpsscholarspace.manoa.hawaii.eduserverapicorebitstreams4bf9154e-7067-4f7f-b60d-36946fd62dd9content> (Internet Archive)
- Sezgin, E., Jackson, D. I., Kocaballi, A. B., Bibart, M., Zupanec, S., Landier, W., Audino, A., Ranalli, M., & Skeens, M. (2025). Can large language models aid caregivers of pediatric cancer patients in information seeking? A cross-sectional investigation. *Cancer Medicine*, 14(1), e70554. <https://doi.org/10.1002/cam4.70554>
- Shen, X., Chen, Z., Backes, M., & Zhang, Y. (2023). In ChatGPT we trust? *Measuring and characterizing the reliability of ChatGPT*. arXiv:2304.08979. <https://arxiv.org/pdf/2304.08979>
- Sidhu, R.S., & Selvamogan, A. (2025). Assessing the quality and readability of AI chatbot responses to frequently asked questions about basal cell carcinoma. *Clinical Surgical Oncology*, 4(3), 1000094. <https://doi.org/10.1016/j.cson.2025.100094>
- Tseng, S., & Fogg, B.J. (1999). Credibility and computing technology. *Communications of the ACM*, 42(5), 39-44. <https://doi.org/10.1145/301353.301402>
- Wang, L., Li, J., Zhuang, B., Huang, S., Fang, M., Wang, C., Li, W., Zhang, M., & Gong, S. (2025). Accuracy of large language models when answering clinical research questions: systematic review and network meta-analysis. *Journal of Medical Internet Research*, 27, e64486. <https://doi.org/10.2196/64486>

- Westbrook, L. (2015). Intimate partner violence online: expectations and agency in question and answer websites. *Journal of the Association for Information Science and Technology*, 66(3), 599-615. <https://doi.org/10.1002/asi.23195>
- Yun, H.S., & Bickmore, T. (2025). Online health information-seeking in the era of large language models: cross-sectional web- based survey study. *Journal of Medical Internet Research*, 27, e68560. <https://doi.org/10.2196/68560>
- Zhang, S., Li, J., Cagiltay, B., Kirkorian, H., Mutlu, B., & Fawaz, K. (2025). A qualitative exploration of parents and their children's uses and gratifications of ChatGPT. *Family Relations*, 74(3), 1056-1071. <https://doi.org/10.1111/fare.13171>
- Zhong, T., & Fang, X. (2026). Understanding user trust in AI-generated content: an elaboration likelihood model perspective. *Online Information Review*, 50(1), 171-188. <https://doi.org/10.1108/OIR-01-2025-0041>

## Copyright

Authors contributing to *Information Research* agree to publish their articles under a [Creative Commons CC BY-NC 4.0 license](https://creativecommons.org/licenses/by-nc/4.0/), which gives third parties the right to copy and redistribute the material in any medium or format. It also gives third parties the right to remix, transform and build upon the material for any purpose, except commercial, on the condition that clear acknowledgment is given to the author(s) of the work, that a link to the license is provided and that it is made clear if changes have been made to the work. This must be done in a reasonable manner, and must not imply that the licensor endorses the use of the work by third parties. The author(s) retain copyright to the work. You can also read more at: <https://publicera.kb.se/ir/openaccess>