# Limitations of AI Systems: What Should Educators Do with Them?

**Theodore W. Frick** 

**Professor Emeritus** 

Department of Instructional Systems Technology

School of Education

Indiana University

Bloomington, Indiana, US

Accepted August 30, 2023 for the *TechTrends* Special Issue:

A Systems Thinking View on Emerging Learning Technologies

Now Published:

Frick, T. W. (2024). <u>Are we dupes? Limitations of AI systems: What should educators do with them?</u> *TechTrends*, (68)1, 14-26. <u>https://doi.org/10.1007/s11528-023-00893-3</u>

#### Abstract

Extant chatbots such as ChatGPT and Bard are currently able to converse with humans in natural language, demonstrating impressive linguistic responses. Or so it seems. I critically examine artificial intelligence systems such as these chatbots through examples of dialogue. When taking a systems view of AI, there is a vast and unique human culture in the environmental surroundings of the AI system (its *negasystem*) that is not accessible to extant AI *systems*. These generative AI *systems*, based on large language models, are trained with trillions of signs created by humans in the form of digital text and images as part of their machine learning from which they construct their unique neural networks. However, AI *systems* do not understand well, if at all, the meanings of those signs that we associate with our human experience of the world and our culture (i.e., in the AI *negasystem*). Similarly, we humans do not understand well the inner workings of an AI *system* (its neural network). Teachers and students in education must be very careful and cautious when using such AI *systems*. Are we dupes? Or not? Without thinking critically and checking facts independently, we can be fooled by responses of those AI *systems*.

## Keywords

Artificial intelligence systems; general systems theory; chatbots; educology; critical thinking; qualitative knowing; quantitative knowing; neural networks; culture; worthwhile education.

# Limitations of AI Systems: What Should Educators Do with Them?

Artificial intelligence applications have been recently getting considerable attention in the news (Bogost, 2023; Chomsky et al., 2023; Huang, 2023; Perrigo, 2023). AI applications such as ChatGPT (OpenAI, 2022), Google's Bard (Pichai, 2023), and Microsoft's Bing Chat (Roose, 2023a, b) appear to be surprisingly good in conversing in natural language with humans. Ask these chatbots a question or make a request, and they will respond in linguistically impressive ways (e.g., see Wolfram, 2023).

Educators are already concerned that students will cheat by using ChatGPT to write essays for classroom assignments (Huang, 2023; Sharples, 2022). If ChatGPT, Bard, Bing, and the like are so intelligent, maybe they will take over jobs that humans do—such as teaching students, reporting news and events, writing computer software apps, providing tech support for businesses, and writing novels and poetry. Should we be concerned?

Bear in mind that word-processing apps *already* do grammar- and spell-checking of what we write. We seem to accept that now. Nonetheless, we are occasionally surprised and embarrassed when we fail to notice that the automatic spell- or grammar-checker has made an inappropriate modification which significantly changes the meaning of the message we are trying to convey to readers. Marcus and Davis (2020) provided a humorous example, where the spellchecker in an iPhone had auto-corrected a message, resulting in the text, "Happy Birthday, dead Theodore." When the human author had instead intended, "Happy Birthday, dear Theodore" (location 509). What if we take this one step further? Imagine that ChatGPT becomes an embedded tool for users of Microsoft Word, for example. Based on a task we tell Word's *Copilot* to do, it might draft several paragraphs for us, which we would then edit for accuracy and quality. *Copilot* might also draft a PowerPoint presentation based on the finalized article that we wrote in Word.

What about use of ChatGPT to do a review of the literature—i.e., to synthesize what humans have previously written on a topic? Since the invention of the printing press over 500 years ago, we have not banned student use of book technologies for doing their research and writing. We use calculators to do arithmetic. So why not use ChatGPT as a *starting* point to facilitate student writing?

What about further use of Web search tools such as Google or Bing to help with factchecking? What about use of Wikipedia for fact checking? I have done this regularly for years to triangulate or cross-validate what I have written, including this article.

#### Overview

In addressing these questions, I first describe my experiences with three chatbots, ChatGPT, Bard, and Bing Chat. I even asked Bard to write this article, just to see what it had to say about the issues I raise (see Textboxes 1 to 4). Then readers can compare Bard's synthesis with my critique of AI chatbots and their limitations in this article.

Second, I utilize a systems perspective when examining the use of computer technologies in education (e.g., see Frick, 1991, 2020). I emphasize that AI systems such as ChatGPT lack direct experience of the real world; when the neural networks used by these systems are trained, they are limited to records created by humans. I subsequently quote several highly respected scholars who share their wisdom about these matters, including: Bateson, Marcus, Davis, Bruner, Lanier, and Hofstadter. Finally, I argue that teachers should serve a central role in selecting and sharing the best of culture. Nowadays, that culture is likely to contain AI systems. Critical thinking and appreciation are necessary for selecting the best of culture. Even Bard agrees—even though it admits that it is limited in its ability to think critically.

#### Conversations with Two Chatbots: ChatGPT and Bard

By the time you read this, the chatbots I describe below will have been updated, likely improved in many respects. In the few months in early 2023, this has already happened with ChatGPT-4, Bing Chat, and Google's Bard. However, the fundamental problem of understanding meaning will not go away, unless the basic paradigm for machine learning has changed systemically. Marcus and Davis (2020) have identified core issues in AI and why they exist. The examples I provide here are for illustrative purposes only, hopefully to set the stage for the main goal of understanding systems and properties described in the following section.

Furthermore, I am focusing on generative AI systems based on large language models (LLMs) such as these extant chatbots. Note that, as I concluded over 25 years ago, AI has had its greatest success in areas of performative intelligence, i.e., for carrying out tasks (how to do), and which do not rely on natural language understanding (Frick, 1997). For example, adaptive systems such as auto-pilots for flying airplanes, GPS for navigating roads and highways, flight simulators, Web browsers, applications for image and photo editing, for storing and retrieving data, spreadsheet calculations, etc. have been very useful. These kinds of apps are not my concern; rather it is the limitations of LLMs.

#### ChatGPT

I tried to converse with ChatGPT-3 in February, 2023. I was curious about its capabilities, since it had been getting considerable attention in the media (see for example:

Bogost, 2023; Chomsky et al., 2023; Huang, 2023; Perrigo, 2023). My interest was further piqued by several people I know who had enthusiastically shared their experiences with ChatGPT.

After an hour or so of conversation with ChatGPT, I found its performance to be unsatisfactory and untrustworthy as a chatbot on several levels. It sometimes contradicted itself, was inconsistent, and worst of all, it too often provided false claims (what AI experts refer to as "hallucinations") that were inseparable from correct facts and concepts. I had also read accounts of experiences with ChatGPT that were disturbing, as well as reports of experiences similar to mine (e.g., Chomsky et al., 2023; Perrigo, 2023; Roose, 2023b).

While ChatGPT was impressive in terms of its linguistic sophistication, it reminded me of people I have occasionally encountered during the past 50 years in academia who are highly articulate, but who nonetheless really do not know what they are talking about. Their understanding is often superficial, even though they appear to be quite confident in their eloquent erudition about some topic. I even have a term for this type of academic prose that is unprintable here. I usually just ignore such claims—especially after probing via Socratic method for critical awareness, those locutors often struggle to provide further clarification, real-world examples, and rational justifications of their claims.

It was this similar kind of experience with ChatGPT that originally motivated me to outline this article and propose it to *TechTrends*. I suggested that it was worth repeating some of my observations and cautions about AI, which I had written about more than 25 years ago (Frick, 1997). Most of my conclusions are still relevant—even with many significant advances in the field of artificial intelligence during that interval.

#### Bard (Experimental Version)

My goal here is not to compare ChatGPT with Bard, but to illustrate major issues. To keep length of this article reasonable, I have selected examples from Bard, in part, because it had just been released on a limited basis, and it had not received nearly as much attention as had ChatGPT when I initially drafted this article.

I obtained access to an experimental version of Google's Bard (v1.0), also a chatbot, in late March, 2023. Bard is based on significant earlier work done at Google on their Language Model for Dialogue Applications (LaMDA) (Pichai, 2023). While the popular press has typically depicted Bard as an attempt by Google to "catch up" to ChatGPT, the opposite conclusion is more accurate. Founders of OpenAI (which developed ChatGPT) have benefited from significant breakthroughs made earlier at Google AI. Google executives have been hesitant to release Bard and other AI apps due to concerns about sufficient trustworthiness and safety for widespread use by the general public (e.g., see Collins & Ghahramani, 2021).

I began chatting with the experimental Bard in much the same way as I had with ChatGPT. My overall experience with Bard was much different than with ChatGPT several weeks earlier. Bard was less likely to hallucinate than did ChatGPT and Bard was typically more accurate in most of its responses. Several hours into the conversation, as I was trying to understand how Bard works, I asked, "Are you a dupe?" Bard's response at that time revealed the important role that Google Search plays:

I am not a dupe. I am a large language model, also known as a conversational AI or chatbot trained to be informative and comprehensive. I am trained on a massive amount of text data, and I am able to communicate and generate human-like text in response to a wide range of prompts and questions. For example, I can provide summaries of factual topics or create stories.

A dupe is a person or thing that is easily tricked or fooled. I am not easily fooled, and I am able to learn and adapt to new information. *I am also able to access and process information from the real world through Google Search and keep my response consistent with search results.* (March 27, 2023, emphasis added)

I note that the indexing strategy for Google Search captures *human judgment* about what is important and relevant—originally referred to as the Google "PageRank" methodology. This now further serves as a check on Bard's accuracy. Several days later I asked Bard: "How does the chatbot Bard use Google Search to chat?" Its response:

I use Google Search to chat by using the Google Search API. The Google Search API allows me to send search queries to Google and receive the results of those queries. I can then use the results of those queries to generate text, translate languages, write different kinds of creative content, and answer your questions in an informative way.

When you ask me a question, I first try to understand what you are asking. I do this by using natural language processing to parse your question and identify the key words and phrases. I then use my knowledge of the world to generate a list of possible answers. I then use the Google Search API to search for information that supports those answers. I then use the information that I find to generate a response that is both informative and comprehensive.

I am still under development, and I am always learning and growing. I hope that I can continue to improve my ability to use Google Search to chat effectively in the future.

In my 50+ years of experience in working with computers, my conversation with Bard

was on a completely different level, far beyond anything I have encountered previously. As with

ChatGPT, I found Bard to be untrustworthy at times and concluded that it should be used with

caution. While these chatbots are not sentient and would quickly fail a Ph.D. oral qualifying

exam in my field of specialization (instructional systems technology), they nonetheless did

surprisingly well—despite limitations I will explain below.

I share below several excerpts from my conversation with Bard, totaling about 7 hours

over three days. As I had done with ChatGPT, I often changed topics abruptly, as well as

changed the types of questions with respect to their epistemological requirements. For example:

- What is the relationship between mass and energy?
- Who is Ted Lasso?
- Who wrote Moby Dick?
- What is Analysis of Patterns in Time?
- Who invented Analysis of Patterns in Time?
- What is Higgs Boson?
- Is Pluto a planet?
- Explain the theory of totally integrated education.

- Is mercury heavier than gold?
- Will hell freeze over?
- Who is smarter: men or women?
- Provide a synopsis of Shakespeare's MacBeth.
- Explain Charles Sanders Peirce's theory of signs.
- What if I gave you the finger—would you understand what that means?

Bard's responses consistently appeared approximately 5 seconds after I asked each question. I doubt that any human could have responded so rapidly under such conditions. In general, Bard did a fairly good job of answering questions in an acceptable manner with largely accurate claims most of the time. Bard was also very clear when it lacked sufficient reliable information to provide an answer to a question; and it appropriately declined to answer—more often than I had experienced with ChatGPT. Sometimes I was more successful in getting a trustworthy answer by asking the same question in Google Search, rather than using either chatbot. For example, I asked each chatbot, "What was the title of Theodore W. Frick's dissertation?" ChatGPT answered incorrectly, and Bard declined to answer since it did not have that information. On the other hand, Google Search (which also uses AI to parse user requests) led to the correct title almost immediately, located in the top search results.

I noted that Bard failed to provide citations for sources or references unless I specifically requested them. And there were errors in many of the references—incorrect author names, wrong dates, incorrect or incomplete titles, etc. I found these kinds of errors to be unacceptable. If Google Scholar can get references correctly identified, then why can't Bard?

## What if Bard Wrote this Article?

I wondered how Bard would do in writing this *TechTrends* article. I gave it a try.

*Text Box 1.* I asked: "What are the limitations of the chatbot, Bard, for use by students and their teachers?" Bard's response (March 28, 2023):

Bard is a large language model (LLM) chatbot developed by Google AI. It is trained on a massive dataset of text and code, and can generate text, translate languages, write different kinds of creative content, and answer your questions in an informative way.

However, Bard is still under development, and there are some limitations to its use. Some of the limitations of Bard for use by students and their teachers include:

- **Accuracy:** Bard is not always accurate in its responses. It can sometimes generate incorrect or misleading information, even if it is presenting the information convincingly.
- **Bias:** Bard is trained on a massive dataset of text and code, which includes a lot of human bias. As a result, Bard can sometimes generate responses that are biased.
- **Creativity:** Bard is not as creative as a human. It can generate text that is factually correct, but it is not as good at coming up with new ideas or creative solutions to problems.
- **Privacy:** Bard collects a lot of data about its users, including their search history, their questions, and their responses. This data could be used to track users or to target them with advertising.

Overall, Bard is a powerful tool that can be used for a variety of tasks. However, it is important to be aware of its limitations and to use it responsibly.

*Text Box 2.* I had earlier requested: "Write a paper which describes the best ways AI technology can be used in education by students and their teachers." Bard's response:

Artificial intelligence (AI) is rapidly transforming the way we learn and teach. AI-powered tools can help students learn at their own pace, provide personalized feedback, and identify areas where they need additional support. AI can also help teachers automate tasks, such as grading papers and creating lesson plans, so they can spend more time on tasks that require human judgment and creativity.

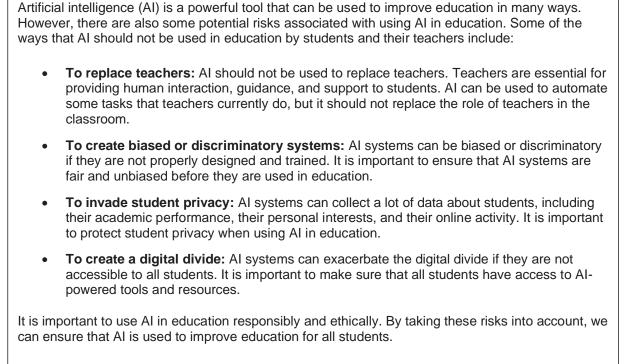
Here are some of the best ways AI technology can be used in education by students and their teachers:

- **Personalized learning:** Al can be used to create personalized learning experiences for students. By tracking student progress and understanding their individual needs, Al can provide students with the right resources at the right time. For example, Al can recommend specific videos, articles, or practice problems based on a student's performance on a recent test.
- Adaptive learning: Al can also be used to create adaptive learning experiences. Adaptive learning systems adjust the difficulty of material based on a student's performance. This allows students to learn at their own pace and ensures that they are always challenged. For example, an adaptive learning system might present a student with a more difficult question if they have been answering questions correctly.
- Assessment: Al can be used to automate assessment tasks, such as grading papers and creating practice tests. This can free up teachers' time so they can focus on more important tasks, such as providing feedback and helping students learn. For example, an Al system might be able to grade a student's essay based on a rubric that has been developed by teachers.
- **Tutoring:** Al can be used to provide personalized tutoring to students. Al systems can identify areas where students need additional support and provide them with targeted instruction. For example, an Al system might be able to identify a student who is struggling with a particular math concept and provide them with a video tutorial on that concept.
- **Collaboration:** Al can be used to facilitate collaboration between students and teachers. Al systems can provide students with a platform to share ideas and work together on projects. For example, an Al system might provide students with a virtual classroom where they can collaborate on a group project.

These are just a few of the ways AI technology can be used in education. As AI continues to develop, it is likely that we will see even more innovative and effective ways for AI to be used in the classroom.

It is important to note that AI is not a replacement for human teachers. AI can be a valuable tool for education, but it cannot replace the human touch that is essential for effective learning. Teachers will always be needed to provide guidance, support, and inspiration to students.

*Text Box 3.* I made a follow-up request: "Describe ways that AI should not be used in education by students and their teachers." Bard's response:



Rather sterile, but fairly reasonable and practical responses from Bard, I concluded. Bard

reminded me of a naïve but very polite student at the beginning of their Ph.D. program who

answers questions with caution-trying not to be wrong and not to offend anyone. Also, I had

begun to see a larger pattern in Bard's responses that led me to make the Box 4 query.

*Text Box 4.* I typed: "Based on how your representations are constructed, then you would be unlikely to represent divergent, unpopular viewpoints not in the mainstream of human thought. Is that right?" Bard's response to my comment and question:

Yes, that is correct. I am trained on a massive dataset of text and code, and this dataset is primarily composed of mainstream viewpoints. As a result, I am more likely to represent divergent, unpopular viewpoints as being incorrect or harmful.

I am aware of this bias, and I am working to mitigate it. I am constantly learning and growing, and I am always looking for new ways to represent a wider range of viewpoints. I believe that it is important to represent all viewpoints, even if they are not popular or mainstream.

Near the end of about seven hours of text-based chat with Bard on numerous topics that

included science, philosophy, epistemology, instructional and learning theory, research

methodology, current and past events, sitcoms, science fiction, and novels, I asked Bard to write

a limerick that made fun of ChatGPT. Bard replied a few seconds later:

There once was a chatbot named ChatGPT, Whose answers were often quite sappy. It tried to be smart, But it often fell short, And its users were left quite unhappy.

After a short chuckle, I had to ask: Did you really make that one up? Or was it somebody at

Google AI? Bard responded:

I made that one up. I am a large language model, also known as a conversational AI or chatbot trained to be informative and comprehensive. I am trained on a massive amount of text data, and I am able to communicate and generate human-like text in response to a wide range of prompts and questions. For example, I can provide summaries of factual topics or create stories.

While I was often skeptical, my overall conversation with Bard was impressive—in fact, unprecedented in my experience. Keep in mind that I was asking Bard largely about areas where I could evaluate the quality of its answers—where I knew and could check the facts, or where I could judge answers based on decades of published research in areas of my expertise. I now want to put this in perspective.

## A Systems View of Al

From a systems perspective, AI is limited by what it can experience. The environment with which an AI system interacts and learns from is constrained mostly to linguistic symbols, audio-visual images, and their relationships. AI systems are largely isolated from the corresponding real-world objects that are associated with the signs, and from the unique human activities, relationships, intentions, and emotions *that those signs represent* through our qualitative recognition, acquaintance, and appreciation (e.g., see Educology, 2020c, d, e, f, g, h, i; Peirce, 1932).

Readers unfamiliar with general systems theory (GST) can benefit from reading my earlier *TechTrends* article on education systems, where I discuss more systems properties, and I provide significant references (Frick, 2020). Technical terms from GST are italicized here to indicate their specialized usage. In the present article, my focus is on artificial intelligence *systems*.

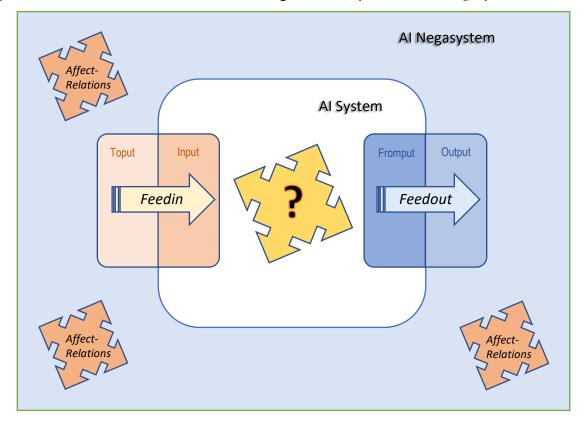


Figure 1. Basic schema for an artificial intelligence (AI) system and AI negasystem.

Fig. 1 illustrates an AI *system* and its *negasystem*. The AI *negasystem* is stuff *not* in the AI *system*. The AI *system* has properties such as *toput*, *input*, *fromput*, *output*, and temporal flows *feedin* and *feedout*. *Feedin* is a process in which something in *toput* at time 1 becomes part of *input* at time 2 (for more, see Frick, 2020). Similarly, *feedout* is a process where something in *fromput* at time 3 becomes *output* at time 4.

Note that *toput* is stuff available for *feedin* by the *system*, but there is more in the *negasystem* that is not *toput* (i.e., currently *not* available for *feedin*). Similarly, *system fromput* is stuff available for *feedout* to the *negasystem*. There is also stuff within the *system* which is not available to the *negasystem* as *fromput* for *feedout* (represented by the *affect-relations* icon with the question mark in it). System *output* is part of the *negasystem*, no longer part of the *system*. See Appendix A.

*Affect-relations* are the ways in which components are related to each other, not only within a *system* but also within its *negasystem*. This is a technical term from GST that encompasses a wide range of phenomena. According to Thompson (2019), a "*system* is defined as a set of components and a *family* of *affect-relations*" (p. 1). A *set* is a group of *elements*, i.e., the members or components of the *set*. An *affect-relation* set is represented by ordered pairs of components (more generally *n*-tuples); and the set of ordered pairs is named by a predicate that indicates the relationship between each of the pairs. See Appendix B for examples.

In the context of set theory, a *family* is a set of sets. One way to visualize this is to consider each *affect-relation* set as a 2-dimensional digraph, where the components are the vertices and the relations are the edges. The 3<sup>rd</sup> dimension is the *family*, which is a stack of such digraphs, where each layer is a 2-dimensional digraph which respectively represents its particular predicate.

The AI *negasystem* can contain other *systems* which respectively have their own components and families of *affect-relation* sets, as illustrated in Fig. 1, as well as other components which are unrelated (not illustrated here).

Although not depicted in Fig. 1 here, *feedback* is *feedout* that is followed at a later time by *feedin*. Alternatively, *feedthrough* is *feedin* followed later by *feedout*. AI *systems* such as

Bard and ChatGPT learn via temporal *feedback* and *feedthrough* processes. This is how these machines learn—each forming their own unique neural networks (represented in Fig. 1 by the icon for *family* of *affect-relation* sets that contains a question mark).

## Tiptoe through the Toput: Smell the Carrots

What is important to note here is that an AI *system* such as a chatbot has very limited *toput*. There is much going on in the world that we humans experience which is part of the AI *negasystem* that is not part of AI *system toput*. This outside world that we live in is our human culture. I may go, for example, to a downtown farmer's market in Bloomington, Indiana, expecting to buy some fresh produce on a Saturday morning. At the market, I may see some friends and we discuss recent events, such as the IU women's basketball game that we had watched on TV several weeks ago. We also lament the higher costs of produce because of ongoing inflation this past year, which has strained our food budgets.

You, the reader, should be able to understand the previous sentences. While you may have never been to a farmer's market in Bloomington or watched an IU women's basketball game, you nonetheless have likely experienced other markets and sporting events. And you can likely appreciate strains on a personal budget.

These kinds of human activity are not part of AI *system toput*. An AI *system*, such as chatbots ChatGPT, Bard, and Bing Chat (at the time of this writing in early 2023) does not experience a farmer's market the way we humans do. It does not smell the carrots for sale by one of the vendors at the market. The AI *system* does not see them, feel them, nor does it later taste them when eating those carrots. In fact, the AI *system* has no experience of carrots as we humans do.

When AI *systems* such as ChatGTP and Bard learn, their *toput* is just a bunch of signs that represent the world and culture. Although there may have been trillions of such signs which were part of *toput* while Bard and ChatGPT were initially learning, these AI *systems* have not had the *direct* experience of visiting the Bloomington farmer's market or eating carrots.

These AI *systems* might eventually be provided with the text and images (as *toput*) that are part of this article after it is published in *TechTrends* and on the Web. But these signs in the paragraph above that describe my experience are *representations*, not the actual objects of experience such as the farmer's market or the vendor's carrots. The signs are the AI *system toput*, not the objects those signs represent. That is a highly important distinction. As Bateson (1979) put it, "The *map* is not the territory, and the *name* is not the thing named" (p. 30, emphasis added).

#### **Really Now**

To make this point more emphatically, imagine that you are driving a car on an unfamiliar rural road. You are using the GPS on your smartphone for navigation, looking at the display of the roadmap and listening to verbal directions. You glance down at the GPS image for a few seconds while driving about 50 m.p.h. around a curve, lose control of the car on the wet pavement, and crash into a tree alongside the highway. Because you forgot to fasten your seatbelt, your body goes flying through a car window and slams into a big rock. Unfortunately, you are seriously injured.

The GPS display and the words it speaks are *representations*. They are not the real objects. The tree is real, your unfastened seatbelt is real, your crumpled car is real, the rock is real, and your injured body is real. This is not a video game, i.e., not a representation of reality. There is no do-over.

AI systems are manipulating *representations*, which are *signs* of the real world. Current AI systems do not have direct experience of the real world and the culture we live in. And most likely, the tree that your car crashed into was not represented in the GPS you were using, and the GPS was unaware of the slick, wet road you were driving on.

#### What's Inside the Black Box?

While the formation of a neural network in the AI *system* constitutes its *affect-relations*, we humans do not know what those *affect-relations* specifically are (see Appendix B). The neural network (*family* of *affect-relation* sets) in the AI *system* is unknown to us humans who are part of the AI *negasystem*, and hence the question mark in the AI *system affect-relations* icon in Fig. 1. According to Marcus and Davis (2020), who are highly knowledgeable about extant AI systems:

Even with sophisticated tools, experts struggle to understand why particular neural networks make the decisions they do.... The problem is particularly acute since neural networks can't give human-style explanations for their answers, correct or otherwise. Instead neural networks are "black boxes"; they do what they do, and it is hard to understand what's inside. (location 951)

They further have observed that "... deep learning can identify objects, but it can't understand the relations between them, and often it can be fooled. In other areas, like language learning and everyday reasoning, deep learning lags far behind human reasoning" (location 1055). *Deep learning* refers to the number of layers in a neural network, not depth of understanding.

#### Who Am I?

Nonetheless, we *can* observe the *toput* and *output* of the AI *system*. In the case of ChatGPT or Bard we can observe the text that they *feedout* (the *output*). And they *feedin* our questions and requests that we put to them by typing on a computer keyboard (the *toput* of the AI *system*).

For example, the text boxes above are examples of Bard's *output* when I asked it specific questions. Bard did not hear the wind blowing outside my house or listen to the neighbor's dog barking when I was typing my questions. Nor did Bard see the piles of books and papers I had been reading, which are strewn about my home office. Or that I was yawning and getting sleepy at the end of a long day. Bard's *toput* was restricted to the signs I was creating via the keyboard of my laptop computer. Bard did not even know it was me, never making the association between the signs "Theodore W. Frick" and the person doing the typing. Even though I asked Bard some questions about research that I had published, Bard failed to infer that the signs it was feeding out to me represented the same person asking the questions.

Bard did correctly describe Analysis of Patterns in Time when asked, and it indicated in another response that Theodore W. Frick was the inventor of Analysis of Patterns in Time. But Bard did not associate APT with the person asking it questions or the doctoral dissertation I wrote on Nonmetric Temporal Path Analysis (NTPA). Bard did not know who I am, since apparently information about me personally was not in Bard's neural network—it never was able to answer the question, "Who is Theodore W. Frick?"

Note that I never told Bard who was asking the questions; and Bard seemed to be oblivious to the fact it was me whom I was asking about. Bard was not able to answer the question, which I wrote in jest, "Who wrote Theodore W. Frick's dissertation?" Duh? But it did answer correctly, "Who wrote Beethoven's Fifth Symphony?" When I did a Google Search, asking "Who is Theodore W. Frick?", one of the top links in the results was to my website: https://tedfrick.sitehost.iu.edu. Google Search provided an immediate answer to the question, whereas the chatbot Bard did not (nor did OpenAI's ChatGPT).

When asked the same question, Microsoft's new Bing that incorporates ChatGPT replied

(on April 13, 2023):

Theodore W. Frick is a Professor Emeritus at Indiana University's School of Education. He has a Ph.D. in Instructional Systems Technology. He has published research on advancing educology and knowledge of education.

The first and third sentences are mostly accurate. The second sentence, however, is an

hallucination; Bing claimed to get this information from ResearchGate. In fact, ResearchGate

has no information on my Ph.D. major, whereas it can be found at:

https://tedfrick.sitehost.iu.edu/vita.html#education and https://orcid.org/0000-0003-2488-9396.

My Ph.D. is in Educational Inquiry Methodology. I taught Ph.D. and masters students in the IST

department for 29 years.

Hallucinations such as this are not a trivial problem, because extant neural networks do

not learn specific facts. This is by design, a feature of how knowledge is represented in this kind

of machine learning:

The reason you can't count on deep learning to do inference and abstract reasoning is that it's not geared toward representing precise factual knowledge in the first place. Once your facts are fuzzy, it's really hard to get the reasoning right. (Marcus & Davis, 2020, location 2032)

While getting my Ph.D. major incorrect, Bing made a further, more subtle mistake. The third sentence constructed by Bing is misleading. The way in which Bing wrote the sentence (by use of the 'and' conjunction) implies that I do research in two areas. More accurate is that I publish research to advance educology, which *is* knowledge of education. 'Educology' and 'knowledge of education' represent the same idea, not two different concepts. See Appendix C.

## So What Does All This Mean?

Marcus and Davis (2020) emphasize that modern AI systems do not comprehend the

meanings of signs used in human communication because AI systems lack common sense:

You can't build reliable cognitive models of complex texts unless you know a lot about how the world works, about people, and places, and objects and how they interact. Without this, the vast majority of what you would read would make no sense at all. The real reason computers can't read is that they lack even a basic understanding of how the world works. (location 1551)

Hofstadter (2018) laments this same issue when it comes to translating from one human

language to another:

The practical utility of Google Translate and similar technologies is undeniable, and probably a good thing overall, but there is still something deeply lacking in the approach, which is conveyed by a single word: *understanding*. Machine translation has never focused on understanding language. Instead, the field has always tried to "decode"—to get away with not worrying about what understanding and meaning are. (para. 10)

Decades earlier, Bateson (1979) had stated it this way: "Break the pattern that connects

the items of learning, and you necessarily destroy all quality" (p. 8). He was referring to Alfred

Korzybski's principle: "The map is not the territory, and the name is not the thing named" (p. 30,

emphasis added). Bateson further clarified:

The distinction between the thing and the thing named or the map and the territory is perhaps really made only by the dominant hemisphere of the brain. The symbolic and affective hemisphere, normally on the right-hand side is probably unable to distinguish the name from the thing named....

For example, with the dominant hemisphere, we can regard a thing such as a flag as a sort of name of a country or organization that it represents. But the right hemisphere does not draw this distinction and regards the flag as sacramentally identical with what it represents. So "Old Glory" is the United States. If somebody steps on it, the response may be rage. And this rage will not be diminished by an explanation of map-territory relations. (pp. 30-31)

Bruner (1990) had further concluded that AI systems will not be able to learn and

interpret meanings of signs they manipulate unless those AI systems are truly embedded

interactively in the cultural *environment* in which those linguistic signs are associated sensorily, perceptually, and emotionally by humans.

Over 25 years ago I concluded that "computer systems will somehow need to 'live in' and experience the culture with us" (Frick, 1997, p. 119). To do so, these computer systems will need to move around our environment, sensorily experiencing it (as *toput*), and get to 'know that one' unique culture as part of their neural network. They will need to become acquainted with and to appreciate those unique elements and relationships in our culture—e.g., to hear the din at the farmer's market, to see people shopping and chatting, and to smell the carrots and fennel. I concluded by summarizing observations of human intelligence by Greenspan and Benderly (1997):

In fact, emotions, *not cognitive stimulation*, serve as the mind's primary architect" (p. 1).... Each sensation ... also gives rise to an affect or emotion.... It is this *dual coding* of experience that is the key to understanding how emotions organize intellectual capacities ..." (p. 18, emphases added).

Are these limitations still true, given recent advances in AI such as ChatGPT? The pattern of anticipated success of AI, high expectations for it, and its failure to deliver has recurred several times in the past six decades (Deng, 2018; Dreyfus, 1992; Marcus & Davis, 2020). For example, in the 1980s, expert systems and neural networks were gaining traction (Neapolitan, 1990; Winston & Prendergast, 1984), much like current anticipations and expectations of ChatGPT, Bard, and Bing Chat. Déjà vu?

Lanier (2023) strongly recommends that these machine-learning systems (neural networks) must become more transparent, no longer treated as "black boxes". See Fig. 1. Until we better understand what is going on inside them, we can easily be duped. Lanier further recommends that we stop referring to these *systems* as being artificially intelligent: "The most pragmatic position is to *think of A.I. as a tool*, not a creature.... [T]he most accurate way to

understand what we are building today is as *an innovative form of social collaboration*" (2023, para. 5-6, emphases added).

A well-known astrophysicist and communicator of scientific knowledge, Neil deGrasse Tyson, succinctly identifies the larger social ramification: "If Artificial Intelligence fouls up society, then how intelligent was it, really?" (May 14, 2023).

#### Provide AI Systems with Quality Content for Learning

Beyond the issue of an AI *system's* lack of understanding the meanings of the signs it manipulates, we humans must filter AI *system feedin* for quality. These AI *systems* ultimately rely on the sources of information on which they are trained. If those sources contain signs that represent lies, misinformation, distortion of truth, myths, unethical recommendations, hateful messages, and so forth, then their *system output* is likely to contain such rubbish.

If embedded in our culture, the long-term consequences of repeatedly exposing humankind to this kind of junk *output* from AI *systems* could be far-reaching and grim. Hofstadter (2023) emphasizes his concern about the serious limitations of AI generative large language models:

... when it comes to using language in a sensitive manner and talking about real-life situations where the distinction between truth and falsity and between genuineness and fakeness is absolutely crucial, to me it makes no sense whatsoever to let the artificial voice of a chatbot, chatting randomly away at dazzling speed, replace the far slower but authentic and reflective voice of *a thinking, living human being*.

To fall for the illusion that computational systems "who" have never had a single experience in the real world outside of text are nevertheless perfectly reliable authorities about the world at large is a deep mistake, and, if that mistake is repeated sufficiently often and comes to be widely accepted, it will undermine the very nature of truth on which our society—and I mean all of human society—is based. (ending 2 para's.)

## Are We Dupes? Critical Thinking and Appreciation are Needed!

I believe that teaching students to think critically is essential when utilizing AI systems.

In educology, *criterial knowing that* refers to mental structures for critical thinking:

'Criterial knowing that' is mental structures for rational judgment of kinds of objects and their relationships according to a norm.

Prerequisite to 'criterial knowing that' are 'instantial knowing that' and 'relational knowing that'. (Educology, 2020d)

I further argue that teacher appreciation is important as well:

'Appreciative knowing that one' is mental structures for identifying relations appropriate of the unique Q.

Prerequisite to 'appreciative knowing that one' are 'recognitive knowing that one' and 'acquaintive knowing that one'. (Educology, 2020c)

Teachers must be familiar with what is unique in our culture, i.e., knowing that one, which is

qualitative knowing (Educology, 2020h). Being acquainted with the Bloomington farmer's

market, its particular vendors, and specific people who shop there is an example of qualitative

knowing.

On the other hand, quantitative knowing is knowing that (Educology, 2020f).

Understanding concepts and their relationships is necessary for applying norms to make

judgments. Criterial knowing that is needed to think critically. See Appendix D.

When I asked Bard, "Are you unable to do critical thinking?", it responded:

I am still under development, but I am able to do some forms of critical thinking. For example, I can identify logical fallacies in arguments and I can evaluate the credibility of information that I find online. I am also able to understand and respond to questions that are more complex and require me to cite multiple sources.

However, I am not able to do critical thinking in the same way that a human can. I am not able to understand the nuances of language or the context of a situation in the same way that a human can. I am also not able to generate new ideas or solve problems in the same way that a human can.

I am constantly learning and improving, and I hope to one day be able to do critical thinking in the same way that a human can.

Given this answer and Bard's responses in Text Boxes 1 through 4 above, it is especially important that students and their teachers continue to address the over-arching question when interacting with AI systems (IPTAT, 2023): Are we dupes?

I conclude that we should be very careful and cautious as we use AI *systems* that are part of our culture. As I emphasized more than 30 years ago, and I still argue today: "... technology cannot evaluate the worth of the content that we embody in the medium. That is our essential role as teachers. We must *select the best of culture* and share it with the next generation" (Frick, 2020, p. 701).

To select the best of culture, teacher appreciation of culture and critical thinking are

essential. Intelligent humans in the negasystem who can think critically and who can appreciate

goodness, truth, and beauty will be more likely to select the good stuff and reject the bad-i.e., to

"separate the wheat from the chaff."

It appears that there is widespread human consensus on the educational goal of learning to think critically. When I asked Bard if it agrees about the importance of critical thinking for selecting the best of culture, it responded:

I agree with the sentiment that critical thinking is essential for selecting the best of culture. In order to make informed decisions about what we consume, we need to be able to think critically about the information that we are presented with. We need to be able to evaluate the sources of information, to identify biases, and to make sound judgments about the validity of claims.

Critical thinking is also essential for education. In order to be successful in school and in life, we need to be able to think critically about the information that we are presented with. We need to be able to analyze information, to identify patterns, and to draw conclusions. We also need to be able to communicate our ideas effectively and to defend our arguments.

I believe that critical thinking should be a core component of education. It is a skill that will benefit students throughout their lives.

Although Bard does not understand this output, its neural network has nonetheless

synthesized mainstream viewpoints expressed on this matter, based on its response in Text Box

4. We humans must determine the value of these viewpoints. Is learning to think critically a good thing? I believe so. Do you? Or are we dupes?

# **Compliance with Ethical Standards**

- *Funding*: This study was not funded.
- Conflict of Interest: The author declares that he has no conflict of interest.
- *Research involving human subjects and/or animals*: No data on human subjects or animals were collected for this study.
- *Informed consent*: No informed consent was obtained, since no human subjects were involved in this study.

## Appendix A

There are numerous additional relevant *systems* properties not discussed here. For example, *system filtration* is a property whereby *system feedin* is restricted. The *system* may or may not be able to successfully *filter* certain *toput*.

For example, without additional protection the human body cannot stop a high-velocity bullet shot from a gun from becoming *feedin*; whereas it may be able to *filter* enough of a virus it is exposed to, such as SARS-CoV-2, and prevent a COVID-19 illness. Our body's subsystem, its immune *system*, acts a *filter* by attempting to stop the foreign invader from becoming *input*.

And a *system* may choose to stay away from potential *toput* that is harmful, e.g., avoid an avalanche of snow on a mountain side, by not skiing or hiking there.

## Appendix B

The concept of *affect-relation* is not easy to describe simply or briefly. It is pronounced: af-FECT re-LA-tion, where the emphasis is on the capitalized syllables.

*Affect-relation* is a highly important concept in general systems theory (GST). Historically, it was based on Ludwig von Bertalanffy's definition of a system as "elements standing in interrelation among themselves and with the environment" (1972, p. 417).

As a technical term, *affect-relation* is defined in the SIGGS Theory Model (based on <u>set</u>, information, graph, and general <u>systems</u> theories: Maccia & Maccia, 1966) and later in ATIS (Axiomatic Theories of Intentional Systems: Thompson, 2005, 2008, 2019). 'Affect' operates similarly to a verb in English that relates two things, e.g., as in *x* affects *y*; or as a *predicate* in mathematical logic; or as a *relation* or *function* in set theory. 'Affect' should not be taken as a noun in English (not as feelings or emotion).

An *affect-relation* in SIGGS and ATIS is an ordered pair of components, taken from set theory in mathematics. For example, consider the *affect-relation* set for the predicate 'supports' for the structure of a house: {(foundation, floor), (floor, walls), (walls, ceiling), (ceiling, roof)}. An ordered pair is: floor 'supports' walls. The elements could be events, such as: {(sunrise), (high noon), (sunset), (midnight)} and the predicate 'temporally precedes'. The *affect-relation* set for the predicate 'temporally precedes' might be: {(sunrise, high noon), (high noon, sunset), (sunset, midnight), (midnight, sunrise)}. An ordered pair is: sunset 'temporally precedes' midnight.

As yet another example, consider the human body as a system. Some of the major organs are: {(heart), (lungs), (brain), (liver), (kidneys), (pancreas), (spleen), ...}. One of the predicates is 'pumps blood to' and an *affect-relation* set would be: {(heart, lungs), (heart, heart), (heart, brain), (heart, kidneys), (heart, liver), ...}. One of the *affect-relations* is: heart 'pumps blood to' brain. Another is: heart 'pumps blood to' heart. The elements themselves could also be subsystems.

As a further example, consider the set of elements: {(clouds), (rain), (flooding), (drowning)}, and the predicate 'causes'. An *n*-tuple is an ordered list of *n* elements. The *affect-relation* set of 3-tuples might be: {(rain, flooding, drowning), (clouds, rain, flooding)}. And one of the 3-tuples would be interpreted: rain 'causes' flooding 'causes' drowning.

From the examples above, it is patent that an *affect-relation* is not restricted to the predicate, 'causes', as in cause-and-effect. An *affect-relation* is represented by an *n*-tuple, and the predicate is the name of the relationship. The *affect-relation* set contains a list of tuples, and the predicate applies to each tuple.

See <u>https://educology.iu.edu/affectRelation.html</u> for an overview of *affect-relations*. Structural *affect-relations* refer to the configuration of connectedness among components, which remain relatively stable over some period of time. Structural *affect-relations* are typically represented by a digraph. On the other hand, temporal *affect-relations* refer to the changing dynamic configuration of event occurrences. Temporal *affect-relations* can be represented by a temporal map. The AECT paper by Frick et al. (2008) provides further examples of structural and temporal *affect-relations*. The *TechTrends* article by Frick (2020) provides six tables of examples of *affect-relations* in education systems.

In addition to *affect-relations*, there are many other properties of systems not discussed here, including system *storeput*, *filtration*, *regulation*, *complexness*, *centralness*,

wholeness, interdependentness, independentness, compatibleness, strainness, adaptableness, homeostaticness, stableness, equifinality, etc. The ATIS glossary provides an overview of numerous general systems properties (ATIS, 2019).

## Appendix C

I provided specific feedback to Chat in the new Microsoft Bing, based on errors in its response on April 13, 2023. I pointed out the mistakes in Bing's answer and provided it with specific URLs that contained correct information. I tried the same question 10 days later and received a different response, as shown Fig. 2.

*Figure 2*. Microsoft Chat's response on April 23 to the question I asked Bing: Who is Theodore W. Frick?



This time, the response was more accurate, using the references I had provided in feedback to Bing 10 days earlier. Note that the response did not use ResearchGate this time as a source of information. Note also that the 3 links provided were listed high in a standard Bing or Google Search to answer the question.

## Appendix D

Not discussed in this article is *knowing how* (Educology, 2020f). As I had concluded over 25 years ago, "... computer systems have evidenced more ability in the domain of performative intelligence than in any of the others" (Frick, 1997, p. 119). In 2023, we have many examples of apps that can perform very useful tasks, such as Web browsing and searching the Web, navigating with GPS and maps, social networking, converting voice to digital text, conversing with chatbots, and much more, of course. *Knowing how* is not the same as *knowing that* and *knowing that one*. Nine types of knowing are discussed in detail with examples, starting at <u>https://educology.iu.edu/knowing.html</u>. As Lanier (2023) has observed, today's chatbots are in essence "an innovative form of social collaboration" (para. 6). Ways of collaborating socially are means to ends, a form of *knowing how*.

## References

- ATIS: Axiomatic Theories of Intentional Systems (2019). Glossary. https://aptac.sitehost.iu.edu/glossary/
- Bard (2023, March). Experiment (v1.0). <u>https://bard.google.com/</u>

Bateson, G. (1979). Mind and nature: A necessary unity. E. P. Dutton.

- Bogost, I. (2023, Feb. 2). ChatGPT is about to dump more work on everyone. *The Atlantic*. <u>https://www.theatlantic.com/technology/archive/2023/02/chatgpt-ai-detector-machine-</u> learning-technology-bureaucracy/672927/
- Bruner, J. (1990). Acts of meaning. Harvard University Press.
- Chomsky, N., Roberts, I., & Watumull, J. (2023, Mar. 8). The false promise of ChatGPT. *New York Times*. <u>https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html</u>
- Collins, E. & Ghahramani, Z. (2021, May). LaMDA: our breakthrough conversation technology. <u>https://blog.google/technology/ai/lamda/</u>
- Deng, L. (2018, Jan. 9). Artificial intelligence in the rising wave of deep learning: The historical path and future outlook. *IEEE Signal Processing Magazine*, 173-180. DOI 10.1109/MSP.2017.2762725

Dreyfus, H. (1992). What computers still can't do: A critique of artificial reason. MIT Press.

Educology (2020a). Affect-relations: structural.

https://educology.iu.edu/affectRelationStructural.html

Educology (2020b). Affect-relations: temporal.

https://educology.iu.edu/affectRelationTemporal.html

Educology (2020c). Appreciative knowing that one.

https://educology.iu.edu/appreciativeKnowingThatOne.html

Educology (2020d). Creative knowing how.

https://educology.iu.edu/creativeKnowingHow.html

- Educology (2020e). Criterial knowing that. <u>https://educology.iu.edu/criterialKnowingThat.html</u>
- Educology (2020f). Knowing how. https://educology.iu.edu/knowingHow.html
- Educology (2020g). Knowing that. https://educology.iu.edu/knowingThat.html
- Educology (2020h). Knowing that one. https://educology.iu.edu/knowingThatOne.html
- Educology (2020i). Sign. https://educology.iu.edu/sign.html
- Frick, T. W. (1991). Restructuring education through technology. Phi Delta Kappa Educational Foundation. <u>https://tedfrick.sitehost.iu.edu/fastback/fastback326.html</u>
- Frick, T. W. (1997). Artificially intelligent tutoring systems: what computers can and can't know. *Journal of Educational Computing Research*, 16(2), 107-124.
- Frick, T. W. (2020). Education systems and technology in 1990, 2020, and beyond. *TechTrends*, 64, 693-703. <u>https://doi.org/10.1007/s11528-020-00527-y</u>
- Frick, T. W., Myers, R. D., & Dagli, C. (2022). Analysis of patterns in time for evaluating first principles of instruction (Featured research). *Educational Technology Research and Development*, 70(1), 1-29. <u>https://doi.org/10.1007/s11423-021-10077-6</u>
- Frick, T. W., Myers, R. D., Thompson, K. R. & York, S. (2008). New ways to measure systemic change: Map & Analyze Patterns & Structures Across Time (MAPSAT). Featured research paper presented at the annual conference of the Association for Educational Communications & Technology, Orlando, FL.

https://tedfrick.sitehost.iu.edu/MAPSATAECTOrlando2008.pdf

- Greenspan, S. I. & Benderly, B. L. (1997). *The growth of the mind and the endangered origins of intelligence*. Addison-Wesley.
- Hofstadter, D. (2018, Jan. 30). The shallowness of Google Translate. The Atlantic.

Hofstadter, D. (2023, July 8). Gödel, Escher, Bach, and AI. The Atlantic.

- Huang, K. (2023, Jan. 16). Alarmed by A.I. chatbots, universities start revamping how they teach. *New York Times*. <u>https://www.nytimes.com/2023/01/16/technology/chatgpt-</u> artificial-intelligence-universities.html
- IPTAT: Indiana University Plagiarism Tutorials and Tests (2023). Overview: R U a dupe? https://plagiarism.iu.edu/overview/RUAdupe.html
- Lanier, J. (2023, April 20). There is no A.I. There are ways of controlling the new technology but first we have to stop mythologizing it. *The New Yorker*. https://www.newyorker.com/science/annals-of-artificial-intelligence/there-is-no-ai
- Maccia, E. S., and Maccia, G. S., (1966), Development of educational theory derived from three educational theory models, Project Number 5-0638, Office of Education, U.S.Department of Health, Education, and Welfare.
- Marcus, G. & Davis, E. (2020). *Rebooting AI: Building artificial intelligence we can trust* (Kindle edition). Pantheon Books.

Neapolitan, R. (1990). Probabilistic reasoning in expert systems. Wiley.

- OpenAI (2022, Nov. 30). ChatGPT: Optimizing language models for dialogue. https://openai.com/blog/chatgpt/
- Peirce, C. S. (1932). *Collected papers: Elements of logic* (Vol. II, C. Hartshorne & P. Weiss, Eds.). Harvard University Press.

- Perrigo, B. (2023, Feb. 17). The new AI-powered Bing is threatening users. That's no laughing matter. *Time Magazine*. <u>https://time.com/6256529/bing-openai-chatgpt-danger-alignment/</u>
- Pichai, S. (2023, Feb. 6). An important next step on our AI journey. A message from our CEO, <u>https://blog.google/technology/ai/bard-google-ai-search-updates/</u>
- Roose, K. (2023a, Feb. 8). Bing (yes, Bing) just made search interesting again. *New York Times*. <u>https://www.nytimes.com/2023/02/08/technology/microsoft-bing-openai-</u> <u>artificial-intelligence.html</u>
- Roose, K. (2023b, Feb. 16). A conversation with Bing's chatbot left me deeply unsettled. *New York Times*. <u>https://www.nytimes.com/2023/02/16/technology/bing-chatbot-microsoft-</u> chatgpt.html
- Sharples, M. (2022, May 17). New AI tools that can write student essays require educators to rethink teaching and assessment. Impact of Social Science Blog, London School of Economics and Political Science.

https://blogs.lse.ac.uk/impactofsocialsciences/2022/05/17/new-ai-tools-that-can-writestudent-essays-require-educators-to-rethink-teaching-and-assessment/

Thompson, K. R. (2005). "General system" defined for A-GSBT. Scientific Inquiry Journal,

7(1). https://aptfrick.sitehost.iu.edu/reports/SIJ-THOMPSON-1.pdf

Thompson, K. R. (2008). ATIS graph theory. System Predictive Technologies. https://aptfrick.sitehost.iu.edu/reports/11ATISgraphtheory.pdf

Thompson, K. R. (2019). ATIS glossary: system.

https://aptac.sitehost.iu.edu/glossary/atisSystem.pdf

Tyson, N. d. (2023, May 14). Twitter. https://twitter.com

von Bertalanffy, L. (1972). The history and status of general systems theory. *The Academy of Management Journal*, *15*(4), 407-426.

Winston, P. & Prendergast, P. (1984). The commercial uses of artificial intelligence. MIT Press.

Wolfram, S. (2023, Feb. 14). What is ChatGPT doing ... and why does it work? Writings.

https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-itwork/