

In A.I. Race, Microsoft and Google Choose Speed Over Caution

Technology companies were once leery of what some artificial intelligence could do. Now the priority is winning control of the industry's next big thing.



By Nico Grant and Karen Weise

Nico Grant reported this story from San Francisco, and Karen Weise reported from Seattle.

Published April 7, 2023 Updated April 10, 2023

In March, two Google employees, whose jobs are to review the company's artificial intelligence products, tried to stop Google from launching an A.I. chatbot. They believed it generated inaccurate and dangerous statements.

Ten months earlier, similar concerns were raised at Microsoft by ethicists and other employees. They wrote in several documents that the A.I. technology behind a planned chatbot could flood Facebook groups with disinformation, degrade critical thinking and erode the factual foundation of modern society.

The companies released their chatbots anyway. Microsoft was first, with a splashy event in February to reveal an A.I. chatbot woven into its Bing search engine. Google followed about six weeks later with its own chatbot, Bard.

The aggressive moves by the normally risk-averse companies were driven by a race to control what could be the tech industry's next big thing — generative A.I., the powerful new technology that fuels those chatbots.

That competition took on a frantic tone in November when OpenAI, a San Francisco start-up working with Microsoft, released ChatGPT, a chatbot that has captured the public imagination and now has an estimated 100 million monthly users.

The surprising success of ChatGPT has led to a willingness at Microsoft and Google to take greater risks with their ethical guidelines set up over the years to ensure their technology does not cause societal problems, according to 15 current and former employees and internal documents from the companies.

The urgency to build with the new A.I. was crystallized in an internal email sent last month by Sam Schillace, a technology executive at Microsoft. He wrote in the email, which was viewed by The New York Times, that it was an “absolutely fatal error in this moment to worry about things that can be fixed later.”

When the tech industry is suddenly shifting toward a new kind of technology, the first company to introduce a product “is the long-term winner just because they got started first,” he wrote. “Sometimes the difference is measured in weeks.”

Last week, tension between the industry's worriers and risk-takers played out publicly as more than 1,000 researchers and industry leaders, including Elon Musk and Apple's co-founder Steve Wozniak, called for a six-month pause in the development of powerful A.I. technology. In a public letter, they said it presented "profound risks to society and humanity."

Regulators are already threatening to intervene. The European Union proposed legislation to regulate A.I., and Italy temporarily banned ChatGPT last week. In the United States, President Biden on Tuesday became the latest official to question the safety of A.I.

"Tech companies have a responsibility to make sure their products are safe before making them public," he said at the White House. When asked if A.I. was dangerous, he said: "It remains to be seen. Could be."

The issues being raised now were once the kinds of concerns that prompted some companies to sit on new technology. They had learned that prematurely releasing A.I. could be embarrassing. Seven years ago, for example, Microsoft quickly pulled a chatbot called Tay after users nudged it to generate racist responses.

Researchers say Microsoft and Google are taking risks by releasing technology that even its developers don't entirely understand. But the companies said that they had limited the scope of the initial release of their new chatbots, and that they had built sophisticated filtering systems to weed out hate speech and content that could cause obvious harm.

Natasha Crampton, Microsoft's chief responsible A.I. officer, said in an interview that six years of work around A.I. and ethics at Microsoft had allowed the company to "move nimbly and thoughtfully." She added that "our commitment to responsible A.I. remains steadfast."

Google released Bard after years of internal dissent over whether generative A.I.'s benefits outweighed the risks. It announced Meena, a similar chatbot, in 2020. But that system was deemed too risky to release, three people with knowledge of the process said. Those concerns were reported earlier by The Wall Street Journal.



Margaret Mitchell, a former Google employee, was blocked from publishing a paper with a former colleague that warned of the risks of newer A.I. technology. Ruth Fremson/The New York Times

Later in 2020, Google blocked its top ethical A.I. researchers, Timnit Gebru and Margaret Mitchell, from publishing a paper warning that so-called large language models used in the new A.I. systems, which are trained to recognize patterns from vast amounts of data, could spew abusive or discriminatory language. The researchers were pushed out after Dr. Gebru criticized the company's diversity efforts and Dr. Mitchell was accused of violating its code of conduct after she saved some work emails to a personal Google Drive account.

Dr. Mitchell said she had tried to help Google release products responsibly and avoid regulation, but instead "they really shot themselves in the foot."

Brian Gabriel, a Google spokesman, said in a statement that "we continue to make responsible A.I. a top priority, using our A.I. principles and internal governance structures to responsibly share A.I. advances with our users."

Concerns over larger models persisted. In January 2022, Google refused to allow another researcher, El Mahdi El Mhamdi, to publish a critical paper.

Dr. El Mhamdi, a part-time employee and university professor, used mathematical theorems to warn that the biggest A.I. models are more vulnerable to cybersecurity attacks and present unusual privacy risks because they've probably had access to private data stored in various locations around the

internet.

Though an executive presentation later warned of similar A.I. privacy violations, Google reviewers asked Dr. El Mhamdi for substantial changes. He refused and released the paper through École Polytechnique.

He resigned from Google this year, citing in part “research censorship.” He said modern A.I.’s risks “highly exceeded” the benefits. “It’s premature deployment,” he added.

After ChatGPT’s release, Kent Walker, Google’s top lawyer, met with research and safety executives on the company’s powerful Advanced Technology Review Council. He told them that Sundar Pichai, Google’s chief executive, was pushing hard to release Google’s A.I.

Jen Gennai, the director of Google’s Responsible Innovation group, attended that meeting. She recalled what Mr. Walker had said to her own staff.

The meeting was “Kent talking at the A.T.R.C. execs, telling them, ‘This is the company priority;’” Ms. Gennai said in a recording that was reviewed by The Times. “‘What are your concerns? Let’s get in line.’”

Mr. Walker told attendees to fast-track A.I. projects, though some executives said they would maintain safety standards, Ms. Gennai said.

Her team had already documented concerns with chatbots: They could produce false information, hurt users who become emotionally attached to them and enable “tech-facilitated violence” through mass harassment online.

In March, two reviewers from Ms. Gennai’s team submitted their risk evaluation of Bard. They recommended blocking its imminent release, two people familiar with the process said. Despite safeguards, they believed the chatbot was not ready.

Ms. Gennai changed that document. She took out the recommendation and downplayed the severity of Bard’s risks, the people said.

Ms. Gennai said in an email to The Times that because Bard was an experiment, reviewers were not supposed to weigh in on whether to proceed. She said she “corrected inaccurate assumptions, and actually added more risks and harms that needed consideration.”

Google said it had released Bard as a limited experiment because of those debates, and Ms. Gennai said continuing training, guardrails and disclaimers made the chatbot safer.

Google released Bard to some users on March 21. The company said it would soon integrate generative A.I. into its search engine.



Satya Nadella, Microsoft's chief executive, made a big bet on A.I. several years ago. Ruth Fremson/The New York Times

Satya Nadella, Microsoft's chief executive, made a bet on generative A.I. in 2019 when Microsoft invested \$1 billion in OpenAI. After deciding the technology was ready over the summer, Mr. Nadella pushed every Microsoft product team to adopt A.I.

Microsoft had policies developed by its Office of Responsible A.I., a team run by Ms. Crampton, but the guidelines were not consistently enforced or followed, said five current and former employees.

Despite having a "transparency" principle, ethics experts working on the chatbot were not given answers about what data OpenAI used to develop its systems, according to three people involved in the work. Some argued that integrating chatbots into a search engine was a particularly bad idea, given how it sometimes served up untrue details, a person with direct knowledge of the conversations said.

Ms. Crampton said experts across Microsoft worked on Bing, and key people had access to the training data. The company worked to make the chatbot more accurate by linking it to Bing search results, she added.

In the fall, Microsoft started breaking up what had been one of its largest technology ethics teams. The group, Ethics and Society, trained and consulted company product leaders to design and build responsibly. In October, most of its members were spun off to other groups, according to four people familiar with the team.

The remaining few joined daily meetings with the Bing team, racing to launch the chatbot. John Montgomery, an A.I. executive, told them in a December email that their work remained vital and that more teams “will also need our help.”

After the A.I.-powered Bing was introduced, the ethics team documented lingering concerns. Users could become too dependent on the tool. Inaccurate answers could mislead users. People could believe the chatbot, which uses an “I” and emojis, was human.

In mid-March, the team was laid off, an action that was first reported by the tech newsletter Platformer. But Ms. Crampton said hundreds of employees were still working on ethics efforts.

Microsoft has released new products every week, a frantic pace to fulfill plans that Mr. Nadella set in motion in the summer when he previewed OpenAI’s newest model.

He asked the chatbot to translate the Persian poet Rumi into Urdu, and then write it out in English characters. “It worked like a charm,” he said in a February interview. “Then I said, ‘God, this thing.’”

Audio produced by Parin Behrooz.

Mike Isaac contributed reporting. Susan C. Beachy contributed research.

A correction was made on April 9, 2023: An earlier version of this article misstated the number of years that have passed since Microsoft shut down a chatbot called Tay. It was seven years ago, not five.

When we learn of a mistake, we acknowledge it with a correction. If you spot an error, please let us know at nytnews@nytimes.com. [Learn more](#)

Nico Grant is a technology reporter covering Google from San Francisco. Previously, he spent five years at Bloomberg News, where he focused on Google and cloud computing. [More about Nico Grant](#)

Karen Weise is a technology correspondent based in Seattle, covering Amazon and Microsoft. Her work aims to help readers better understand two of the most powerful companies in America and their growing influence on society. [More about Karen Weise](#)

A version of this article appears in print on , Section A, Page 1 of the New York edition with the headline: A.I. Race Leads Tech Giants to Take Ethics Risks