# *Ego, Fear and Money: How the A.I. Fuse Was Lit*

The people who were most afraid of the risks of artificial intelligence decided they should be the ones to build it. Then distrust fueled a spiraling competition.

Larry Page and Elon Musk were on opposite sides in the debate over the risks of artificial intelligence. Hokyoung Kim

**By Cade Metz, Karen Weise, Nico Grant and Mike Isaac**
Reporting from San Francisco

Dec. 3, 2023

Elon Musk celebrated his 44th birthday in July 2015 at a three-day party thrown by his wife at a California wine country resort dotted with cabins. It was family and friends only, with children racing around the upscale property in Napa Valley.

This was years before Twitter became X and Tesla had a profitable year. Mr. Musk and his wife, Talulah Riley — an actress who played a beautiful but dangerous robot on HBO's science fiction series "Westworld" — were a year from throwing in the towel on their second marriage. Larry Page, a party guest, was still the chief executive of Google. And artificial intelligence had pierced the public consciousness only a few years before, when it was used to identify cats on YouTube — with 16 percent accuracy.

A.I. was the big topic of conversation when Mr. Musk and Mr. Page sat down near a firepit beside a swimming pool after dinner the first night. The two billionaires had been friends for more than a decade, and Mr. Musk sometimes joked that he occasionally crashed on Mr. Page's sofa after a night playing video games.

But the tone that clear night soon turned contentious as the two debated whether artificial intelligence would ultimately elevate humanity or destroy it.

As the discussion stretched into the chilly hours, it grew intense, and some of the more than 30 partyers gathered closer to listen. Mr. Page, hampered for more than a decade by an unusual ailment in his vocal cords, described his vision of a digital utopia in a whisper. Humans would eventually merge with artificially intelligent machines, he said. One day there would be many kinds of intelligence competing for resources, and the best would win.

If that happens, Mr. Musk said, we're doomed. The machines will destroy humanity.

With a rasp of frustration, Mr. Page insisted his utopia should be pursued. Finally he called Mr. Musk a "specieist," a person who favors humans over the digital life-forms of the future.

That insult, Mr. Musk said later, was "the last straw."

Many in the crowd seemed gobsmacked, if amused, as they dispersed for the night, and considered it just another one of those esoteric debates that often break out at Silicon Valley parties.

But eight years later, the argument between the two men seems prescient. The question of whether artificial intelligence will elevate the world or destroy it — or at least inflict grave damage — has framed an ongoing debate among Silicon Valley founders, chatbot users, academics, legislators and regulators about whether the technology should be controlled or set free.

That debate has pitted some of the world's richest men against one another: Mr. Musk, Mr. Page, Mark Zuckerberg of Meta, the tech investor Peter Thiel, Satya Nadella of Microsoft and Sam Altman of OpenAI. All have fought for a piece of the business — which one day could be worth trillions of dollars — and the power to shape it.

At the heart of this competition is a brain-stretching paradox. The people who say they are most worried about A.I. are among the most determined to create it and enjoy its riches. They have justified their ambition with their strong belief that they alone can keep A.I. from endangering Earth.

Mr. Musk and Mr. Page stopped speaking soon after the party that summer. A few weeks later, Mr. Musk dined with Mr. Altman, who was then running a tech incubator, and several researchers in a private room at the Rosewood hotel in Menlo Park, Calif., a favored deal-making spot close to the venture capital offices of Sand Hill Road.

That dinner led to the creation of a start-up called OpenAI later in the year. Backed by hundreds of millions of dollars from Mr. Musk and other funders, the lab promised to protect the world from Mr. Page's vision.

Thanks to its ChatGPT chatbot, OpenAI has fundamentally changed the technology industry and has introduced the world to the risks and potential of artificial intelligence. OpenAI is valued at more than $80 billion, according to two people familiar with the company's latest funding round, though Mr. Musk and Mr. Altman's partnership didn't make it. The two have since stopped speaking.

"There is disagreement, mistrust, egos," Mr. Altman said. "The closer people are to being pointed in the same direction, the more contentious the disagreements are. You see this in sects and religious orders. There are bitter fights between the closest people."

Last month, that infighting came to OpenAI's boardroom. Rebel board members tried to force out Mr. Altman because, they believed, they could no longer trust him to build A.I. that would benefit humanity. Over five chaotic days OpenAI looked as if it were going to fall apart, until the board — pressured by giant investors and employees who threatened to follow Mr. Altman out the door — backed down.

The drama inside OpenAI gave the world its first glimpse of the bitter feuds among those who will determine the future of A.I.

But years before OpenAI's near meltdown, there was a little-publicized but ferocious competition in Silicon Valley for control of the technology that is now quickly reshaping the world, from how children are taught to how wars are fought. The New York Times spoke with more than 80 executives, scientists and entrepreneurs, including two people who attended Mr. Musk's birthday party in 2015, to tell that story of ambition, fear and money.

## The Birth of DeepMind

Five years before the Napa Valley party and two before the cat breakthrough on YouTube, Demis Hassabis, a 34-year-old neuroscientist, walked into a cocktail party at Peter Thiel's San Francisco townhouse and realized he'd hit pay dirt. There in Mr. Thiel's living room, overlooking the city's Palace of Fine Arts and a swan pond, was a chess board. Dr. Hassabis had once been the second-best player in the world in the under-14 category.

"I was preparing for that meeting for a year," Dr. Hassabis said. "I thought that would be my unique hook in: I knew that he loved chess."

In 2010, Dr. Hassabis and two colleagues, who all lived in Britain, were looking for money to start building "artificial general intelligence," or A.G.I., a machine that could do anything the brain could do. At the time, few people were interested in A.I. After a half century of research, the artificial intelligence field had failed to deliver anything remotely close to the human brain.

Still, some scientists and thinkers had become fixated on the downsides of A.I. Many, like the three young men from Britain, had a connection to Eliezer Yudkowsky, an internet philosopher and self-taught A.I. researcher. Mr. Yudkowsky was a leader in a community of people who called themselves Rationalists or, in later years, effective altruists.

They believed that A.I. could find a cure for cancer or solve climate change, but they worried that A.I. bots might do things their creators had not intended. If the machines became more intelligent than humans, the Rationalists argued, the machines could turn on their creators.

Mr. Thiel had become enormously wealthy through an early investment in Facebook and through his work with Mr. Musk in the early days of PayPal. He had developed a fascination with the singularity, a trope of science fiction that describes the moment when intelligent technology can no longer be controlled by humanity.

With funding from Mr. Thiel, Mr. Yudkowsky had expanded his A.I. lab and created an annual conference on the singularity. Years before, one of Dr. Hassabis's two colleagues had met Mr. Yudkowsky, and he snagged them speaking spots at the conference, ensuring they'd be invited to Mr. Thiel's party.

Mr. Yudkowsky introduced Dr. Hassabis to Mr. Thiel. Dr. Hassabis assumed that lots of people at the party would be trying to squeeze their host for money. His strategy was to arrange another meeting. There was a deep tension between the bishop and the knight, he told Mr. Thiel. The two pieces carried the same value, but the best players understood that their strengths were

vastly different.

Hokyoung Kim

It worked. Charmed, Mr. Thiel invited the group back the next day, where they gathered in the kitchen. Their host had just finished his morning workout and was still sweating in a shiny tracksuit. A butler handed him a Diet Coke. The three made their pitch, and soon Mr. Thiel and his venture capital firm agreed to put 1.4 million British pounds (roughly $2.25 million) into their start-up. He was their first major investor.

They named their company DeepMind, a nod to "deep learning," a way for A.I. systems to learn skills by analyzing large amounts of data; to neuroscience; and to the Deep Thought supercomputer from the sci-fi novel "The Hitchhiker's Guide to the Galaxy." By the fall of 2010, they were building their dream machine. They wholeheartedly believed that because they understood the risks, they were uniquely positioned to protect the world.

"I don't see this as a contradictory position," said Mustafa Suleyman, one of the three DeepMind founders. "There are huge benefits to come from these technologies. The goal is not to eliminate them or pause their development. The goal is to mitigate the downsides."

Having won over Mr. Thiel, Dr. Hassabis worked his way into Mr. Musk's orbit. About two years later, they met at a conference organized by Mr. Thiel's investment fund, which had also put money into Mr. Musk's company SpaceX. Dr. Hassabis secured a tour of SpaceX headquarters. Afterward, with rocket hulls hanging from the ceiling, the two men lunched in the cafeteria and talked.

Mr. Musk explained that his plan was to colonize Mars to escape overpopulation and other dangers on Earth. Dr. Hassabis replied that the plan would work — so long as superintelligent machines didn't follow and destroy humanity on Mars, too.

Mr. Musk was speechless. He hadn't thought about that particular danger. Mr. Musk soon invested in DeepMind alongside Mr. Thiel so he could be closer to the creation of this technology.

Flush with cash, DeepMind hired researchers who specialized in neural networks, complex algorithms created in the image of the human brain. A neural network is essentially a giant mathematical system that spends days, weeks or even months identifying patterns in large amounts of digital data. First developed in the 1950s, these systems could learn to handle tasks on their own. After analyzing names and addresses scribbled on hundreds of envelopes, for instance, they could read handwritten text.

DeepMind took the concept further. It built a system that could learn to play classic Atari games like Space Invaders, Pong and Breakout to illustrate what was possible.

This got the attention of another Silicon Valley powerhouse, Google, and specifically Larry Page. He saw a demonstration of Deep Mind's machine playing Atari games. He wanted in.

## The Talent Auction

In the fall of 2012, Geoffrey Hinton, a 64-year-old professor at the University of Toronto, and two graduate students published a research paper that showed the world what A.I. could do. They trained a neural network to recognize common objects like flowers, dogs and cars.

Scientists were surprised by the accuracy of the technology built by Dr. Hinton and his students. One who took particular notice was Yu Kai, an A.I. researcher who had met Dr. Hinton at a research conference and had recently started working for Baidu, the giant Chinese internet company. Baidu offered Dr. Hinton and his students $12 million to join the company in Beijing, according to three people familiar with the offer.

Dr. Hinton turned Baidu down, but the money got his attention.

The Cambridge-educated British expatriate had spent most of his career in academia, except for occasional stints at Microsoft and Google, and was not especially driven by money. But he had a neurodivergent child, and the money would mean financial security.

"We did not know how much we were worth," Dr. Hinton said. He consulted lawyers and experts on acquisitions and came up with a plan: "We would organize an auction, and we would sell ourselves." The auction would take place during an annual A.I. conference at the Harrah's hotel and casino on Lake Tahoe.

Big Tech took notice. Google, Microsoft, Baidu and other companies were beginning to believe that neural networks were a path to machines that could not only see, but hear, write, talk and — eventually — think.

Mr. Page had seen similar technology at Google Brain, his company's A.I. lab, and he thought Dr. Hinton's research could elevate his scientists' work. He gave Alan Eustace, Google's senior vice president of engineering, what amounted to a blank check to hire any A.I. expertise he needed.

Mr. Eustace and Jeff Dean, who led the Brain lab, flew to Lake Tahoe and took Dr. Hinton and his students out to dinner at a steakhouse inside the hotel the night before the auction. The smell of old cigarettes was overpowering, Dr. Dean recalled. They made the case for coming to work at Google.

The next day, Dr. Hinton ran the auction from his hotel room. Because of an old back injury, he rarely sat down. He turned a trash can upside down on a table, put his laptop on top and watched the bids roll in over the next two days.

Google made an offer. So did Microsoft. DeepMind quickly bowed out as the price went up. The industry giants pushed the bids to $20 million and then $25 million, according to documents detailing the auction. As the price passed $30 million, Microsoft quit, but it rejoined the bidding at $37 million.

"We felt like we were in a movie," Dr. Hinton said.

Then Microsoft dropped out a second time. Only Baidu and Google were left, and they pushed the bidding to $42 million, $43 million. Finally, at $44 million, Dr. Hinton and his students stopped the auction. The bids were still climbing, but they wanted to work for Google. And the money was staggering.

It was an unmistakable sign that deep-pocketed companies were determined to buy the most talented A.I. researchers — which was not lost on Dr. Hassabis at DeepMind. He had always told his employees that DeepMind would remain an independent company. That was, he believed, the best way to ensure its technology didn't turn into something dangerous.

But as Big Tech entered the talent race, he decided he had no choice: It was time to sell.

By the end of 2012, Google and Facebook were angling to acquire the London lab, according to three people familiar with the matter. Dr. Hassabis and his co-founders insisted on two conditions: No DeepMind technology could be used for military purposes, and its A.G.I. technology must be overseen by an independent board of technologists and ethicists.

Google offered $650 million. Mark Zuckerberg of Facebook offered a bigger payout to DeepMind's founders, but would not agree to the conditions. DeepMind sold to Google.

Mr. Zuckerberg was determined to build an A.I. lab of his own. He hired Yann LeCun, a French computer scientist who had also done pioneering A.I. research, to run it. A year after Dr. Hinton's auction, Mr. Zuckerberg and Dr. LeCun flew to Lake Tahoe for the same A.I. conference. While padding around a suite at the Harrah's casino in his socks, Mr. Zuckerberg personally interviewed top researchers, who were soon offered millions of dollars in salary and stock.

A.I. was once laughed off. Now the richest men in Silicon Valley were shelling out billions to keep from being left behind.

## The Lost Ethics Board

When Mr. Musk invested in DeepMind, he broke his own informal rule — that he would not invest in any company he didn't run himself. The downsides of his decision were already apparent when, only a month or so after his birthday spat with Mr. Page, he again found himself face to face with his former friend and fellow billionaire.

The occasion was the first meeting of DeepMind's ethics board, on Aug. 14, 2015. The board had been set up at the insistence of the start-up's founders to ensure that their technology did no harm after the sale. The members convened in a conference room just outside Mr. Musk's office at SpaceX, with a window looking out onto his rocket factory, according to three people familiar with the meeting.

But that's where Mr. Musk's control ended. When Google bought DeepMind, it bought the whole thing. Mr. Musk was out. Financially he had come out ahead, but he was unhappy.

Three Google executives now firmly in control of DeepMind were there: Mr. Page; Sergey Brin, a Google co-founder and Tesla investor; and Eric Schmidt, Google's chairman. Among the other attendees were Reid Hoffman, another PayPal founder, and Toby Ord, an Australian philosopher studying "existential risk."

The DeepMind founders reported that they were pushing ahead with their work, but that they were aware the technology carried serious risks.

Hokyoung Kim

Mr. Suleyman, the DeepMind co-founder, gave a presentation called "The Pitchforkers Are Coming." A.I. could lead to an explosion in disinformation, he told the board. He fretted that as the technology replaced countless jobs in the coming years, the public would accuse Google of stealing their livelihoods. Google would need to share its wealth with the millions who could no longer find work and provide a "universal basic income," he argued.

Mr. Musk agreed. But it was pretty clear that his Google guests were not prepared to embark on a redistribution of (their) wealth. Mr. Schmidt said he thought the worries were completely overblown. In his usual whisper, Mr. Page agreed. A.I. would create more jobs than it took away, he argued.

Eight months later, DeepMind had a breakthrough that stunned the A.I community and the world. A DeepMind machine called AlphaGo beat one of the world's best players at the ancient game of Go. The game, streamed over the internet, was watched by 200 million people across the globe. Most researchers had assumed that A.I. needed another 10 years to muster the ingenuity to do that.

Rationalists, effective altruists and others who worried about the risks of A.I. claimed the computer's win validated their fears.

"This is another indication that A.I. is progressing faster than even many experts anticipated," Victoria Krakovna, who would soon join DeepMind as an "A.I. safety" researcher, wrote in a blog post.

DeepMind's founders were increasingly worried about what Google would do with their inventions. In 2017, they tried to break away from the company. Google responded by increasing the salaries and stock award packages of the DeepMind founders and their staff. They stayed put.

The ethics board never had a second meeting.

## The Breakup

Convinced that Mr. Page's optimistic view of A.I. was dead wrong, and angry at his loss of DeepMind, Mr. Musk built his own lab.

OpenAI was founded in late 2015, just a few months after he met with Sam Altman at the Rosewood hotel in Silicon Valley. Mr. Musk pumped money into the lab, and his former PayPal buddies, Mr. Hoffman and Mr. Thiel, came along for the ride. The three men and others pledged to put $1 billion into the project, which Mr. Altman, who was 30 at the time, would help run. To get them started, they poached Ilya Sutskever from Google. (Dr. Sutskever was one of the graduate students Google "bought" in Dr. Hinton's auction.)

Initially, Mr. Musk wanted to operate OpenAI as a nonprofit, free from the economic incentives that were driving Google and other corporations. But by the time Google wowed the tech community with its Go stunt, Mr. Musk was changing his mind about how it should be run. He desperately wanted OpenAI to invent something that would capture the world's imagination and close the gap with Google, but it wasn't getting the job done as a nonprofit.

In late 2017, he hatched a plan to wrest control of the lab from Mr. Altman and the other founders and transform it into a commercial operation that would join forces with Tesla and rely on supercomputers the car company was developing, according to four people familiar with the matter.

When Mr. Altman and others pushed back, Mr. Musk quit and said he would focus on his own A.I. work at Tesla. In February 2018, he announced his departure to OpenAI's staff on the top floor of the start-up's offices in a converted truck factory, three people who attended the meeting said. When he said that OpenAI needed to move faster, one researcher retorted at the meeting that Mr. Musk was being reckless.

Mr. Musk called the researcher a "jackass" and stormed out, taking his deep pockets with him.

OpenAI suddenly needed new financing in a hurry. Mr. Altman flew to Sun Valley for a conference and ran into Satya Nadella, Microsoft's chief executive. A tie-up seemed natural. Mr. Altman knew Microsoft's chief technology officer, Kevin Scott. Microsoft had bought LinkedIn from Mr. Hoffman, an OpenAI board member. Mr. Nadella told Mr. Scott to get it done. The deal closed in 2019.

Mr. Altman and OpenAI had formed a for-profit company under the original nonprofit, they had $1 billion in fresh capital, and Microsoft had a new way to build artificial intelligence into its vast cloud computing service.

Not everyone inside OpenAI was happy.

Dario Amodei, a researcher with ties to the effective altruist community, had been on hand at the Rosewood hotel when OpenAI was born. Dr. Amodei, who endlessly twisted his curls between his fingers as he talked, was leading the lab's efforts to build a neural network called a large language model that could learn from enormous amounts of digital text. By analyzing countless Wikipedia articles, digital books and message boards, it could generate text on its own. It also had the unfortunate habit of making things up. It was called GPT-3, and it was released in the summer of 2020.

Researchers inside OpenAI, Google and other companies thought this rapidly improving technology could be a path to A.G.I.

But Dr. Amodei was unhappy about the Microsoft deal because he thought it was taking OpenAI in a really commercial direction. He and other researchers went to the board to try to push Mr. Altman out, according to five people familiar with the matter. After they failed, they left. Like DeepMind's founders before them, they worried that their new corporate overlords would favor commercial interests over safety.

In 2021, the group of about 15 engineers and scientists created a new lab called Anthropic. The plan was to build A.I. the way the effective altruists thought it should done — with very tight controls.

"There was no attempt to remove Sam Altman from OpenAI by the co-founders of Anthropic," said an Anthropic spokeswoman, Sally Aldous. "The co-founders themselves came to the conclusion that they wished to depart OpenAI to start their own company, made this known to OpenAI's leadership, and over several weeks negotiated an exit on mutually agreeable terms."

Anthropic accepted a $4 billion investment from Amazon and another $2 billion from Google two years later.

## The Reveal

After OpenAI received another $2 billion from Microsoft, Mr. Altman and another senior executive, Greg Brockman, visited Bill Gates at his sprawling mansion on the shores of Lake Washington, outside Seattle. The Microsoft founder was no longer involved in the company day to day but kept in regular touch with its executives.

Hokyoung Kim

Over dinner, Mr. Gates told them he doubted that large language models could work. He would stay skeptical, he said, until the technology performed a task that required critical thinking — passing an A.P. biology test, for instance.

Five months later, on Aug. 24, 2022, Mr. Altman and Mr. Brockman returned and brought along an OpenAI researcher named Chelsea Voss. Ms. Voss had been a medalist in an international biology Olympiad as a high schooler. Mr. Nadella and other Microsoft executives were there, too.

On a huge digital display on a stand outside Mr. Gates's living room, the OpenAI crew presented a technology called GPT-4.

Mr. Brockman gave the system a multiple-choice advanced biology test, and Ms. Voss graded the answers. The first question involved polar molecules, groups of atoms with a positive charge at one end and a negative charge at the other. The system answered correctly and explained its choice. "It was only trained to provide an answer," Mr. Brockman said. "The

conversational nature kind of fell out, almost magically." In other words, it was doing things they hadn't really designed it to do.

There were 60 questions. GPT-4 got only one answer wrong.

Mr. Gates sat up in his chair, his eyes opened wide. In 1980, he had a similar reaction when researchers showed him the graphical user interface that became the basis for the modern personal computer. He thought GPT was that revolutionary.

By October, Microsoft was adding the technology across its online services, including its Bing search engine. And two months later OpenAI released its ChatGPT chatbot, which is now used by 100 million people every week.

OpenAI had beat the effective altruists at Anthropic. Mr. Page's optimists at Google scurried to release their own chatbot, Bard, but were widely perceived to have lost the race to OpenAI. Three months after ChatGPT's release, Google stock was down 11 percent. Mr. Musk was nowhere to be found.

But it was just the beginning.

Susan Beachy contributed research.

**Cade Metz** is a technology reporter and the author of "Genius Makers: The Mavericks Who Brought A.I. to Google, Facebook, and The World." He covers artificial intelligence, driverless cars, robotics, virtual reality and other emerging areas. More about Cade Metz

**Karen Weise** writes about technology and is based in Seattle. Her coverage focuses on Amazon and Microsoft, two of the most powerful companies in America. More about Karen Weise

**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

**Mike Isaac** is a technology correspondent for The Times based in San Francisco. He regularly covers Facebook and Silicon Valley. More about Mike Isaac