Praise for the Hardcover Edition of IN THE PLEX

"Levy is America's premier technology journalist. . . . He has produced the most interesting book ever written about Google. He makes the biggest intellectual challenges of computer science seem endlessly fun and fascinating. . . . We can expect many more books about Google. But few will deliver the lively, idea-based journalism of *In the Plex*."

—Siva Vaidhyanathan, *The Washington Post*

"[Steven Levy] spent much of the past three years playing anthropologist at one of the Internet's most interesting villages and set of inhabitants—the Googleplex and the tribe of Googlers who inhabit it. . . . A deep dive into Google's culture, history and technology."

—Mike Swift, *San Jose Mercury News*

"The rise of Google is an engrossing story, and nobody's ever related it in such depth."

—Hiawatha Bray, *The Boston Globe*

"The most comprehensive, intelligent and readable analysis of Google to date. Levy is particularly good on how those behind Google think and work. . . . What's more, his lucid introductions to Google's core technologies—the search engine and the company's data centres—are written in non-geek English and are rich with anecdotes and analysis. . . . *In the Plex* teems with original insight into Google's most controversial affairs."

—Andrew Keen, *New Scientist*

"Steven Levy's new account [of Google], *In the Plex*, is the most authoritative to date and in many ways the most entertaining."

—James Gleick, *The New York Review of Books*

"An instructive primer on how the minds behind the world's most influential internet company function."

—Richard Waters, *The Wall Street Journal*

"Almost nothing can stop a remarkable idea executed well at the right time, as Steven Levy's brisk-but-detailed history of Google, *In the Plex*, convincingly proves. . . . [It] makes obsolete previous books on the company."

—Jack Shafer, *San Francisco Chronicle*

"[A] dense, driven examination of the pioneering search engine that changed the face of the Internet. . . . Outstanding reportage delivered in the upbeat, informative fashion for which Levy is well known."

—*Kirkus Reviews* (starred review)

ALSO BY STEVEN LEVY

Facebook: The Inside Story

The Perfect Thing: How the iPod Shuffles Commerce, Culture, and Coolness

Crypto: How the Code Rebels Beat the Government—Saving Privacy in the Digital Age

Insanely Great: The Life and Times of Macintosh, the Computer That Changed Everything

Artificial Life: The Quest for a New Creation

The Unicorn's Secret: Murder in the Age of Aquarius

Hackers: Heroes of the Computer Revolution

IN THE PLEX

How Google Thinks, Works, and Shapes Our Lives

STEVEN LEVY

SIMON & SCHUSTER PAPERBACKS
New York London Toronto Sydney New Delhi

Simon & Schuster
1230 Avenue of the Americas
New York, NY 10020

This Simon & Schuster trade paperback edition February 2021

SIMON & SCHUSTER and colophon are registered trademarks of Simon & Schuster, Inc.

For information about special discounts for bulk purchases, please contact Simon & Schuster Special Sales at 1-866-506-1949 or business@simonandschuster.com.

The Simon & Schuster Speakers Bureau can bring authors to your live event.
For more information or to book an event, contact the Simon & Schuster Speakers Bureau at 1-866-248-3049 or visit our website at www.simonspeakers.com.

Interior design by Ruth Lee Mui

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data is available for the hardcover.

ISBN 978-1-4165-9658-5
ISBN 978-1-4165-9659-2 (pbk)
ISBN 978-1-4165-9671-4 (ebook)

CONTENTS

In memory of Philip Klass (1920–2010)

PROLOGUE

SEARCHING FOR GOOGLE

"Have you heard of Google?"

It was a blazing hot July day in 2007, in the rural Indian village of Ragihalli, located thirty miles outside Bangalore. Twenty-two people from a company based in Mountain View, California, had driven in SUVs and vans up an unpaved road to this enclave of seventy threadbare huts with cement floors, surrounded by fields occasionally trampled by unwelcome elephants. Though electricity had come to Ragihalli some years earlier, there was not a single personal computer in the community. The visit had begun awkwardly, as the outsiders piled out of the cars and faced the entire population of the village, about two hundred people, who had turned out to welcome them. It was as if these well-dressed Westerners had dropped in from another planet, which in a sense they had. Young schoolchildren were pushed forward, and they performed a song. The visitors, in turn, gave the children notebooks and candy. There was an uncomfortable silence, broken when Marissa Mayer, the delegation's leader, a woman of thirty-two, said, "Let's interact with them." The group fanned out and began to engage the villagers in awkward conversation.

That is how Alex Vogenthaler came to ask a spindly young man with a wide smile whether he had heard of Google, Vogenthaler's employer. It was a question that he would never have had to ask in his home country: virtually everyone in the United States and everywhere in the wired-up world

knew Google. Its uncannily effective Internet search product had changed the way people accessed information, changed the way they *thought* about information. Its 2004 IPO had established it as an economic giant. And its founders themselves were the perfect examples of the superbrainy engineering mentality that represented the future of business in the Internet age.

The villager admitted that, no, he had never heard of this Google. "What is it?" he asked. Vogenthaler tried to explain in the simplest terms that Google was a company that operated on the Internet. People used it to search for information. You would ask it a question, and it would immediately give you the answer from huge repositories of information it had gathered on the World Wide Web.

The man listened patiently but clearly was more familiar with rice fields than search fields.

Then the villager held up a cell phone. "Is this you what mean?" he seemed to ask.

The little connectivity meter on the phone display had four bars. There are significant swaths of the United States of America where one can barely pull in a signal—or gets no bars at all. But here in rural India, the signal was strong.

Google, it turns out, was on the verge of a multimillion-dollar mobile effort to make smart phones into information prostheses, adjuncts to the human brain that would allow people to get information to a vast swath of all the world's knowledge instantly. This man might not know Google yet, but the company would soon be in Ragihalli. And then he *would* know Google.

I witnessed this exchange in 2007 as an observer on the annual trip of Google associate product managers, a select group pegged as the company's future leaders. We began our journey in San Francisco and touched down in Tokyo, Beijing, Bangalore, and Tel Aviv before returning home sixteen days later.

My participation on the trip had been a consequence of a long relationship with Google. In late 1998, I'd heard buzz about a smarter search engine and tried it out. Google was miles better than anything I'd used before. When I heard a bit about the site's method of extracting such good results—it relied on sort of a web-based democracy—I became even more intrigued. This is how I put it in the February 22, 1999, issue of *Newsweek:* "Google, the Net's hottest search engine, draws on feedback from the web itself to deliver more relevant results to customer queries."

Later that year, I arranged with Google's newly hired director of corporate communications, Cindy McCaffrey, to visit its Mountain View headquarters. One day in October I drove to 2400 Bayshore Parkway, where Google had just moved from its previous location above a Palo Alto bicycle shop. I'd visited a lot of start-ups and wasn't really surprised by the genial chaos—a vast room, with cubicles yet unfilled and a cluster of exercise balls. However, I hadn't expected that instead of being attired in traditional T-shirts and jeans, the employees were decked out in costumes. I had come on Halloween.

"Steven, meet Larry Page and Sergey Brin," said Cindy, introducing me to the two young men who had founded the company as Stanford graduate students. Larry was dressed as a Viking, with a long-haired fur vest and a hat with long antlers protruding. Sergey was in a cow suit. On his chest was a rubber slab from which protruded huge, wart-specked teats. They greeted me cheerfully and we all retreated to a conference room where the Viking and the cow explained the miraculous powers of Google's PageRank technology.

That was the first of many interviews I would conduct at Google. Over the next few years, the company became a focus of my technology reporting at *Newsweek*. Google grew from the small start-up I had visited to a behemoth of more than 20,000 employees. Every day, billions of people used its search engine, and Google's remarkable ability to deliver relevant results in milliseconds changed the way the world got its information. The people who clicked on its ads made Google wildly profitable and turned its founders into billionaires—and triggered an outcry among traditional beneficiaries of ad dollars.

Google also became known for its irreverent culture and its data-driven approach to business decision making; management experts rhapsodized about its unconventional methods. As the years went by, Google began to interpret its mission—to gather and make accessible and useful the world's information—in the broadest possible sense. The company created a series of web-based applications. It announced its intention to scan all the world's books. It became involved in satellite imagery, mobile phones, energy generation, photo storage. Clearly, Google was one of the most important contributors to the revolution of computers and technology that marked a turning point in civilization. I knew I wanted to write a book about the company but wasn't sure how.

Then in early July 2007, I was asked to join the associate product managers on their trip. It was an unprecedented invitation from a company

that usually limits contact between journalists and its employees. The APM program, I learned, was a highly valued initiative. To quote the pitch one of the participants made in 2006 to recent and upcoming college graduates: "We invest more into our APMs than any other company has ever invested into young employees. . . . We envision a world where everyone is awed by the fact that Google's executives, the best CEOs in the Silicon Valley, and the most respected leaders of global non-profits all came through the Google APM program." Eric Schmidt, Google's CEO, told me, "One of these people will probably be our CEO one day—we just don't know which one."

The eighteen APMs on the trip worked all over Google: in search, advertising, applications, and even stealth projects such as Google's attempt to capture the rights to include magazines in its index. Mayer's team, along with the APMs themselves, had designed the agenda of the trip. Every activity had an underlying purpose to increase the participants' understanding of a technology or business issue, or make them more (in the parlance of the company) "Googley." In Tokyo, for instance, they engaged in a scavenger hunt in the city's legendary Akihabara electronics district. Teams of APMs were each given $50 to buy the weirdest gadgets they could find. Ducking into backstreets with stalls full of electronic parts and gizmos, they wound up with a cornucopia: USB-powered ashtrays shaped like football helmets that suck up smoke; a plate-sized disk that simulated the phases of the moon; a breathalyzer you could install in your car; and a stubby wand that, when waved back and forth, spelled out words in LED lights. In Bangalore, there was a different shopping hunt—an excursion to the market area where the winner of the competition would be the one who haggled best. (Good training for making bulk purchases of computers or even buying an Internet start-up.) Another Tokyo high point was the 5 A.M. trip to the Tsukiji fish market. It wasn't the fresh sushi that fascinated the APMs but the mechanics of the fish auction, in some ways similar to the way Google works its AdWords program.

In China, Google's top executive there, Kai-Fu Lee, talked of balancing Google's freewheeling style with government rules—and censorship. But during interviews with Chinese consumers, the APMs were discouraged to hear the perception of the company among locals: "Baidu [Google's local competitor] knows more [about China] than Google," said one young man to his APM interlocutors.

At every office the APMs visited, they attended meetings with local Googlers, first learning about projects under way and then explaining to

the residents what was going on at Mountain View headquarters. I began to get an insider's sense of Google's product processes—and how serving its users was akin to a crusade. An interesting moment occurred in Bangalore when Mayer was taking questions from local engineers after presenting an overview of upcoming products. One of them asked, "We've heard the road map for products, what's the road map for revenues?" She almost bit his head off. "That's *not* the way to think," she said. "We are focused on our *users.* If we make them happy, we will have revenues."

The most fascinating part of the trip was the time spent with the young Googlers. They were generally from elite colleges, with SAT scores approaching or achieving perfection. Carefully culled from thousands of people who would have killed for the job, their personalities and abilities were a reflection of Google's own character. During a bus ride to the Great Wall of China, one of the APMs charted the group demographics and found that almost all had parents who were professionals and more than half had parents who taught at a university—which put them in the company of Google's founders. They all grew up with the Internet and considered its principles to be as natural as the laws of gravity. They were among the brightest and most ambitious of a generation that was better equipped to handle the disruptive technology wave than their elders were. Their minds hummed like tuning forks in resonance with the company's values of speed, flexibility, and a deep respect for data.

Yet even while immersed in an optimism bubble with these young people, I could see the strains that came with Google's abrupt growth from a feisty start-up to a market-dominating giant with more than 20,000 employees. The APMs had spent a year navigating the folkways of a complicated corporation, albeit a determinedly different one—and now they were almost senior employees. What's more, I was stunned when a poll of my fellow travelers revealed that not a single one of them saw him- or herself working for Google in five years. Marissa Mayer took this news calmly, claiming that such ambition was why they had been hired in the first place. "This is the gene that Larry and Sergey look for," she told me. "Even if they leave, it's still good for us. They're going to take the Google DNA with them." (Almost five years to the day later, Mayer herself would leave Google, to become CEO of the struggling Internet company Yahoo.)

After covering the company for almost a decade, I thought I knew it pretty well, but the rare view of the company I got in those two weeks made me see it in a different, wider light. Still, there were considerable mysteries. Google was a company built on the values of its founders, who

harbored ambitions to build a powerful corporation that would impact the entire world, at the same time loathing the bureaucracy and commitments that running such a company would entail. Google professed a sense of moral purity—as exemplified by its informal motto, "Don't be evil"—but it seemed to have a blind spot regarding the consequences of its own technology on privacy and property rights. A bedrock principle of Google was serving its users—but a goal was building a giant artificial intelligence learning machine that would bring uncertain consequences to the way all of us live. From the very beginning, its founders said that they wanted to change the world. But who were they, and what did they envision this new world order to be?

After the trip I realized that the best way to answer these questions was to report as much as possible from inside Google. Just as I'd had a rare glimpse into its inner workings during that summer of 2007, I would try to immerse myself more deeply into its engineering, its corporate life, and its culture, to report how it really operated, how it developed its products, and how it was managing its growth and public exposure. I would be an outsider with an insider's view.

To do this, of course, I'd need cooperation. Fortunately, based on our long relationship, Google's executives, including "LSE"—Larry Page, Sergey Brin, and Eric Schmidt—agreed to let me in. During the next two years—a critical time when Google's halo lost some of its glow even as the company grew more powerful—I interviewed hundreds of current and former Googlers and attended a variety of meetings in the company. These included product development meetings, "interface reviews," search launch meetings, privacy council sessions, weekly TGIF all-hands gatherings, and the gatherings of the high command known as Google Product Strategy (GPS) meetings, where projects and initiatives are approved or rejected. I also ate a lot of meals at Andale, the burrito joint in Google's Building 43.

What I discovered was a company exulting in creative disorganization, even if the creativity was not always as substantial as hoped for. Google had massive goals, and the entire company channeled its values from the founders. Its mission was collecting and organizing all the world's information—and that's only the beginning. From the very start, its founders saw Google as a vehicle to realize the dream of artificial intelligence in augmenting humanity. To realize their dreams, Page and Brin had to build a huge company. At the same time, they attempted to maintain as much as possible the nimble, irreverent, answer-to-no-one freedom of a small start-

up. In the two years I researched this book, the clash between those goals reached a peak, as David had become a Goliath.

My inside perspective also provided me the keys to unlock more of the secrets of Google's two "black boxes"—its search engine and its advertising model—than had previously been disclosed. Google search is part of our lives, and its ad system is the most important commercial product of the Internet age. In this book, for the first time, readers can learn the full story of their development, evolution, and inner workings. Understanding those groundbreaking products helps us understand Google and its employees because their operation embodies both the company's values and its technological philosophy. More important, understanding them helps us understand our own world—and tomorrow's.

The science fiction writer William Gibson once said that the future is already here—just not evenly distributed. At Google, the future is already under way. To understand this pioneering company and its people is to grasp our technological destiny. And so here is Google: how it works, what it thinks, why it's changing, how it will continue to change us. And how it hopes to maintain its soul.

PART ONE

THE WORLD ACCORDING TO GOOGLE

Biography of a Search Engine

1

"It was science fiction more than computer science."

On February 18, 2010, Judge Denny Chin of the New York Southern District federal court took stock of the packed gallery in Courtroom 23B. It was going to be a long day. He was presiding over a hearing that would provide only a gloss to hundreds of submissions he had already received on this case. "There is just too much to digest," he said. He shook his head, preparing himself to hear the arguments of twenty-seven representatives of various interest groups or corporations, as well as presentations by some of the lawyers for various parties, lawyers who filled every place in two long tables before him.

The case was *The Authors Guild, Inc., Association of American Publishers, et al. v. Google Inc.* It was a lawsuit tentatively resolved by a class settlement agreement in which an authors' group and a publishers' association set conditions for a technology company to scan and sell books. Judge Chin's decision would involve important issues affecting the future of digital works, and some of the speakers before the court engaged on those issues. But many of the objectors—and most who addressed the court were

objectors to the settlement—focused on a young company headquartered on a sprawling campus in Mountain View, California. That company was Google. The speakers seemed to distrust it, fear it, even despise it.

"A major threat to . . . freedom of expression and participation in cultural diversity"
"An unjustified monopoly"
"Eviscerates privacy protections"
"Concealment and misdirection"
"Price fixing . . . a massive market distortion . . . preying on the desperate"
"May well be a per se violation of the antitrust laws"

(That last statement held special weight, as it came from the U.S. deputy assistant attorney general.)

But the federal government was only one of Google's surprising opponents. Some of the others were supporters of the public interest, monitoring the privacy rights and pocketbooks of citizens. Others were advocates of free speech. There was even an objector representing the folk-singer Arlo Guthrie.

The irony was that Google itself explicitly embraced the lofty values and high moral standards that it was being attacked for flouting. Its founders had consistently stated that their goal was to make the world better, specifically by enabling humanity's access to information. Google had created an astonishing tool that took advantage of the interconnected nature of the burgeoning World Wide Web, a tool that empowered people to locate even obscure information within seconds. This search engine transformed the way people worked, entertained themselves, and learned. Google made historic profits from that product by creating a new form of advertising—nonintrusive and even useful. It hired the sharpest minds in the world and encouraged them to take on challenges that pushed the boundaries of innovation. Its focus on engineering talent to accomplish difficult goals was a national inspiration. It even warned its shareholders that the company would sometimes pursue business practices that serve humanity even at the expense of lower profits. It accomplished all those achievements with a puckish irreverence that captivated the public and made heroes of its employees.

But that didn't matter to the objectors in Judge Chin's courtroom. Those people were Google's natural allies, and they thought that Google was no longer . . . good. The mistrust and fear in the courtroom were

reflected globally by governments upset by Google's privacy policies and businesses worried that Google's disruptive practices would target them next. Everywhere Google's executives turned, they were faced with protests and lawsuits.

The course of events was baffling to Google's two founders, Larry Page and Sergey Brin. Of all Google's projects, the one at issue in the hearing—Google's Book Search project—was perhaps the most idealistic. It was an audacious attempt to digitize every book ever printed, so that anyone in the world could locate the information within. Google would not give away the full contents of the books, so when users discovered them, they would have reason to buy them. Authors would have new markets; readers would have instant access to knowledge. After being sued by publishers and authors, Google made a deal with them that would make it even easier to access the books and to buy them on the spot. Every library would get a free terminal to connect to the entire corpus of the world's books. To Google, it was a boon to civilization.

Didn't people understand?

By all metrics, the company was still thriving. Google still retained its hundreds of millions of users, hosted billions of searches every day, and had growing businesses in video and wireless devices. Its employees were still idealistic and ambitious in the best sense. But a shadow now darkened Google's image. To many outsiders, the corporate motto that Google had taken seriously—"Don't be evil"—had become a joke, a bludgeon to be used against it.

What had happened?

Doing good was Larry Page's plan from the very beginning. Even as a child, he wanted to be an inventor, not simply because his mind aligned perfectly with the nexus of logic and technology (which it did) but because, he says, "I really wanted to change the world."

Page grew up in Lansing, Michigan, where his father taught computer science at Michigan State. His parents divorced when he was eight, but he was close with both his father and mother—who had her own computer science degree. Naturally, he spoke computers as a primary language. As he later told an interviewer, "I think I was the first kid in my elementary school to turn in a word-processed document."

Page was not a social animal—people who talked to him often wondered if there were a jigger of Asperger's in the mix—and could unnerve people by simply not talking. But when he did speak, more often than not

he would come out with ideas that bordered on the fantastic. Attending a summer program in leadership (motto: "A healthy disregard for the impossible") helped move him to action. At the University of Michigan, he became obsessed with transportation and drew up plans for an elaborate monorail system in Ann Arbor, replacing the mundane bus system with a "futuristic" commute between the dorms and the classrooms. It seemed to come as a surprise to him that a fanciful multimillion-dollar transit fantasy from an undergraduate would not be quickly embraced and implemented. (Fifteen years after he graduated, Page would bring up the issue again in a meeting with the university's president.)

His intelligence and imagination were clear. But when you got to know him, what stood out was his ambition. It expressed itself not as a personal drive (though there was that, too) but as a general principle that everyone should think big and then make big things happen. He believed that the only true failure was not attempting the audacious. "Even if you fail at your ambitious thing, it's very hard to fail completely," he says. "That's the thing that people don't get." Page *always* thought about that. When people proposed a short-term solution, Page's instinct was to think long term. There would eventually be a joke among Googlers that Page "went to the future and came back to tell us about it."

Page earned a degree in computer science like his father did. But his destiny was in California, specifically in the Silicon Valley. In a way, Page's arrival at Stanford was a homecoming. He'd lived there briefly in 1979 when his dad had spent a sabbatical at Stanford; some faculty members still remembered him as an insatiably curious seven-year-old. In 1995, Stanford was not only the best place to pursue cutting-edge computer science but, because of the Internet boom, was also the world capital of ambition. Fortunately, Page's visions extended to the commercial: "Probably from when I was twelve, I knew I was going to start a company eventually," he'd later say. Page's brother, nine years older, was already in Silicon Valley, working for an Internet start-up.

Page chose to work in the department's Human-Computer Interaction Group. The subject would stand Page in good stead in the future with respect to product development, even though it was not in the HCI domain to figure out a new model of information retrieval. On his desk and permeating his conversations was Apple interface guru Donald Norman's classic tome *The Psychology of Everyday Things*, the bible of a religion whose first, and arguably only, commandment is "The user is always right." (Other Norman disciples, such as Jeff Bezos at Amazon.com, were adopt-

ing this creed on the web.) Another influential book was a biography of Nikola Tesla, the brilliant Serb scientist; though Tesla's contributions arguably matched Thomas Edison's—and his ambitions were grand enough to impress even Page—he died in obscurity. "I felt like he was a great inventor and it was a sad story," says Page. "I feel like he could've accomplished much more had he had more resources. And he had trouble commercializing the stuff he did. Probably more trouble than he should've had. I think that was a good lesson. I didn't want to just invent things, I also wanted to make the world better, and in order to do that, you need to do more than just invent things."

The summer before entering Stanford, Page attended a program for accepted candidates that included a tour of San Francisco. The guide was a grad student Page's age who'd been at Stanford for two years. "I thought he was pretty obnoxious," Page later said of the guide, Sergey Brin. The content of the encounter is now relegated to legend, but their argumentative banter was almost certainly good-natured. Despite the contrast in personalities, in some ways they were twins. Both felt most comfortable in the meritocracy of academia, where brains trumped everything else. Both had an innate understanding of how the ultraconnected world that they enjoyed as computer science (CS) students was about to spread throughout society. Both shared a core belief in the primacy of data. And both were rock stubborn when it came to pursuing their beliefs. When Page settled in that September, he became close friends with Brin, to the point where people thought of them as a set: LarryAndSergey.

Born in Russia, Brin was four when his family immigrated to the United States. His English still maintained a Cyrillic flavor, and his speech was dotted with anachronistic Old World touches such as the use of "whatnot" when peers would say "stuff like that." He had arrived at Stanford at nineteen after whizzing through the University of Maryland, where his father taught, in three years; he was one of the youngest students ever to start the Stanford PhD program. "He skipped a million years," says Craig Silverstein, who arrived at Stanford a year later, and would eventually become Google's first employee. Sergey was a quirky kid who would zip through Stanford's hallways on omnipresent Rollerblades. He also had an interest in trapeze. But the professors understood that behind the goofiness was a formidable mathematical mind. Soon after arriving at Stanford, he knocked off all the required tests for a doctorate and was free to sample the courses until he found a suitable entree for a thesis. He supplemented his academics with swimming, gymnastics, and sailing. (When his father asked

him in frustration whether he planned to take advanced courses, he said that he might take advanced swimming.) Donald Knuth, a Stanford professor whose magisterial series of books on the art of computer programming made him the Proust of computer code, recalls driving down the Pacific coast to a conference with Sergey one afternoon and being impressed at his grasp of complicated issues. His adviser, Hector Garcia-Molina, had seen a lot of bright kids go through Stanford, but Brin stood out. "He was brilliant," Garcia-Molina says.

One task that Brin took on was a numbering scheme for the new Gates Computer Science Building, which was to be the home of the department. (His system used mathematical flourishes.) The structure was named after William Henry Gates III, better known as Bill, the cofounder of Microsoft. Though Gates had spent a couple of years at Harvard and endowed a building named after his mother there, he went on a small splurge of funding palatial new homes for computer science departments at top technical institutions that he *didn't* attend, including MIT and Carnegie Mellon—along with Stanford, the trifecta of top CS programs. Even as they sneered at Windows, the next generation of wizards would study in buildings named after Bill Gates.

Did Gates ever imagine that one of those buildings would incubate a rival that might destroy Microsoft?

The graduate computer science program at Stanford was built around close relationships between students and faculty members. They would team up to work on big, real-world problems; the fresh perspective of the young people maintains the vitality of the professor's interests. "You always follow the students," says Terry Winograd, who was Page's adviser. (Page would often remind him that they had met during his dad's Stanford sabbatical.) Over the years Winograd had become an expert at figuring out where students stood on the spectrum of brainiacs who found their way into the department. Some were kids whose undergrad record was straight A pluses, GRE scores scraping perfection, who would come in and say, "What thesis should I work on?" On the other end of the spectrum were kids like Larry Page, who would come in and say, "Here's what I think I can do." And his proposals were *crazy.* He'd come into the office and talk about doing something with space tethers or solar kites. "It was science fiction more than computer science," recalls Winograd. But an outlandish mind was a valuable asset, and there was definitely a place in the current science to channel wild creativity.

In 1995, that place was the World Wide Web. It had sprung from the

restless brain of a (then)-obscure British engineer named Tim Berners-Lee, who was working as a technician at the CERN physics research lab in Switzerland. Berners-Lee could sum up his vision in a sentence: "Suppose all the information stored on computers everywhere were linked . . . there would be a single global information space."

The web's pedigree could be traced back to a 1945 paper by the American scientist Vannevar Bush. Entitled "As We May Think," it outlined a vast storage system called a "memex," where documents would be connected, and could be recalled, by information breadcrumbs called "trails of association." The timeline continued to the work of Douglas Engelbart, whose team at the Stanford Research Institute devised a linked document system that lived behind a dazzling interface that introduced the metaphors of windows and files to the digital desktop. Then came a detour to the brilliant but erratic work of an autodidact named Ted Nelson, whose ambitious Xanadu Project (though never completed) was a vision of disparate information linked by "hypertext" connections. Nelson's work inspired Bill Atkinson, a software engineer who had been part of the original Macintosh team; in 1987 he came up with a link-based system called HyperCard, which he sold to Apple for $100,000 on the condition that the company give it away to all its users. But to really fulfill Vannevar Bush's vision, you needed a huge system where people could freely post and link their documents.

By the time Berners-Lee had his epiphany, that system was in place: the Internet. While the earliest websites were just ways to distribute academic papers more efficiently, soon people began writing sites with information of all sorts, and others created sites just for fun. By the mid-1990s, people were starting to use the web for profit, and a new word, "e-commerce," found its way into the lexicon. Amazon.com and eBay became Internet giants. Other sites positioned themselves as gateways, or portals, to the wonders of the Internet.

As the web grew, its linking structure accumulated a mind-boggling value. It treated the aggregate of all its contents as a huge compost of ideas, any one of which could be reached by the act of connecting one document to another. When you looked at a page you could see, usually highlighted in blue, the pointers to other sites that the webmaster had coded on the page—that was the hypertext idea that galvanized Bush, Nelson, and Atkinson. But for the first time, as Berners-Lee had intended, the web was coaxing a critical mass of these linked sites and documents into a single network. In effect, the web was an infinite database, a sort of crazily expanding universe of human knowledge that, in theory, could hold every

insight, thought, image, and product for sale. And all of it had an intricate lattice of cross-connections created by the independent linking activity of anyone who had built a page and coded in a link to something elsewhere on the web.

In retrospect, the web was to the digital world what the Louisiana Purchase was to the young United States: the opportunity of a century.

Berners-Lee's creation was so new that when Stanford got funding from the National Science Foundation in the early 1990s to start a program called the Digital Library Project, the web wasn't mentioned in the proposal. "The theme of that project was interoperability—how can we make all these resources work together?" recalls Hector Garcia-Molina, who cofounded the project. By 1995 though, Garcia-Molina knew that the World Wide Web would inevitably be part of the projects concocted by the students who worked with the program, including Page and Brin.

Brin already had a National Science Foundation fellowship and didn't need funding, but he was trying to figure out a dissertation topic. His loose focus was data mining, and with Rajeev Motwani, a young professor he became close with, he helped start a research group called MIDAS, which stood for Mining Data at Stanford. In a résumé he posted on the Stanford site in 1995, he talked about "a new project" to generate personalized movie ratings. "The way it works is as follows," he wrote. "You rate the movies you have seen. Then the system finds other users with similar tastes to extrapolate how much you like other movies." Another project he worked on with Garcia-Molina and another student was a system that detected copyright violations by automating searches for duplicates of documents. "He came up with some good algorithms for detecting copies," says Garcia-Molina. "Now you use Google."

Page was also seeking a dissertation topic. One idea he presented to Winograd, a collaboration with Brin, seemed more promising than the others: creating a system where people could make annotations and comments on websites. But the more Page thought about annotation, the messier it got. For big sites, there would probably be a lot of people who wanted to mark up a page. How would you figure out who gets to comment or whose comment would be the one you'd see first? For that, he says, "We needed a rating system."

Having a human being determine the ratings was out of the question. First, it was inherently impractical. Further, humans were unreliable. Only algorithms—well drawn, efficiently executed, and based on sound data—could deliver unbiased results. So the problem became finding the right

data to determine whose comments were more trustworthy, or interesting, than others. Page realized that such data already existed and no one else was really using it. He asked Brin, "Why don't we use the links on the web to do that?"

Page, a child of academia, understood that web links were like citations in a scholarly article. It was widely recognized that you could identify which papers were really important without reading them—simply tally up how many other papers cited them in notes and bibliographies. Page believed that this principle could also work with web pages. But getting the right data would be difficult. Web pages made their outgoing links transparent: built into the code were easily identifiable markers for the destinations you could travel to with a mouse click from that page. But it wasn't obvious at all what linked *to* a page. To find that out, you'd have to somehow collect a database of links that connected to some other page. Then you'd go *backward*.

That's why Page called his system BackRub. "The early versions of hypertext had a tragic flaw: you couldn't follow links in the other direction," Page once told a reporter. "BackRub was about reversing that."

Winograd thought this was a great idea for a project, but not an easy one. To do it right, he told Page, you'd really have to capture a significant chunk of the World Wide Web's link structure. Page said, sure, he'd go and download the web and get the structure. He figured it would take a week or something. "And of course," he later recalled, "it took, like, years." But Page and Brin attacked it. Every other week Page would come to Garcia-Molina's office asking for disks and equipment. "That's fine," Garcia-Molina would say. "This is a great project, but you need to give me a budget." He asked Page to pick a number, to say how much of the web he needed to crawl, and to estimate how many disks that would take. "I want to crawl the *whole* web," Page said.

Page indulged in a little vanity in naming the part of the system that rated websites by the incoming links: he called it PageRank. But it was a sly vanity; many people assumed the name referred to web pages, not a surname.

Since Page wasn't a world-class programmer, he asked a friend to help out. Scott Hassan was a full-time research assistant at Stanford, working for the Digital Library Project program while doing part-time grad work. Hassan was also good friends with Brin, whom he'd met at an Ultimate Frisbee game during his first week at Stanford. Page's program "had so many bugs in it, it wasn't funny," says Hassan. Part of the problem was that Page was

using the relatively new computer language Java for his ambitious project, and Java kept crashing. "I went and tried to fix some of the bugs in Java itself, and after doing this ten times, I decided it was a waste of time," says Hassan. "I decided to take his stuff and just rewrite it into the language I knew much better that didn't have any bugs."

He wrote a program in Python—a more flexible language that was becoming popular for web-based programs—that would act as a "spider," so called because it would crawl the web for data. The program would visit a web page, find all the links, and put them into a queue. Then it would check to see if it had visited those link pages previously. If it hadn't, it would put the link on a queue of future destinations to visit and repeat the process. Since Page wasn't familiar with Python, Hassan became a member of the team. He and another student, Alan Steremberg, became paid assistants to the project.

Brin, the math prodigy, took on the huge task of crunching the mathematics that would make sense of the mess of links uncovered by their monster survey of the growing web.

Even though the small team was going somewhere, they weren't quite sure of their destination. "Larry didn't have a plan," says Hassan. "In research you explore something and see what sticks."

By March 1996, they began a test, starting at a single page, the Stanford computer science department home page. The spider located the links on the page and fanned out to all the sites that linked to Stanford, then to the sites that linked to *those* websites. "That first one just used the titles of documents because collecting the documents themselves required a lot of data and work," says Page. After they snared about 15 million of those titles, they tested the program to see which websites it deemed more authoritative.

"Even the first set of results was very convincing," Hector Garcia-Molina says. "It was pretty clear to everyone who saw this demo that this was a very good, very powerful way to order things."

"We realized it worked really, really well," says Page. "And I said, 'Wow, the big problem here is not annotation. We should now use it not just for ranking annotations, but for ranking *searches*.'" It seemed the obvious application for an invention that gave a ranking to every page on the web. "It was pretty clear to me and the rest of the group," he says, "that if you have a way of ranking things based not just on the page itself but based on what the world thought of that page, that would be a really valuable thing for search."

▼ ▼ ▼

The leader in web search at that time was a program called AltaVista that came out of Digital Equipment Corporation's Western Research Laboratory. A key designer was Louis Monier, a droll Frenchman and idealistic geek who had come to America with a doctorate in 1980. DEC had been built on the minicomputer, a once innovative category now rendered a dinosaur by the personal computer revolution. "DEC was very much living in the past," says Monier. "But they had small groups of people who were very forward-thinking, experimenting with lots of toys." One of those toys was the web. Monier himself was no expert in information retrieval but a big fan of data in the abstract. "To me, that was the secret—data," he says. What the data was telling him was that if you had the right tools, it was possible to treat everything in the open web like a single document.

Even at that early date, the basic building blocks of web search had been already set in stone. Search was a four-step process. First came a sweeping scan of all the world's web pages, via a spider. Second was indexing the information drawn from the spider's crawl and storing the data on racks of computers known as servers. The third step, triggered by a user's request, identified the pages that seemed best suited to answer that query. That result was known as search quality. The final step involved formatting and delivering the results to the user.

Monier was most concerned with the second step, the time-consuming process of crawling through millions of documents and scooping up the data. "Crawling at that time was slow, because the other side would take on average four seconds to respond," says Monier. One day, lying by a swimming pool, he realized that you could get everything in a timely fashion by parallelizing the process, covering more than one page at a time. The right number, he concluded, was a thousand pages at once. Monier figured out how to build a crawler working on that scale. "On a single machine I had one thousand threads, independent processes asking things and not stepping on each other's toes."

By late 1995, people in DEC's Western Research Lab were using Monier's search engine. He had a tough time convincing his bosses to open up the engine to the public. They argued that there was no way to make money from a search engine but relented when Monier sold them on the public relations aspect. (The system would be a testament to DEC's powerful new Alpha processing chip.) On launch day, AltaVista had 16 million documents in its indexes, easily besting anything else on the net. "The big ones then had maybe a million pages," says Monier. That was the power of

AltaVista: its breadth. When DEC opened it to outsiders on December 15, 1995, nearly 300,000 people tried it out. They were dazzled.

AltaVista's actual search quality techniques—what determined the ranking of results—were based on traditional information retrieval (IR) algorithms. Many of those algorithms arose from the work of one man, a refugee from Nazi Germany named Gerard Salton, who had come to America, got a PhD at Harvard, and moved to Cornell University, where he cofounded its computer science department. Searching through databases using the same commands you'd use with a human—"natural language" became the term of art—was Salton's specialty.

During the 1960s, Salton developed a system that was to become a model for information retrieval. It was called SMART, supposedly an acronym for "Salton's Magical Retriever of Text." The system established many conventions that still persist in search, including indexing and relevance algorithms. When Salton died in 1995, his techniques still ruled the field. "For thirty years," wrote one academic in tribute a year later, "Gerry Salton *was* information retrieval."

The World Wide Web was about to change that, but the academics didn't know it—and neither did AltaVista. While its creators had the insight to gather all of the web, they missed the opportunity to take advantage of the link structure. "The innovation was that I was not afraid to fetch as much of the web as I could, store it in one place, and have a really fast response time. *That* was the novelty," says Monier. Meanwhile, AltaVista analyzed what was on each individual page—using metrics like how many times each word appeared—to see if a page was a relevant match to a given keyword in a query.

Even though there was no clear way to make money from search, AltaVista had a number of competitors. By 1996, when I wrote about search for *Newsweek*, executives from several companies were all boasting the most useful service. When pressed, all of them would admit that in the race between the omnivorous web and their burgeoning technology, the web was winning. "Academic IR had thirty years to get to where it is—we're breaking new ground, but it's difficult," complained Graham Spencer, the engineer behind the search engine created by a start-up called Excite. AltaVista's director of engineering, Barry Rubinson, said that the best approach was to throw massive amounts of silicon toward the problem and then hope for the best. "The first problem is that relevance is in the eye of the beholder," he said. The second problem, he continued, is making sense

of the infuriatingly brief and cryptic queries typed into the AltaVista search field. He implied that the task was akin to voodoo. "It's all wizardry and witchcraft," he told me. "Anyone who tells you it's scientific is just pulling your leg."

No one at the web search companies mentioned using links.

The links were the reason that a research project running on a computer in a Stanford dorm room had become the top performer. Larry Page's PageRank was powerful because it cleverly analyzed those links and assigned a number to them, a metric on a scale of 1 to 10, that allowed you to see the page's prominence in comparison to every other page on the web. One of the early versions of BackRub had simply counted the incoming links, but Page and Brin quickly realized that it wasn't merely the number of links that made things relevant. Just as important was who was doing the linking. PageRank reflected that information. The more prominent the status of the page that made the link, the more valuable the link was and the higher it would rise when calculating the ultimate PageRank number of the web page itself. "The idea behind PageRank was that you can estimate the importance of a web page by the web pages that link to it," Brin would say. "We actually developed a lot of math to solve that problem. Important pages tended to link to important pages. We convert the entire web into a big equation with several hundred million variables, which are the PageRanks of all the web pages, and billions of terms, which are all the links." It was Brin's mathematic calculations on those possible 500 million variables that identified the important pages. It was like looking at a map of airline routes: the hub cities would stand out because of all the lines representing flights that originated and terminated there. Cities that got the most traffic from other important hubs were clearly the major centers of population. The same applied to websites. "It's all recursive," Page later said. "In a way, how good you are is determined by who links to you and who you link to determines how good you are. It's all a big circle. But mathematics is great. You can solve this."

The PageRank score would be combined with a number of more traditional information retrieval techniques, such as comparing the keyword to text on the page and determining relevance by examining factors such as frequency, font size, capitalization, and position of the keyword. (Those factors help determine the importance of a keyword on a given page—if a term is prominently featured, the page is more likely to satisfy a query.) Such factors are known as *signals*, and they are critical to search quality.

There are a few crucial milliseconds in the process of a web search during which the engine interprets the keyword and then accesses the vast index, where all the text on billions of pages is stored and ordered just like an index of a book. At that point the engine needs some help to figure out how to rank those pages. So it looks for signals—traits that can help the engine figure out which pages will satisfy the query. A signal says to the search engine, "Hey, consider me for your results!" PageRank itself is a signal. A web page with a high PageRank number sends a message to the search engine that it's a more reputable source than those with lower numbers.

Though PageRank was BackRub's magic wand, it was the combination of that algorithm with other signals that created the mind-blowing results. If the keyword matched the title of the web page or the domain name, that page would go higher in the rankings. For queries consisting of multiple words, documents containing all of the search query terms in close proximity would typically get the nod over those in which the phrase match was "not even close." Another powerful signal was the "anchor text" of links that led to the page. For instance, if a web page used the words "Bill Clinton" to link to the White House, "Bill Clinton" would be the anchor text. Because of the high values assigned to anchor text, a BackRub query for "Bill Clinton" would lead to www.whitehouse.gov as the top result because numerous web pages with high PageRanks used the president's name to link the White House site. "When you did a search, the right page would come up, even if the page didn't include the actual words you were searching for," says Scott Hassan. "That was pretty cool." It was also something other search engines failed to do. Even though www.whitehouse.gov was the ideal response to the Clinton "navigation query," other commercial engines didn't include it in their results. (In April 1997, Page and Brin found that a competitor's top hit was "Bill Clinton Joke of the Day.")

PageRank had one other powerful advantage. To search engines that relied on the traditional IR approach of analyzing content, the web presented a terrible challenge. There were millions and millions of pages, and as more and more were added, the performance of those systems inevitably degraded. For those sites, the rapid expansion of the web was a problem, a drain on their resources. But because of PageRank, BackRub got *better* as the web grew. New sites meant more links. This additional information allowed BackRub to identify even more accurately the pages that might be relevant to a query. And the more recent links would improve the freshness of the site. "PageRank has the benefit of learning from the whole of the World Wide Web," Brin would explain.

Of course, Brin and Page had the logistical problem of capturing the whole web. The Stanford team did not have the resources of DEC. For a while, BackRub could access only the bandwidth available to the Gates Building—10 megabits of traffic per second. But the entire university ran on a giant T3 line that could operate at 45 megabits per second. The BackRub team discovered that by retoggling an incorrectly set switch in the basement, it could get full access to the T3 line. "As soon as they toggled that, we were all the way up to the maximum of the entire Stanford network," says Hassan. "We were using all the bandwidth of the network. And this was from a single machine doing this, on a desktop in my dorm room."

In those days, people who ran websites—many of them with minimal technical savvy—were not used to their sites being crawled. Some of them would look at their logs, and see frequent visits from www.stanford.edu, and suspect that the university was somehow stealing their information. One woman from Wyoming contacted Page directly to demand that he stop, but Google's "bot" kept visiting. She discovered that Hector Garcia-Molina was the project's adviser and called him, charging that the Stanford computer was doing terrible things to her computer. He tried to explain to her that being crawled is a harmless, nondestructive procedure, but she'd have none of it. She called the department chair and the Stanford security office. In theory, complainants could block crawlers by putting a little piece of code on their sites called /robots.txt, but the angry webmasters weren't receptive to the concept. "Larry and Sergey got annoyed that people couldn't figure out /robots.txt," says Winograd, "but in the end, they actually built an exclusion list, which they didn't want to." Even then, Page and Brin believed in a self-service system that worked in scale, serving vast populations. Handcrafting exclusions was anathema.

Brin and Page fell into a pattern of rapid iterating and launching. If the pages for a given query were not quite in the proper order, they'd go back to the algorithm and see what had gone wrong. It was a tricky balancing act to assign the proper weights to the various signals. "You do the ranking initially, and then you look at the list and say, 'Are they in the right order?' If they're not, we adjust the ranking, and then you're like, 'Oh this looks really good,'" says Page. Page used the ranking for the keyword of "university" as a litmus test. He paid particular attention to the relative ranking of his alma mater, Michigan, and his current school, Stanford. Brin and Page assumed that Stanford would be ranked higher, but Michigan topped it. Was that a flaw in the algorithm? No. "We decided that Michigan had more stuff on the web, and that was reasonable," says Page.

This listing showed the power of PageRank. It made BackRub much more useful than the results you'd get from the commercial search engines. Their list of institutions for the "university" query seemed totally random. The number one result for that generic term in AltaVista would give you the Oregon Center for Optics. Page recalls a conversation back then with an AltaVista engineer who told him that with the way pages were scored, a query for "university" was likely to get a page where that word appeared twice in the headline. "That doesn't make any sense," Page said, noting that such a search was more likely to get a minor university with redundancy in its title.

"If you want major universities, you should type 'major universities,'" said the engineer. Page was appalled. "I'm like, well, they teach you in human computer interaction, which is my branch, that the user is never wrong. The person in the system is never wrong."

Until that moment, the task of compiling a list of universities and ranking them in significance had been complicated, intellectually challenging, and labor-intensive. Some magazines employed large teams working for months to do just that. If you were to try to teach a computer to do that, your instinct would be to feed it data about SAT scores, graduation rates, prizewinners among faculty, and a thousand other factors. Then you'd have to figure out how to weigh them. The odds were low that a machine would crank out a rating that squared with the gut feeling of a well-educated citizen. But BackRub knew nothing about those statistics. It just knew how to take advantage of the fact that links created by the web community had implicitly produced a ranking that was better than any group of magazine editors or knowledge curators could come up with. Larry Page and Sergey Brin had figured out how to mine that knowledge before the information retrieval establishment and commercial search engines even realized that it existed.

"The whole field had suffered blinders," says the computer scientist Amit Singhal, then a Bell Labs researcher who had been a protégé of Gerry Salton. "In some sense, search really did need two people who were never tainted by people like me to come up with that shake-up."

Larry Page was not the only person in 1996 who realized that exploiting the link structure of the web would lead to a dramatically more powerful way to find information. In the summer of that year, a young computer scientist named Jon Kleinberg arrived in California to spend a yearlong postdoctoral fellowship at IBM's research center in Almaden, on the south-

ern edge of San Jose. With a new PhD from MIT, he had already accepted a tenure-track job in the CS department at Cornell University.

Kleinberg decided to look at web search. The commercial operations didn't seem effective enough and were further hobbled by spam. AltaVista's results in particular were becoming less useful because websites had gamed it by "word stuffing"—inserting multiple repetitions of desirable keywords, often in invisible text at the bottom of the web page. "The recurring refrain," says Kleinberg, "was that search doesn't work." But he had an intuition of a more effective approach. "One thing that was not being used at all was the fact that the web was a network," he says. "You could find people saying in the academic papers that links ought to be taken advantage of, but by 1996 it still hadn't been."

Kleinberg began to play around with ways to analyze links. Since he didn't have the assistance, the resources, the time, or the inclination, he didn't attempt to index the entire web for his link analysis. Instead he did a kind of prewash. He typed a query into AltaVista, took the first two hundred results, and then used that subset for his own search.

Interestingly, the best results for the query were often not included in those AltaVista solutions. For instance, if you typed in "newspaper," AltaVista would not give you links for *The New York Times* or *The Washington Post*. "That's not surprising, because AltaVista is about matching strings, and unless *The New York Times* happened to say, 'I'm a newspaper!' AltaVista is not going to find it," Kleinberg explains. But, he suspected, he'd have more luck if he checked out what those 200 sites pointed to. "Among those 200 people who were saying 'newspapers,' someone was going to point to *The New York Times*," he says. "In fact, a *bunch* of people were going to point to *The New York Times*, because among those 200 pages were some people who really liked to collect links for newspapers on the web. If you pulled in those links, and got a set of 5,000 to 10,000 of them, in a sense, you'd have a vote. The winner would be the one with the most inlinks from the group." It was the same lightbulb that had brightened over Larry Page's head.

Sometime in December 1996, Kleinberg got the balance right. One of his favorite queries was "Olympics." The summer games had been held in Atlanta that year, and there were thousands of sites that in some way dealt with the athletic contests, the politics, the bomb that a domestic terrorist had planted. The AltaVista results for that keyword were riddled with spam and were generally useless. But Kleinberg's top result was the official Olympics site.

Kleinberg began showing his breakthrough around IBM. His managers quickly put him in touch with the patent lawyers. Most people took a look at what Kleinberg had set up and wanted him to find stuff for them. Even the patent attorney wanted Kleinberg to help him find sources for his hobby, medieval siege devices. By February 1997, he says, "all sorts of IBM vice presidents were trooping through Almaden to look at demos of this thing and trying to think about what they could do with it." Ultimately, the answer was . . . not much. IBM was a $70 billion business, and it was hard to see how a research project about links on this World Wide Web could make a difference. Kleinberg shrugged it off. He was going to teach computer science at Cornell.

Through mutual friends at Stanford, Kleinberg heard about Larry Page's project, and in July 1997 they met at Page's office in the Gates Building. Kleinberg was impressed with BackRub. "In academia, when there's a hard problem everyone wants to solve, you're always implicitly competing with the other people who are working on it," says Kleinberg. But neither mentioned that issue. Kleinberg encouraged Page to publish his findings, but Page wasn't receptive. "Larry was worried about writing a paper," says Kleinberg. "He was wary because he wanted to see how far he could get with it while he refined it."

Kleinberg could see that his goals were different from Page's. "They wanted to crawl the whole web and get it on racks of servers that they would accumulate," Kleinberg says. "My view was 'How can I solve this problem without having to sink three months into indexing the web?' We had the same core idea, but how we went about it was almost diametrically opposite." Kleinberg was trying to understand network behavior. Page and Brin were *building* something. "Kleinberg had this notion of authority, where your page can become good just by linking to the right pages," says Page. "Whereas what I was doing was more of a traffic simulation, which is actually how people might search the web."

Kleinberg kept up with Google. He turned down job feelers in 1999 and again in 2000. He was happy at Cornell. He'd win teaching awards and a MacArthur fellowship. He led the life in academia he'd set out to lead, and not becoming a billionaire didn't seem to bother him.

There was yet a third person with the idea, a Chinese engineer named Yanhong (Robin) Li. In 1987, he began his studies at Beijing University, an institution that claimed prominence in the country by way of a metric: The Science Citation Index, which ranked scientific papers by the number of

other papers that cited them. The index was used in China to rank universities. "Beijing University, measured by the number of citations its professors got from their papers, was ranked number one," said Li.

Li came to the United States in 1991 to get a master's degree at SUNY Buffalo, and in 1994 took a job at IDD Information Services in Scotch Plains, New Jersey, a division of Dow Jones. Part of his job was improving information retrieval processes. He tried the search engines at the time—AltaVista, Excite, Lycos—and found them ineffectual and spam-ridden. One day in April 1996 he was at an academic conference. Bored by the presentation, he began to ponder how search engines could be improved. He realized that the Science Citation Index phenomenon could be applied to the Internet. The hypertext link could be regarded as a citation! "When I returned home, I started to write this down and realized it was revolutionary," he says. He devised a search approach that calculated relevance from both the frequency of links and the content of anchor text. He called his system RankDex.

When he described his scheme to his boss at Dow Jones, urging the company to apply for a patent, he was at first encouraged, then disappointed when nothing happened. "So a couple of months later, I decided to write the application by myself." He bought a self-help book on patent applications and filed his in June 1996. But when he told his boss, Dow Jones reasserted itself and hired a lawyer to review the patent, which it refiled in February 1997. (Stanford University would not file its patent for Larry Page's PageRank system until January 1998.) Nonetheless, Dow Jones did nothing with Li's system. "I tried to convince them it was important, but their business had nothing to do with Internet search, so they didn't care," he says.

Robin Li quit and joined the West Coast search company called Infoseek. In 1999, Disney bought the company and soon thereafter Li returned to China. It was there in Beijing that he would later meet—and compete with—Larry Page and Sergey Brin.

Page and Brin had launched their project as a stepping-stone to possible dissertations. But it was inevitable that they began to eye their creation as something that could make them money. The Stanford CS program was as much a corporate incubator as an academic institution. David Cheriton, one of the professors, once put it this way: "The unfair advantage that Stanford has over any other place in the known universe is that we're surrounded by Silicon Valley." It was not uncommon for its professors to

straddle both worlds, maintaining posts in the department while playing in the high-tech scrum of start-ups striving for the big score. There was even a joke that faculty members couldn't get tenure until they started a company.

Cheriton himself was a prime example of how the Stanford network launched companies and enriched the founders. One of the earlier gold strikes from Stanford was the founding of Sun Microsystems by a group that included Andy Bechtolsheim, Vinod Khosla, and Bill Joy. Cheriton was close to Bechtolsheim, so in 1995, when the latter decided to start Granite Systems, a networking start-up, the two collaborated. Eighteen months later, Cisco bought the company for $220 million.

Sergey Brin, Rollerblading his way around the corridors of Gates Hall, took notice. Though Brin and Page didn't have classes with Cheriton, they headed to his office for some advice. They specifically wanted to know how they might interest a company into using PageRank in its own search technology. Cheriton told them that it would be difficult—Sun Microsystems, he reminded them, had been started out of frustration when companies had spurned Bechtolsheim's attempts to sell his workstation technology.

Yet Brin and Page were reluctant at that point to strike out on their own. They had both headed to Stanford intending to become PhDs like their dads.

But licensing their search engine wasn't easy. Though Brin and Page had a good meeting with Yahoo founders Jerry Yang and David Filo, former Stanford students, Yahoo didn't see the need to buy search engine technology. They also met with an AltaVista designer, who seemed interested in BackRub. But the wise men back in DEC headquarters in Maynard, Massachusetts, nixed the idea. Not Invented Here.

Maybe the closest Page and Brin came to a deal was with Excite, a search-based company that had begun—just like Yahoo—with a bunch of sharp Stanford kids whose company was called Architext before the venture capitalists (VCs) got their hands on it and degeekified the name. Terry Winograd, Sergey's adviser, accompanied them to a meeting with Vinod Khosla, the venture capitalist who had funded Excite.

That led to a meeting with Excite's founders, Joe Kraus and Graham Spencer, at Fuki Sushi, a Palo Alto restaurant. Larry insisted that the whole BackRub team come along. "He always likes to have more people on his side than the opposite side, to get the upper hand," says Scott Hassan, who

attended along with Page, Brin, and Alan Steremberg. "They sent two people, so we had four." The Excite people began comparison tests with BackRub, plugging in search queries such as "Bob Marley." The results were a lot better than Excite's.

Larry Page laid out an elaborate plan, which he described in detail in emails to Khosla in January 1997. Excite would buy BackRub, and then Larry alone would go to work there. Excite's adoption of BackRub technology, he claimed, would boost its traffic by 10 percent. Extrapolating that in terms of increased ad revenue, Excite would take in $130,000 more every day, for a total of $47 million in a year. Page envisioned his tenure at Excite lasting for seven months, long enough to help the company implement the search engine. Then he would leave, in time for the fall 1997 Stanford semester, resuming his progress toward a doctorate. Excite's total outlay would be $1.6 million, including $300,000 to Stanford for the license, a $200,000 salary, a $400,000 bonus for implementing it within three months, and $700,000 in Excite stock. (Since Page and Brin were working for Stanford while developing their work, the school owned the PageRank patent. Stanford would commonly make financial arrangements so that such inventors could hold exclusive licenses to the intellectual property they created. Eventually Stanford did so with Google, in exchange for 1.8 million shares.) "With my help," wrote the not-quite-twenty-four-year-old student, "this technology will give Excite a substantial advantage and will propel it to a market leadership position."

Khosla made a tentative counteroffer of $750,000 total. But the deal never happened. Hassan recalls a key meeting that might have sunk it. Though Excite had been started by a group of Stanford geeks very much like Larry and Sergey, its venture capital funders had demanded they hire "adult supervision," the condescending term used when brainy geeks are pushed aside as top executives and replaced by someone more experienced and mature, someone who could wear a suit without looking as though he were attending his Bar Mitzvah. The new CEO was George Bell, a former Times Mirror magazine executive. Years later, Hassan would still laugh when he described the meeting between the BackRub team and Bell. When the team got to Bell's office, it fired up BackRub in one window and Excite in the other for a bake-off.

The first query they tested was "Internet." According to Hassan, Excite's first results were Chinese web pages where the English word "Internet" stood out among a jumble of Chinese characters. Then the team typed

"Internet" into BackRub. The first two results delivered pages that told you how to use browsers. It was exactly the kind of helpful result that would most likely satisfy someone who made the query.

Bell was visibly upset. The Stanford product was *too* good. If Excite were to host a search engine that instantly gave people information they sought, he explained, the users would leave the site instantly. Since his ad revenue came from people *staying* on the site—"stickiness" was the most desired metric in websites at the time—using BackRub's technology would be counterproductive. "He told us he wanted Excite's search engine to be 80 percent as good as the other search engines," says Hassan. And we were like, "Wow, these guys don't know what they're talking about."

Hassan says that he urged Larry and Sergey right then, in early 1997, to leave Stanford and start a company. "Everybody else was doing it," he says. "I saw Hotmail and Netscape doing really well. Money was flowing into the Valley. So I said to them, 'The search engine is the idea. We *should do this*.' They didn't think so. Larry and Sergey were both very adamant that they could build this search engine at Stanford."

"We weren't . . . in an entrepreneurial frame of mind back then," Sergey later said.

Hassan quit the project. He got a job with a new company called Alexa and worked part-time on a start-up called eGroups. In fact, Larry and Sergey—this was before they had gotten a dollar in funding for Google—pitched in $5,000 each to help him buy computers for eGroups. (The investment paid off less than three years later when Yahoo bought eGroups for an estimated $413 million.)

But for the next year and a half, all the companies they approached turned them down. "We couldn't get anyone interested," says Page. "We did get offers, but they weren't for much money. So we said, 'Whatever,' and went back to Stanford to work on it some more. It wasn't like we wanted a lot of money, but we wanted the stuff to get really used. And they would want us to work there and we'd ask, 'Do we really want to work for this company?' These companies weren't going to focus on search—they were becoming portals. They didn't understand search, and they weren't technology people."

In September 1997, Page and Brin renamed BackRub to something they hoped would be suitable for a business. They gave serious consideration to "The Whatbox," until they realized that it sounded too much like "wetbox," which wasn't family-friendly. Then Page's dorm roommate suggested they call it "googol." The word was a mathematical term referring to

the number 1 followed by 100 zeros. Sometimes the word "googolplex" was used generically to refer to an insanely large number. "The name reflected the scale of what we were doing," Brin explained a few years later. "It actually became a better choice of name later on, because now we have billions of pages and images and groups and documents, and hundreds of millions of searches a day." Page misspelled the word, which was just as well since the Internet address for the correct spelling was already taken. "Google" was available. "It was easy to type and memorable," says Page.

One night, using a new open-source graphics program called GIMP, Sergey designed the home page, spelling the new company name in different colors, making a logo that resembled something made from children's blocks. It conveyed a sense of amiable whimsy. He put an exclamation point after the name, just like Yahoo, another Internet company founded by two Stanford PhD dropouts. "He wanted it to be playful and young," says Page. Unlike a lot of other web pages, the Google home page was so sparse it looked unfinished. The page had a box to type in requests and two buttons underneath, one for search and another labeled I'm Feeling Lucky, a startling bid of confidence that implied that, unlike the competition, Google was capable of nailing your request on the first try. (There was another reason for the button. "The point of I'm Feeling Lucky was to replace the domain name system for navigation," Page said in 2002. Both Page and Brin hoped that instead of guessing what was the address of their web destination, they'd just "go to Google.") The next day Brin ran around the CS department at Stanford, showing off his GIMP creation. "He was asking everybody whether it made any sense to put other stuff on the page," says Dennis Allison, a Stanford CS lecturer. "And everybody said no." That was fine with Page and Brin. The more stuff on the page, the slower it would run, and both of them, especially Page, believed that speed was of the essence when it came to pleasing users. Page later found it humorous that people praised the design for its Zen-like use of white space. "The minimalism is that we didn't have a webmaster and had to do it ourselves," he says.

Meanwhile, BackRub-turned-Google was growing to the point where it was difficult to run using Stanford's facilities. It was becoming less a research project than an Internet start-up run from a private university. Page and Brin's reluctance to write a paper about their work had become notorious in the department. "People were saying, 'Why is this so secret? This is an academic project, we should be able to know how it worked,'" says Terry Winograd.

Page, it seemed, had a conflict about information. On one hand, he subscribed heartily to the hacker philosophy of shared knowledge. That was part of what his project was all about: making human knowledge accessible, making the world a better place. But he also had a strong sense of protecting his hard-won proprietary information. He remembered Nikola Tesla, who had died in poverty even as his inventions enriched others. Later, there would be speculation whether Page, a private person to begin with, had pulled back a little more after his father's death in June 1996. Scott Hassan recalls that the team conveyed its condolences to Page that month, but Hassan didn't speak much about the loss with Page. "Mostly we talked about technical stuff," he would recall. Mike Moritz, one of the venture capitalists who would fund Google, later surmised that "a large part" of Page's later wariness could be associated with that loss. "He felt that the world was pulled out from underneath him," Moritz said. "It makes it hard to trust anything again."

But it wasn't just the secrecy that stalled Brin and Page. Writing a paper wasn't as interesting to them as building something. "Inherently, Larry and Sergey aren't paper-oriented—they're product-oriented," says Winograd. "If they have another ten minutes, they want to make something better. They don't want to take ten minutes to tell you something they did." But finally Winograd convinced them to explain PageRank in a public forum. They presented a paper called "The Anatomy of a Large-Scale Hypertextual Web Search Engine" at a conference in Australia in May 1998.

Arthur Clarke once remarked that the best technology was indistinguishable from magic. The geeks of Silicon Valley, assuming he was talking about them, have never forgotten that and have invoked the quote in countless press releases about their creations. But Google search really did feel like magic. At Stanford, Larry's and Sergey's professors and friends were using the search engine to answer questions and telling their friends about it. Google was handling as many as 10,000 queries a day. At times it was consuming half of Stanford's Internet capacity. Its appetite for equipment and bandwidth was voracious. "We just begged and borrowed," says Page. "There were tons of computers around, and we managed to get some." Page's dorm room was essentially Google's operations center, with a motley assortment of computers from various manufacturers stuffed into a homemade version of a server rack—a storage cabinet made of Legos. Larry and Sergey would hang around the loading dock to see who on campus was getting computers—companies like Intel and Sun gave lots of free machines

to Stanford to curry favor with employees of the future—and then the pair would ask the recipients if they could share some of the bounty.

That still wasn't enough. To store the millions of pages they had crawled, the pair had to buy their own high-capacity disk drives. Page, who had a talent for squeezing the most out of a buck, found a place that sold refurbished disks at prices so low—a tenth of the original cost—that something was clearly wrong with them. "I did the research and figured out that they were okay as long as you replaced the [disk] operating system," he says. "We got 120 drives, about nine gigs each. So it was about a terabyte of space." It was an approach that Google would later adopt in building infrastructure at low cost.

Larry and Sergey would be sitting by the monitor, watching the queries—at peak times, there would be a new one every second—and it would be clear that they'd need even more equipment. *What next?* they'd ask themselves. *Maybe this is real.*

Stanford wasn't kicking them out—the complications of running the nascent Google were outweighed by pride that something interesting was brewing in the department. "It wasn't like our lights were dimming when they would run the crawler," says Garcia-Molina, who was still hoping that Larry and Sergey would develop their work academically. "I think it would have made a great thesis," he says. "I think their families were behind them to get PhDs, too. But doing a company became too much of an attraction."

There was no alternative; no one would pay enough for Google. And the happy visitors they were attracting gave them confidence that their efforts could make a difference. After years of dreaming how his ideas could change the world, Larry Page realized that he'd done something that might do just that. "If the company failed, too bad," says Page. "We were really going to be able to do something that *mattered*."

They went back to Dave Cheriton, who encouraged them to just get going. "Money shouldn't be a problem," he said. Cheriton suggested that they meet with Andy Bechtolsheim. Brin dashed off an email to Bechtolsheim that evening around midnight and got an immediate reply asking if the two students could show up at eight the next morning at Cheriton's house, which was on the route Bechtolsheim used to go to work each day. At that ungodly hour Page and Brin demoed their search engine for Bechtolsheim on Cheriton's porch, which had an ethernet connection. Bechtolsheim, impressed but eager to get to the office, cut the meeting short by offering to write the duo a $100,000 check.

"We don't have a bank account yet," said Brin.

"Deposit it when you get one," said Bechtolsheim, who raced off in his Porsche. With as little fanfare as if he were grabbing a latte on the way to work, he had just invested in an enterprise that would change the way the world accessed information. Brin and Page celebrated with a Burger King breakfast. The check remained in Page's dorm room for a month.

Soon afterward, Bechtolsheim was joined by other angel investors, including Dave Cheriton. One was a Silicon Valley entrepreneur named Ram Shriram, whose own company had recently been purchased by Amazon.com. Shriram had met Brin and Page in February 1998; although he had been skeptical about a business model for search engines, he was so impressed with Google that he had been advising them. After the Bechtolsheim meeting, Shriram invited them to his house to meet his boss Jeff Bezos, who was enthralled with their passion and "healthy stubbornness," as they explained why they would never put display ads on their home page. Bezos joined Bechtolsheim, Cheriton, and Shriram as investors, making for a total of a million dollars of angel money.

On September 4, 1998, Page and Brin filed for incorporation and finally moved off campus. Sergey's girlfriend at the time was friendly with a manager at Intel named Susan Wojcicki, who had just purchased a house on Santa Margarita Street in Menlo Park with her husband for $615,000. To help meet the mortgage, the couple charged Google $1,700 a month to rent the garage and several rooms in the house. At that point they'd taken on their first employee, fellow Stanford student Craig Silverstein. He'd originally connected with them by offering to show them a way to compress all the crawled links so they could be stored in memory and run faster. ("It was basically to get my foot in the door," he says.) They also hired an office manager. But almost as if they were still hedging on their PhDs, they maintained a presence at Stanford that fall, coteaching a course, CS 349, "Data Mining, Search, and the World Wide Web," which met twice a week that semester. Brin and Page announced it as a "project class" in which the students would work with the repository of 25 million web pages that they had captured as part of what was now a private company. They even had a research assistant. The first assigned reading was their own paper, but later in the semester a class was devoted to a comparison of PageRank and Kleinberg's work.

In December, after the final projects were due, Page emailed the students a party invitation that also marked a milestone: "The Stanford Research Project is now Google.com: The Next Generation Internet Search Company."

"Dress is Tiki Lounge wear," the invitation read, "and bring something for the hot tub."

2

"We want Google to be as smart as you."

Larry Page did not want to be Tesla'd. Google had quickly become a darling of everyone who used it to search the net. But at first so had AltaVista, and that search engine had failed to improve. How was Google, led by two talented but inexperienced youngsters, going to tackle the devilishly difficult problems of improving its service?

"If we aren't a lot better next year, we will already be forgotten," Page said to one of the first reporters visiting the company.

The web was growing like digital kudzu. People were coming to Google in droves. Google's plan was to get even more traffic. "When we started the company, we had two computers," says Craig Silverstein. "One was the web service, and one was doing everything else—the page rank, the searches. And there was a giant chain of disks that went off the back of the computer that stored twenty-five million web pages. Obviously that was not going to scale very well." Getting more computers was no problem. Google needed brainpower, especially since Brin and Page had reached the limits of what they could do in writing the software that would enable the search engine to grow and improve. "Coding is not where their interests are," says Silverstein.

The founders also knew that Google had to be a lot smarter to keep satisfying users—and to fulfill the world-changing ambitions of its founders. "We don't always produce what people want," Page explained in Google's early days. "It's really difficult. To do that you have to be smart—you have to understand everything in the world. In computer science, we call that artificial intelligence."

Brin chimed in. "We want Google to be as smart as you—you should be getting an answer the minute you think of it."

"The ultimate search engine," said Page. "We're a long way from that."

Page and Brin both held a core belief that the success of their company would hinge on having world-class engineers and scientists committed to their ambitious vision. Page believed that technology companies

can thrive only by "an understanding of engineering at the highest level." Somehow Page and Brin had to identify such a group and impress them enough to have them sign on to a small start-up. Oh, and they had a policy that limited the field: no creeps. They were already thinking of the culture of their company and making sure that their hires would show traits of hard-core wizardry, user focus, and starry-eyed idealism.

"We just hired people like us," says Page.

Some of Google's early hires were simply brainy recent grads, people like Marissa Mayer, a hard-driving math whiz and ballet dancer in her high school in Wausau, Wisconsin, who had become an artificial intelligence star at Stanford. (During her interview with Silverstein, she was asked for three things Google could do better; ten years later, she was still kicking herself that she listed only two.) But Page and Brin also went after people with résumés more often seen in the recruitment offices of Microsoft Research or Carnegie Mellon's CS department. One of their first coups was a professor at the University of California at Santa Barbara named Urs Hölzle. He'd played with the earlier crop of search engines such as AltaVista and Inktomi and concluded that, as a computer scientist familiar with Boolean syntax and other techniques, he could use those techniques to find what he wanted on the Internet. But he assumed that search would never be something his mother would use. Google instantly changed his mind about that: you just typed in what you wanted, and, bang, the first thing was right. Mom would like that! "They definitely seemed to know what they were doing," he says of Larry and Sergey.

More important to him, when he visited the new company in early 1999, he understood that though he had no background in information retrieval, the problems Brin and Page were working on had a lot in common with his own work in big computer systems. This little search engine was butting up against issues in performance and scalability that only huge projects had previously grappled with. That was Google's secret weapon to lure world-class computer scientists: in a world where corporate research labs were shutting down, this small start-up offered an opportunity to break ground in computer science.

Hölzle, still wary, accepted the offer but kept his position at UCSB by taking a yearlong leave. He would never return. In April he arrived at Google with Yoshka, a big floppy Leonberger dog, in tow, and dived right in to help shore up Google's overwhelmed infrastructure. (By then Google had moved from Wojcicki's Menlo Park house to a second-floor office over a bicycle shop in downtown Palo Alto.) Though Google had a hundred

computers at that point—it was buying them as quickly as it could—it could not handle the load of queries. Hundreds of thousands of queries a day were coming in.

The average search at that time, Hölzle recalls, took three and a half seconds. Considering that speed was one of the core values of Page and Brin—it was like motherhood, and scale was apple pie—this was a source of distress for the founders. "Basically during the middle of the day we were maxed out," says Hölzle. "Nothing was happening for some users, because it would just never get a page basically back. It was all about scalability, performance improvements." Part of the problem was that Page and Brin had written the system in what Hölzle calls "university code," a nice way of saying amateurish. "The web server couldn't handle more than ten requests or so a second because it was written in Python, which is a great idea for a research system, but it's not a high-performance solution," he says. He immediately set about rewriting the code.

Hölzle was joined by other computer scientists who were more daring in taking the leap to permanent Google employment. This included a minimigration of engineers from DEC's research division. Established legend in Silicon Valley cited Xerox's Palo Alto Research Center (PARC) as the canonical lab brimming with breakthrough innovation that had been misunderstood, buried, or otherwise fumbled by the clueless parent company. (Its inventions included the modern computer interface with windows and file folders.) But when it came to missed opportunities, PARC had nothing on DEC's Western Research Laboratory, which was handed over to Compaq when that personal computer company bought Digital Equipment Corporation in 1998. (In 2002, Hewlett-Packard would acquire Compaq.) In 1998, two years before Apple even began work on the iPod, DEC engineers were developing a digital music player that could store a whole music collection and fit in your pocket. In addition, DEC had some of the founding fathers of the Internet, as well as scientists writing pioneering papers on network theory. But DEC never used its engineers' ideas to help AltaVista become Google. ("From the moment I left DEC, I never used AltaVista," says Louis Monier, who split in 1998. "It was just pathetic. It was completely obvious that Google was better.") So it was little wonder that some of them went to Google. "The number [of former DEC scientists at Google] is really kind of staggering," says Bill Weihl, a DEC refugee who came to the company in 2004.

One of the DEC engineers had already independently discovered the power of web links in search. Jeffrey Dean suspected that it would be help-

ful to web users if a software program could point them to pages that were related to the ones they liked. In his vision, you would be reading an article in *The New York Times* and his program would pop up, asking if you'd like to see ten other interesting pages related to the one you were reading.

Dean had never been much interested in information retrieval. Now that he suspected a revolution was afoot, he was. But his attempts to join up with the AltaVista crew ended ignominiously. "The AltaVista team had grown really fast," he says, "and hired a bunch of people who I think were not as technically good as they could have been." In other words—*get me away from here*. In February 1999, Dean bailed from DEC to join a start-up called mySimon.

Within a few months, though, he was bored. Then he heard that Urs Hölzle, whom he'd known through his grad school adviser, had joined up with the guys who did PageRank. "I figured Google would be better because I knew more of the people there, and they seemed like they were more technically savvy," he says. He was so excited about working there that even though his official starting date wasn't until August 1999, in July he began coming to Google after his workday at mySimon ended.

Dean's hiring got the attention of another DEC researcher, Krishna Bharat. He had also been thinking of ways to get web search results from links. Bharat was working on something called the Hilltop algorithm, which algorithmically identified "expert sites" and used those to point to the most relevant results. It was something like Jon Kleinberg's hub approach, but instead of using AltaVista as a prewash to get top search results and then figure out who the expert sites were, Bharat went straight to a representation of the web—links and some bits from the pages—stored in computer memory. Bharat's algorithms would roam around the "neighborhood of the query" to find the key sites.

The India-born computer scientist had already been on Google's radar: when he ate lunch at a joint called World Wraps in Palo Alto, he'd run into Sergey Brin, who would invariably hand him a business card and urge him to apply to Google. Bharat was impressed with Google—he'd actually presented his Hilltop algorithm in the same session at the conference in Australia when Brin and Page showed off Google to a bowled-over audience of IR people. He also liked Sergey. Their mutual friend Rajeev Motwani once hosted a seminar where Brin had arrived on Rollerblades and began rhapsodizing about PageRank without missing a beat. Bharat thought that was incredibly cool. But Google was so *small*. It was hard for