Google’s true origin partly lies in CIA and NSA research grants for mass surveillance

They're always watching. (Reuters/Brian Snyder)

SHARE

WRITTEN BY

Jeff Nesbit
Former director of legislative and public affairs, National Science Foundation

December 08, 2017

Two decades ago, the US intelligence community worked closely with Silicon Valley in an effort to track citizens in cyberspace. And Google is at the heart of that origin story. Some of the research that led to Google’s ambitious creation was funded and coordinated by a research group established by the intelligence community to find ways to track individuals and groups online.

The intelligence community hoped that the nation's leading computer scientists could take non-classified information and user data, combine it with what would become known as the internet, and begin to create for-profit, commercial enterprises to suit the needs of both the intelligence community and the public. They hoped to direct the supercomputing revolution from the start in order to make sense of what millions of human beings did inside this digital information network. That collaboration has made a comprehensive public-private mass surveillance state possible today.

The story of the deliberate creation of the modern mass-surveillance state includes elements of Google's surprising, and largely unknown, origin. It is a somewhat different creation story than the one the public has heard, and explains what Google cofounders Sergey Brin and Larry Page set out to build, and why.

But this isn't just the origin story of Google: It's the origin story of the mass-surveillance state, and the government money that funded it.

Backstory: The intelligence community and Silicon Valley

In the mid 1990s, the intelligence community in America began to realize that they had an opportunity. The supercomputing community was just beginning to migrate from university settings into the private sector, led by investments from a place that would come to be known as Silicon Valley.

nted to alley's

y for rity

A digital revolution was underway: one that would transform the world of data gathering and how we make sense of massive amounts of information. The intelligence community wanted to shape Silicon Valley's supercomputing efforts at their inception so they would be useful for both military and homeland security purposes. Could this supercomputing network, which would become capable of storing terabytes of information, make intelligent sense of the digital trail that human beings leave behind?

Answering this question was of great interest to the intelligence community.

Intelligence-gathering may have been *their* world, but the Central Intelligence Agency (CIA) and the National Security Agency (NSA) had come to realize that their future was likely to be profoundly shaped

outside the government. It was at a time when military and intelligence budgets within the Clinton administration were in jeopardy, and the private sector had vast resources at their disposal. If the intelligence community wanted to conduct mass surveillance for national security purposes, it would require cooperation between the government and the emerging supercomputing companies.

To do this, they began reaching out to the scientists at American universities who were creating this supercomputing revolution. These scientists were developing ways to do what no single group of human beings sitting at work stations in the NSA and the CIA could ever hope to do: gather huge amounts of data and make intelligent sense of it.

A rich history of the government's science funding

There was already a long history of collaboration between America's best scientists and the intelligence community, from the creation of the atomic bomb and satellite technology to efforts to put a man on the moon.

**elf was
e of an
ort.**

In fact, the internet itself was created because of an intelligence effort: In the 1970s, the agency responsible for developing emerging technologies for military, intelligence, and national security purposes—the Defense Advanced Research Projects Agency (DARPA)—linked four supercomputers to handle massive data transfers. It handed the operations off to the National Science Foundation (NSF) a decade or so later, which proliferated the network across thousands of universities and, eventually, the public, thus creating the architecture and scaffolding of the World Wide Web.

Silicon Valley was no different. By the mid 1990s, the intelligence community was seeding funding to the most promising supercomputing efforts across academia, guiding the creation of efforts to make massive amounts of information useful for both the private sector as well as the intelligence community.

They funded these computer scientists through an unclassified, highly compartmentalized program that was managed for the CIA and the NSA by large military and intelligence contractors. It was called the Massive Digital Data Systems (MDDS) project.

The Massive Digital Data Systems (MDDS) project

MDDS was introduced to several dozen leading computer scientists at Stanford, CalTech, MIT, Carnegie Mellon, Harvard, and others in a white paper that described what the CIA, NSA, DARPA, and other agencies hoped to achieve. The research would largely be funded and managed by unclassified science agencies like NSF, which would allow the architecture to be scaled up in the private sector if it managed to achieve what the intelligence community hoped for.

"Not only are activities becoming more complex, but changing demands require that the IC [Intelligence Community] process different types as well as larger volumes of data," the intelligence community said in its 1993 MDDS white paper. "Consequently, the IC is taking a proactive role in stimulating research in the efficient management of massive databases and ensuring that IC requirements can be incorporated or adapted into commercial products. Because the challenges are not unique to any one agency, the Community Management Staff (CMS) has commissioned a Massive Digital Data Systems [MDDS] Working Group to address the needs and to identify and evaluate possible solutions."

Over the next few years, the program's stated aim was to provide more than a dozen grants of several million dollars each to advance this research concept. The grants were to be directed largely through the NSF so that the most promising, successful efforts could be captured as intellectual property and form the basis of companies attracting investments from Silicon Valley. This type of public-to-private innovation system helped launch powerful science and technology companies like Qualcomm, Symantec, Netscape, and others, and funded the pivotal research in areas like Doppler radar and fiber optics, which are central to large companies like AccuWeather, Verizon, and AT&T today. Today, the NSF provides nearly 90% of all federal funding for university-based computer-science research.

The CIA and NSA's end goal

The research arms of the CIA and NSA hoped that the best computer-science minds in academia could identify what they called "birds of a feather:" Just as geese fly together in large V shapes, or flocks of sparrows make sudden movements together in harmony, they predicted that like-minded groups of humans would move together online. The intelligence community named their first unclassified briefing for scientists the "birds of a feather" briefing, and the "Birds of a Feather Session on the Intelligence Community Initiative in Massive Digital Data Systems" took place at the Fairmont Hotel in San Jose in the spring of 1995.

ned ssified ntists eather"

Their research aim was to track digital fingerprints inside the rapidly expanding global information network, which was then known as the World Wide Web. Could an entire world of digital information be organized so that the requests humans made inside such a network be tracked and sorted? Could their queries be linked and ranked in order of importance? Could "birds of a feather" be identified inside this sea of information so that communities and groups could be tracked in an organized way?

By working with emerging commercial-data companies, their intent was to track like-minded groups of people across the internet and identify them from the digital fingerprints they left behind, much like forensic scientists use fingerprint smudges to identify criminals. Just as "birds of a feather flock together," they predicted that potential terrorists would communicate with each other in this new global, connected world—and they could find them by identifying patterns in this massive amount of new information. Once these groups were identified, they could then follow their digital trails everywhere.

Sergey Brin and Larry Page, computer-science boy wonders

In 1995, one of the first and most promising MDDS grants went to a computer-science research team at Stanford University with a decade-long history of working with NSF and DARPA grants. The primary objective of this grant was "query optimization of very complex queries that are described using the 'query flocks' approach." A second grant—the DARPA-NSF grant most closely associated with Google's origin—was part of a coordinated effort to build a massive digital library using the internet as its backbone. Both grants funded research by two graduate students who were making rapid advances in web-page ranking, as well as tracking (and making sense of) user queries: future Google cofounders Sergey Brin and Larry Page.

The research by Brin and Page under these grants became the heart of Google: people using search functions to find precisely what they wanted inside a very large data set. The intelligence community, however, saw a slightly different benefit in their research: Could the network be organized so efficiently that individual users could be uniquely identified and tracked?

This process is perfectly suited for the purposes of counter-terrorism and homeland security efforts: Human beings and like-minded groups who might pose a threat to national security can be uniquely identified online before they do harm. This explains why the intelligence community found Brin's and Page's research efforts so appealing; prior to this time, the CIA largely used human intelligence efforts in the field to identify people and groups that might pose threats. The ability to track them virtually (in conjunction with efforts in the field) would change everything.

It was the beginning of what in just a few years' time would become Google. The two intelligence-community managers charged with leading the program met regularly with Brin as his research progressed, and he was an author on several other research papers that resulted from this MDDS grant before he and Page left to form Google.

The grants allowed Brin and Page to do their work and contributed to their breakthroughs in web-page ranking and tracking user queries. Brin didn't work for the intelligence community—or for anyone else. Google had not yet been incorporated. He was just a Stanford researcher taking

advantage of the grant provided by the NSA and CIA through the unclassified MDDS program.

Left out of Google's story

The MDDS research effort has never been part of Google's origin story, even though the principal investigator for the MDDS grant specifically named Google as directly resulting from their research: "Its core technology, which allows it to find pages far more accurately than other search engines, was partially supported by this grant," he wrote. In a published research paper that includes some of Brin's pivotal work, the authors also reference the NSF grant that was created by the MDDS program.

Instead, every Google creation story only mentions just one federal grant: the NSF/DARPA "digital libraries" grant, which was designed to allow Stanford researchers to search the entire World Wide Web stored on the university's servers at the time. "The development of the Google algorithms was carried on a variety of computers, mainly provided by the NSF-DARPA-NASA-funded Digital Library project at Stanford," Stanford's Infolab says of its origin, for example. NSF likewise only references the digital libraries grant, not the MDDS grant as well, in its own history of Google's origin. In the famous research paper, "The Anatomy of a Large-Scale Hypertextual Web Search Engine," which describes the creation of Google, Brin and Page thanked the NSF and DARPA for its digital library grant to Stanford. But the grant from the intelligence community's MDDS program—specifically designed for the breakthrough that Google was built upon—has faded into obscurity.

Google has said in the past that it was not funded or created by the CIA. For instance, when stories circulated in 2006 that Google had received funding from the intelligence community for years to assist in counter-terrorism efforts, the company told Wired magazine founder John Battelle, "The statements related to Google are completely untrue."

Did the CIA directly fund the work of Brin and Page, and therefore create Google? No. But were Brin and Page researching precisely what the NSA, the CIA, and the intelligence community hoped for, assisted by their grants? Absolutely.

A
ssified,

To understand this significance, you have to consider what the intelligence community was trying to achieve as it seeded grants to the best computer-science minds in academia: The CIA and NSA funded an unclassified,

ized
ed
on to
that
actly

compartmentalized program designed from its inception to spur the development of something that looks almost exactly like Google. Brin's breakthrough research on page ranking by tracking user queries and linking them to the many searches conducted—essentially identifying "birds of a feather"—was largely the aim of the intelligence community's MDDS program. And Google succeeded beyond their wildest dreams.

The intelligence community's enduring legacy within Silicon Valley

Digital privacy concerns over the intersection between the intelligence community and commercial technology giants have grown in recent years. But most people still don't understand the degree to which the intelligence community relies on the world's biggest science and tech companies for its counter-terrorism and national-security work.

Civil-liberty advocacy groups have aired their privacy concerns for years, especially as they now relate to the Patriot Act. "Hastily passed 45 days after 9/11 in the name of national security, the Patriot Act was the first of many changes to surveillance laws that made it easier for the government to spy on ordinary Americans by expanding the authority to monitor phone and email communications, collect bank and credit reporting records, and track the activity of innocent Americans on the Internet," says the ACLU. "While most Americans think it was created to catch terrorists, the Patriot Act actually turns regular citizens into suspects."

When asked, the biggest technology and communications companies—from Verizon and AT&T to Google, Facebook, and Microsoft—say that they never deliberately and proactively offer up their vast databases on their customers to federal security and law enforcement agencies: They say that they only respond to subpoenas or requests that are filed properly under the terms of the Patriot Act.

But even a cursory glance through recent public records shows that there is a treadmill of constant requests that could undermine the intent behind this privacy promise. According to the data-request records that the companies make available to the public, in the most recent reporting period between 2016 and 2017, local, state and federal government authorities seeking information related to national security, counter-terrorism or criminal concerns issued more than 260,000 subpoenas, court orders, warrants, and other legal requests to Verizon, more than 250,000 such requests to AT&T, and nearly 24,000 subpoenas, search warrants, or court orders to Google. Direct national security or counter-terrorism requests are a small fraction of this overall group of requests, but the Patriot Act legal process has now become so routinized that the companies each have a group of employees who simply take care of the stream of requests.

In this way, the collaboration between the intelligence community and big, commercial science and tech companies has been wildly successful. When national security agencies need to identify and track people and groups, they know where to turn – and do so frequently. That was the goal in the beginning. It has succeeded perhaps more than anyone could have imagined at the time.

Learn how to write for Quartz Ideas. We welcome your comments at ideas@qz.com.