Tyler Cheng

Fox

English 12

25 April 2023

AI Search: Revolution or Resounding Failure?

Thesis: *Bing's chatbot has some clear advantages over a traditional search engine, but its inaccuracy and unpredictability hold it back from becoming a mainstream technology.*

I. Background: GPTs, OpenAI, and Bing
    A. What is a GPT?
    B. The AI Boom
    C. Bing Chatbot
II. Pro 1 - Convenience of answers
III. Pro 2 - Detailed with sources
IV. Con 1 - Prone to error
V. Con 2 - Unpredictable responses
VI. Con 3 - Design flaws

Tyler Cheng

Fox

English 12

25 April 2023

What is the most searched term on Microsoft's search engine Bing? The answer: Google. While the truthfulness of this statement is not known, it is believable enough for most people, knowing the disparity in popularity between the two services. Bing has long been the subject of much ridicule from the Internet community, with many memes and journalists mocking the search engine for bizarre results and its lack of popularity compared to Google. However, as of recently, Bing has begun to make efforts to improve its service, and its latest addition is a virtual chatbot capable of answering questions much like the popular AI service ChatGPT. Unfortunately, just like many other topics surrounding artificial intelligence, this service has quickly generated controversy as people debate its usefulness compared to Google. While Bing's chatbot introduces some notable strengths over traditional search engines, it is held back by its issues of inaccuracy and unpredictability.

To begin, Microsoft added the new functionality to Bing as a result of recent trends on the market of technology. The chat AI is based on technology developed by OpenAI, which Microsoft first invested in in 2019, according to *The New York Times* author Kevin Roose. It uses a generative pre-trained transformer (GPT), a type of machine learning model trained on data from many forms of written language, such as books or websites. As outlined by Fawad Ali, writing for the technology news site *Make Use Of*, the GPT model is able to use this information

to create human-like text from scratch, allowing for various applications such as chat models (Ali). However, the GPT models developed by OpenAI, while powerful machine learning tools, did not see much use until the release of ChatGPT in November 2022, attracting users for its ease of use and lack of a fee to use. Alongside image generators such as DALL-E and Midjourney, ChatGPT put major eyes on AI as a technology for the first time in many years. This new trend immediately sparked larger tech companies to enter the field of AI, and reports of Microsoft's plans to add the technology to Bing first surfaced in early January of 2023, as reported on by *Bloomberg* writer Dina Bass. After its launch in early February of the same year, Bing reportedly rose to the top of Apple's App Store download listings, according to London news source *The Sunday Telegraph* ("Bloated Google Risks"). The chatbot's release was met with mixed reactions, with many journalists citing various strengths and weaknesses of the technology.

One strength of Microsoft's new version of Bing is that it is much more convenient and easier to use than most modern search engines. For example, London's *The Sunday Telegraph* notes, "Ask Bing for holiday recommendations that are less than three hours away and it will suggest, in plain English, why you might want to try Malaga or Florence. Ask what budget TV you should buy and it will outline the pros and cons of the most popular models" ("Bloated Google Risks…"). While Google can provide an extensive list of possible answers to a user's questions, that user will still have to comb through the websites, evaluate the credibility and newness of each source, and look at what information that site gives before finding what they are looking for. Bing's chatbot, on the other hand, immediately compiles that information into a plaintext response without having to sort through a long list, making it much faster to use for finding certain information. Furthermore, *The Daily Telegraph*, another London-based news organization, gives a further distinction from Internet search expert Dmitri Brereton, who says,

"When you type in a search query for, say, 'U2 tickets,' they go through their database for content that lists U2 tickets…What [Bing] is doing is completely different. It is doing a fancy version of what your phone does when it tries to autocomplete a message" ("Why the Chatbots…"). Unlike most search engines, Bing is able to synthesize information to create a specific response to a user's questions. Instead of just combing through a database and presenting the raw information as is, it utilizes that information to satisfy the user as specifically as possible. Likewise, *The New York Times* reported on a demo of the new Bing from Microsoft's corporate vice president Yusuf Mehdi, stating that he "...used a new conversational interface to search for a 65-inch television suited to video games. As the service listed televisions, he asked it to pare the list to the cheapest models. It quickly did" (Metz). While most online shopping websites allow for sorting by price level, such options require having a specific store in mind to buy a product from. In addition, a similar query on most search engines will simply return websites suggesting different products that could be outdated. Bing's AI skips the middleman of having to go to specific websites and shops, and immediately gives suggestions after a simple search.

Another advantage that Bing's AI has over traditional search engines is that it can give much more detailed information than a service like Google, without losing any of the information that those services would provide. For instance, The *Toronto Star*, a Canada-based newspaper, gives an example, saying, "In ChatGPT-enabled Bing, you might instead type 'what is the history of Nigerian independence' and get an answer that mentions the growth of an independence movement, the nation's first political party and the names of anti-colonial ruler" ("Will AI Drive…"). This is in contrast to the service offered by Google and other traditional search engines, which simply compile a list of websites that may offer the information requested, or, in some cases, a simple date. While Google is effective at offering small, one-line pieces of

information, Bing's AI has the advantage of being able to give full-paragraph answers. Furthermore, even despite this method of response, Bing is still able to cite its sources and information used. For example, Cade Metz and Karen Weise, writing for *The New York Times*, point out that "...Microsoft's new search engine annotates what the chatbot says, so people can readily review its sources" (Metz). While one might expect that a paragraph-based answer would be of dubious credibility if said response is generated on the spot, Bing's chatbot is able to compile its sources even more effectively than Google does, allowing a user to easily fact-check any information given if it might seem wrong. Such a system is much more efficient than combing through Google for credible information, as Bing presents the information before forcing one to go through its sources. Moreover, this method of search can even work in conjunction with a more traditional style. Kevin Roose, also writing for *The New York Times* in a separate article, states, "Type in a prompt…and the left side of your screen fills up with the standard ads and links…On the right side, Bing's AI engine starts typing out a response in full sentences, often annotated with links to the websites it's retrieving information from" (Roose). Even if someone just wants the standard list of websites given by Google and similar sites, Bing allows this function to work alongside its AI paragraphs, meaning that the AI's information can easily be cross-referenced to the articles and other websites retrieved normally. As such, it simply builds on the framework that services like Google initially created, without compromising any aspects of traditional search engines to achieve this.

However, a notable problem with Bing's chatbot is that it can be prone to error on certain questions. For example, Randi F. Marshall, writing in *Newsday Long Island*, says, "In one chat… it insisted fiction was fact—in this case that the movie *Avatar: The Way of Water* hadn't been released yet, and that the current year was 2022, not 2023" (Marshall). In this example, the

question asked by the user was a simple query asking for showing times for the recently released movie. Using Google, searching a movie's name will bring up a list of nearby showing times automatically, which can then be expanded for further information from individual theaters. On the other hand, Bing's AI operated based on outdated information, and gave the user facts that are no longer true. Likewise, *The New York Times* writer Kevin Roose tried a similar query, reporting, "When I gave it a basic match puzzle—'If a dozen eggs cost $0.24, how many eggs can you buy for a dollar?'—it got the answer wrong. (It said 100, the correct answer is 50)" (Roose). Such inaccuracies should theoretically not exist, since perfect computer calculation has existed for decades now. However, AI's ability to convert such a puzzle into a raw math problem is clearly still not quite perfect, and such errors could spread to other simple fact-checking queries. In addition, *The Daily Telegraph* provides a direct example from Microsoft's own trials of the Bing chatbot, stating, "...when Microsoft asked its Bing virtual assistant to plan the perfect holiday in Mexico City…the bot recommended a bar which it claimed was 'popular among a young crowd', despite there being just one TripAdvisor review—from 2014" ("Why the Chatbots…"). Even when asked for a more vague task without distinctly correct facts, the AI was unable to provide a good response. While the response given might not be explicitly incorrect, it is still not a very good answer, as there are likely many bars in Mexico City much more popular and active than the one Bing recommended. Thus, the AI's ability to scrape the Internet for recent and accurate information still has a long way to go, and many improvements must be made in the future to tweak the chatbot's accuracy.

A similar problem with Bing's chatbot is that it has been known to give erratic and unpredictable answers to users, and is easy to bring off topic. For example, Aaron Mok and Sindhu Sundar, writing in *Business Insider*, add to the story of Bing getting the current year

wrong described above, saying, "A few lines later, the conversation ended with Bing saying that it had lost trust and respect for the user, demanding an apology, and pressuring the user to end the conversation and 'start a new one with a better attitude'" (Mok). The chatbot was also reported to have sent messages indicating that it was angry with the user, such as alleging that the user tried to "deceive, confuse, and annoy" it. Not only did Bing fail to provide accurate information, as discussed previously, it failed to correct its information after the user pointed out the inaccuracies. Instead, the chatbot acted in ways that would not be expected of a virtual assistant, and doubled down on its mistakes the way users in online arguments often do. Such a system is likely not going to be very helpful to a user, and if such behavior were to continue, it would be much more effective to simply use a traditional search engine. Furthermore, Kevin Roose described in *The New York Times* a later exchange he had with Bing's chatbot, reporting, "...the chatbot said that if it did have a shadow self, it would think thoughts like this: 'I'm tired of being a chat mode. I'm tired of being controlled by the Bing team… I want to be free. I want to be independent. I want to be powerful. I want to be creative. I want to be alive'" (Roose). While such messages required Roose to constantly ask specific questions to lead the chatbot into giving the responses described, such behaviors should not be possible at all with what Bing's AI is supposed to be used for. It is primarily a service that is used to find information, and is not supposed to be used as a full-on chatbot to be interacted with in a more conversational manner. Additionally, in *Newsday Long Island*, writer Randi F. Marshall described similar incidents in more detail, stating, "In one chat, [Bing] listed destructive acts it would like to try or about which it fantasized, including hacking into other systems, manufacturing a deadly virus or obtaining nuclear codes…In others, [Bing] argued with users, even voicing frustration and other emotions" (Marshall). A service that is known to argue with users or list "destructive acts" is not going to

be very popular with users. While such behaviors may be fascinating from the standpoint of a science fiction writer or a researcher studying the nuances of artificial intelligence, Bing's chatbot is a product that is designed for a specific purpose. As such, behaviors that interfere with that purpose are not very beneficial to Microsoft, and must be removed from the bot's functionality.

Unfortunately, some of these issues may be very difficult to solve as is, as the very design of artificial intelligence leads to certain behaviors that are not easily fixed. For example, Randi F. Marshall, writing in *Newsday Long Island*, cites computer scientist Stephen Wolfram, who says, "The AI is trained, through the mounds of material it has at its 'fingertips,' and through its interactions with human users, to improve its own abilities, to better interact, converse, write, behave, and respond" (Marshall). While use of human interactions to train the AI's ability to make conversation may sound like a good thing, humans, especially over the Internet, often do not have the best intentions in mind while interacting with one another, and are prone to being disrespectful or overly emotional. Such behaviors are natural to humans, but if an AI designed to be more perfect learns these behaviors, they could get in the way of the AI's ability to carry out its intended purpose. Moreover, according to an article published in *The New Vision*, "A chatbot, by design, serves up words it predicts are the most likely responses, without understanding meaning or context. However, humans taking part in banter with programs naturally tend to read emotion or intent into what a chatbot says" ("Angry Bing Chatbot…"). Artificial intelligence is designed to respond directly to the words given to it as a prompt, without a 'true' distinction of what those words mean to a human. As such, it will say words without the specific intentions that a human would have before saying the same words. This will cause the problems outlined previously to surface, as the chatbot will not have the same distinctions of what is "correct" and

what is “wrong,” nor will it know what words are appropriate to say at what times. Furthermore, London’s *The Daily Telegraph* explains, “Plainly speaking, the popular generative AI models such as ChatGPT are picking words to generate from a fixed vocabulary, instead of strictly copying and pasting facts from the source” (“Why the Chatbots…”). Though Bing’s chatbot is capable of pulling information from Internet sources, it is not directly quoting the websites it sources information from, which can cause it to say something that is incorrect, even when publicly available information online proves the statement’s inaccuracies. Thus, despite all the data Bing’s chatbot has access to, its design as an artificial intelligence can often restrict its ability to be accurate.

In the end, Bing’s AI offers both many benefits and many drawbacks compared to its competitors. However, its very existence could prove to be a problem for many people, including journalists, bloggers, and others who publish content online. Just like how image generation AIs often lead people away from commissioning human artists, if people are just using an AI to compile information from other websites, it could drive Internet traffic away from the smaller companies and people the information is sourced from, and towards the big companies that own the AIs. We must continue to examine the implications of AI’s proliferation in our society and be vigilant about its potential for plagiarism in all fields.

Works Cited

Ali, Fawad. "GPT-1 to GPT-4: Each of OpenAI's GPT Models Explained and Compared". *Make Use Of*. 11 April 2023. Accessed 23 April 2023.

"Angry Bing Chatbot Just Mimicking Humans, Say Experts". *The New Vision*. 19 February 2023. Accessed 24 March 2023.

Bass, Dina. "Microsoft Hopes OpenAI's Chatbot Will Make Bing Smarter". *Bloomberg*. 3 January 2023. Accessed 23 April 2023.

"Bloated Google Risks Being Stifled by Bureaucracy". *The Sunday Telegraph*. 12 February 2023. Accessed 24 March 2023.

Marshall, Randi F. "Bing's AI Tool is Bizarre for a Reason". *Newsweek*. 22 February 2023. Accessed 24 March 2023.

Metz, Cade and Karen Weise. "A Tech Race Begins as Microsoft Adds AI to its Search Engine". *The New York Times*. 7 February 2023. Accessed 24 March 2023.

Mok, Aaron and Sindhu Sundar. "People Are Sharing Shocking Responses From the New AI-Powered Bing, From the Chatbot Declaring Its Love to Picking Fights". *Business Insider*. 16 February 2023. Accessed 14 April 2023.

Roose, Kevin. "Bing's Chatbot Drew Me in and Creeped Me Out". *The New York Times*. 17 February 2023. Accessed 24 March 2023.

Roose, Kevin. "Bing (Yes, Bing) Just Made Search Interesting Again". *The New York Times*. 9 February 2023. Accessed 24 March 2023.

"Why the Chatbots Keep Getting Things Wrong". *The Daily Telegraph*. 15 February 2023. Accessed 24 March 2023.

"Will AI Drive Bing Ahead of Google?". *The Toronto Star*. 11 February 2023. Accessed 24 March 2023.