Google Gets A Second Brain, Changing Everything About Search

Wade Roush12/12/12Follow @wroush

Share and Comment

In the 1983 sci-fi/comedy flick The Man with Two Brains, Steve Martin played Michael Hfuhruhurr, a neurosurgeon who marries one of his patients but then falls in love with the disembodied brain of another woman, Anne. Michael and Anne share an entirely telepathic relationship, until Michael’s gold-digging wife is murdered, giving him the opportunity to transplant Anne’s brain into her body.

Well, you may not have noticed it yet, but the search engine you use every day—by which I mean Google, of course—is also in the middle of a brain transplant. And, just as Dr. Hfuhruhurr did, you’re probably going to like the new version a lot better.

You can think of Google, in its previous incarnation, as a kind of statistics savant. In addition to indexing hundreds of billions of Web pages by keyword, it had grown talented at tricky tasks like recognizing names, parsing phrases, and correcting misspelled words in users’ queries. But this was all mathematical sleight-of-hand, powered mostly by Google’s vast search logs, which give the company a detailed day-to-day picture of the queries people type and the links they click. There was no real understanding underneath; Google’s algorithms didn’t know that “San Francisco” is a city, for instance, while “San Francisco Giants” is a baseball team.

Now that’s changing. Today, when you enter a search term into Google, the company kicks off two separate but parallel searches. One runs against the traditional keyword-based Web index, bringing back matches that are ranked by statistical relevance—the familiar “ten blue links.” The other search runs against a much newer database of named entities and relationships.

Type in the query “Philadelphia,” and this second search will produce a new “knowledge panel” in the right-hand margin of the results page, complete with a map and other basic facts about the city William Penn founded. (Hedging its bets, however, Google will also include a thumbnail of the movie poster from the 1993 Tom Hanks film Philadelphia.) To use Google’s own description, the new database helps the search engine understand “things, not strings.”
This second brain is called the Knowledge Graph. English speakers began to see the newly supercharged searches that it powers back in May, and last week the service was rolled out to speakers of seven other languages. But the story behind the knowledge panels goes back to mid-2010, when Google bought a San Francisco search startup called Metaweb Technologies and decided to use its massive semantic database, called Freebase, as the nucleus for its own project to approximate the way humans understand the world.

Metaweb’s creation doesn’t boil down to a collection of guesses about how documents are related to one another, the way Google’s other databases do. Rather, it’s a human-curated encyclopedia of verified facts about things in the world and relationships between them—more than 570 million things and 3.5 billion relationships, at last count. (Philadelphia is a city that’s part of a state that’s part of a nation; it has a known population, a typical set of weather patterns, et cetera.)

While the knowledge panels are the most visible manifestation of the Knowledge Graph, the new information is helping to order and rationalize almost everything else Google does. The consequences will be sweeping. While true AI is still a long way off, the Knowledge Graph represents a shortcut to a time when software will be better at meeting, or even anticipating, our information needs. In essence, Google’s engineers are building toward a future when the company’s famous “I’m Feeling Lucky” option is all you need, and the search engine returns the right result the first time, every time.

Amit Singhal, Google's top search guru

“This is a baby step toward having an understanding computer,” says Amit Singhal, a senior vice president of engineering at Google and the man with ultimate responsibility for improving Google’s core search algorithms. “Now, when you encounter the letters T-A-J-M-A-H-A-L on any Web page, the computers suddenly start understanding that this document is about the monument, and this one is about the musician, and this one is about a restaurant. That ‘aboutness’ is foundational to building the search of tomorrow.”

In a recent interview with Fortune, Google CEO Larry Page said he’s been pushing for such changes for the last 10 years. “The perfect search engine would really understand whatever your need is,” Page said. “It would understand everything in the world deeply, give you back kind of exactly what you need.”
Of course, Google (NASDAQ: GOOG) isn’t just a search engine—it’s also an advertising marketplace that generated $37 billion in revenue last year, and a media platform (YouTube), and a provider of cloud computing services (Gmail and Google Drive), and the leading maker of browser software (Chrome) and mobile operating systems (Android). Having a search engine that understands “aboutness” at its center will alter this whole empire. There are few hints so far of exactly how, but the changes will likely be at least as far-reaching as previous large-scale overhauls of the company’s core technology.

Principal engineer Shashi Thakur compares the Knowledge Graph project to the introduction of universal search in 2007. That change, which made it possible for the first time for users to search Web pages, videos, maps, images, and books from a single unified interface, resulted in a huge boost to Google’s overall traffic. Ultimately, the Knowledge Graph project could have an even “larger possible strategic impact” than universal search, Thakur says.

**Statistical Magic Hits a Limit**

In the search business, you couldn’t ask for a better pedigree than Amit Singhal’s. He studied at Cornell with information retrieval pioneer Gerry Salton, who’s been called “the father of digital search” and was himself a student of Howard Aiken, the Harvard professor who designed IBM’s first computer in 1944. After Google recruited Singhal away from AT&T Labs in 2000, his first job was to rewrite Sergey Brin’s original ranking algorithms to go beyond PageRank and take many new types of relevancy signals into account. The improvements were so dramatic that Singhal was later made a Google Fellow and awarded a prize “in the millions of dollars,” according to journalist Steve Levy’s account of Google’s early years, *In the Plex*.

Still, despite these accomplishments, Singhal says the history of search is basically one big kludge designed to simulate actual human understanding of language.

“The compute power was not there and various other pieces were not there, and the most effective way to search ended up being what today is known as keyword-based search,” Singhal explains. “You give us a query, we find out what is important in that query, and we find out if those important words are also important in a document, using numerous heuristics. This process worked incredibly well—we built the entire field of search on it, including every search company you know of, Google included. But the dream to actually go farther and get closer to human understanding was always there.”

After his initial rewrite of Google’s relevance algorithms, Singhal went on to tackle other problems like morphological analysis: figuring out how to reduce words like “runner” and “running” to their roots (“run,” in this case), in order to perform broader searches, while at the same time learning how to sidestep anomalies (apple and Apple obviously come form the same root, but have very different meanings in the real world). Universal search came next, then autocomplete and Google Instant, which begins to return customized search results even before a user finishes typing a query. (Type “wea,” for example, and you’ll get a local weather forecast.)
“But throughout this process, one thing always bothered us,” Singhal says. “It was that we didn’t ever represent the real world properly in the computer. It was still all a lot of statistical magic, built on top of runs of letters. Even though it almost looked like an incredibly intelligent computer, and we did it far better than anyone, the truth was it was still working on strings of letters.”

This frustration wasn’t just a matter of intellectual aesthetics. Singhal says that by 2009 or 2010, Google had run up against a serious barrier. The goal of the company’s search engineers had always been to connect users with the information they need as efficiently as possible. But for a large group of ambiguous search terms, statistical correlations alone couldn’t help Google intuit the user’s intent. Take Singhal’s favorite example, Taj Mahal. Is the user who types that query searching for the famous mausoleum in Uttar Pradesh (Singhal’s home state), the Grammy-winning blues musician, or the Indian restaurant down the street? Google’s engineers realized that using statistics alone, “we would never be able to say that one of those [interpretations] was more important than the other,” Singhal says.

“I’m very proud of what we achieved using the statistical method, and we still have huge components of our system that are built upon that,” Singhal says. “But we couldn’t take that to the system that we would all want five years from now. Those statistical matching approaches were starting to hit some fundamental limits.”

What Google needed was a way to know more about all of the world’s Taj Mahals, so that it could get better at guessing which one a user wants based on other contextual clues such as their location. And that’s where Metaweb comes into the story. “They were on this quest to represent real-world things, entities, and what is important, what should be known about them,” says Singhal. When Google came across the startup, it had just 12 million entities in its database, which Singhal calls “a toy” compared to real world. “But we saw the promise in the representation technology, and the process they had built to scale that to what we really needed to build a representation of the real world.”

The Database of Everything

Metaweb Technologies has a fascinating history of its own. The company was born as a 2005 spinoff of Applied Minds, the Glendale, CA-based consulting firm and invention factory founded five years before by former Disney R&D head Bran Ferren and former Thinking Machines CEO Danny Hillis. John Giannandrea, a director of engineering at Google who was Metaweb’s chief technology officer, says the idea behind the startup was to build “a machine-readable encyclopedia” to help computers mimic human understanding.

“If you and I are having a conversation, we share a vocabulary,” says Giannandrea, who came to Metaweb after CTO roles at Tellme Networks and Netscape/AOL. “If I say ‘fiscal cliff,’ you know what I mean, and the reason is that you have a dictionary in your head about ideas. Computers don’t have that. That’s what we set about doing.”

The knowledge base Metaweb built is called Freebase, and it’s still in operation today. It’s a collaborative database—technically, a semantic graph—that grows through the contributions of
volunteers, who carefully specify the properties of each new entity and how it fits into existing knowledge categories. (For example, Freebase knows that Jupiter is an entity of type Planet, that it has properties such as a mean radius of 69,911 km, and that it is the fictional setting of two Arthur C. Clarke novels.) While Freebase now hosted by Google, it’s still open to submissions from anyone, and the information in it can be freely reused under a Creative Commons license. In fact, Microsoft uses Freebase to give its Bing search engine an understanding of entities, which is the same role now played by the Knowledge Graph at Google.

Freebase bears some of the markings of earlier knowledge bases such as Cyc, a project begun almost three decades ago by AI researcher Doug Lenat to build a comprehensive ontology of common-sense knowledge. But Giannandrea is careful to point out that Metaweb wasn’t trying to build an AI system. “We explicitly avoided hard problems about reasoning or complicated logic structures,” he says. “We just wanted to build a big enough data set that it could be useful. A lot of these vocabularies and ontologies, they don’t cover roller coasters, they don’t know how to represent the ingredients in food or the recipe for a cocktail. We wanted to cover all of it.”

One implication of covering “all of it” was that Metaweb had to break away from the classic relational-database model, in which data is stored in orderly tables of rows and columns, and build its own proprietary graph database. In a semantic graph, there are no rows and columns, only “nodes” and “edges,” that is, entities and relationships between them. Because it’s impossible to specify in advance what set of properties and relationships you might want to assign to a real-world entity (what’s known in database lingo as the “schema”), graph databases are far better than relational databases for representing practical knowledge.

John Giannandrea, director of engineering at Google and former CTO of Metaweb Technologies

“Suppose you have people and the schema is the place they were born and the date of birth, and you have a million people,” explains Giannandrea. “Now you want to add date of death. Changing the schema after the people are loaded: traditional databases aren’t very good at that.
That’s why semantic graphs are very powerful—you can keep coming up with new kinds of edges.”

To fill up its knowledge base, Metaweb didn’t rely just on volunteers. It also looked for public databases that it could suck up—Wikipedia, for example, and the CIA World Factbook and the MusicBrainz open music database. “We added entities any way we could,” Giannandrea says.

The real challenge for the startup was weeding out the duplicate entities. In a semantic graph, an entity can only be represented once, or everything falls apart. “The process of reconciliation is the key, hard, expensive thing to do,” Giannandrea says. To help pay for that, Metaweb built and sold software tools drawing on Freebase that partners could use to make their own information products more useful. The Wall Street Journal, for example, hired Metaweb to build a database to help its readers pivot between different types of related content.

By the time Google came knocking in 2010, Metaweb and outside contributors had spent five years loading entities into Freebase. The search giant’s acquisition offer was attractive, Giannandrea says, in part because adding to the database was becoming more difficult. Metaweb couldn’t ingest all the world’s knowledge at once, but it didn’t know which sources were most important. “We wanted to make it easier for people to find stuff,” he says. “But that problem is harder if you don’t know what people are looking for. And one of the things a search engine has a good understanding of is what people are trying to figure out. That can help with the prioritization.”

Freebase has doubled in size to about 24 million entities since Google acquired Metaweb. But the Knowledge Graph—which Freebase helped to nucleate—has grown much faster, shooting past half a billion entities in less than three years. There are two reasons for this rapid growth, Giannandrea says. One is that Google itself owns huge databases of real-world things like products (Google Catalogs) and geographical locations (Google Maps). “There is lots and lots of data at Google, not all of which we can free up, but a lot of it, which explains why the Knowledge Graph is much, much larger than the original Freebase,” Giannandrea says. The other reason is Google’s search logs—its real-time picture of what people are searching for, in every country where Google is available, which helps Giannandrea’s team decide which corners of the Knowledge Graph it needs to fill out next.

Being part of Google has brought a few technical advantages as well. With help from Singhal’s team, the Metaweb engineers have been able to improve the algorithms that pull new data into the Knowledge Graph and vet it for accuracy. Not every new entity is reviewed by a human—with about 40 times as many entities as Wikipedia, the Knowledge Graph is far too large for that—but Google has developed quality-assurance systems that let human workers sample a statistically significant fraction of them for precision.

At the same time, the reconciliation problem gets easier with scale. “If I say to you there is a person called Harrison Ford, you couldn’t reconcile that because your database might have 10 Harrison Fords,” says Giannandrea. “But if I said he was a movie actor you’d get closer. Then if I said he was born in a certain year, you’d say okay, fine.” The same principle applies to aliases...
like “the Fresh Prince,” which is really the same entity as Will Smith. The more facts the Knowledge Graph contains, in other words, the easier it is to eliminate duplication.

But is 570 million entities enough to build a working representation of the world—or is Google still just getting started? “I think it’s a lot,” Giannandrea says. For comparison, he points to biologist E.O. Wilson’s Encyclopedia of Life project, which contains listings for 2 million species so far, on its way to 10 million at the most. “That is not a billion things, that is 10 million,” says Giannandrea. By the same token, “There are thousands of models of cars, not billions. There are a lot of sub-genres of Mexican restaurants, but you would be hard pressed to come up with more than 100.”

Google’s oft-repeated mission is to organize all the world’s information, and if you listen to Giannandrea long enough, you start to realize that the company really means it. The startling truth is that for Google, reducing our whole world to a semantic graph isn’t a storage challenge or a computing challenge. The only real question is knowing how many entities is enough for practical purposes, and recognizing when the project has reached a point of diminishing returns. In fact, the Knowledge Graph is already so large that its growth curve is beginning to flatten out, Giannandrea says.

“I do think there is such a thing has having enough knowledge to be useful,” he says. “If you’ve already got every written work, every product, and every populated place on Earth represented, a lot of the big items are covered. There are certainly more things in the Knowledge Graph than I will ever know about in my lifetime. So it’s already big enough for one person.”

**Doodling with the Knowledge Panel**

Of course, to be useful across Google’s network, the Knowledge Graph needs to be big enough for *everyone*. That’s why the company has been putting a lot of effort this year into internationalizing the graph. That’s a twofold problem, according to Shashi Thakur. There’s the language aspect—every entity in the graph needs to be associated with its correct name, not just in English but in French, Italian, Japanese, German, Russian, Spanish, and Portuguese. But just as crucially, there’s a local aspect. Queries against the Knowledge Graph often need to be interpreted differently depending on the user’s location. If you type in “Corinthians” from the U.S., for example, you’re probably searching for information about the seventh book of the New Testament, so that’s what you’ll see in the knowledge panel. But in Brazil, you’ll see links to Sport Club Corinthians, the famous São Paulo-based soccer team.

Internationalization is just one of the problems Google has had to tackle to make the Knowledge Graph useful. Another is how to present the information in the first place. At the moment, Knowledge Graph results can take one of three forms:

1. If your question has a definite, simple answer, the result will often appear in the form of a card at the very top of the regular search results. Say you’re meeting a friend at the airport and you’re wondering when their flight will arrive. Type in the flight number, and Google will show you the arrival time in big type, along with a simple graphic showing the flight’s progress.
2. Most often, Knowledge Graph results show up in the form of a knowledge panel to the right of the regular search results. The panel may include factoids about your search term, maps, photos, event listings, or links and thumbnails for related entities in the graph (the query “Vatican” produces links for “Pope,” “Holy See,” “Vatican Museum,” and “Vatican City”). A knowledge panel will appear whenever “we have high confidence that we know what you’re searching for,” according to Giannandrea.

3. If the best answer to your query is a list rather than a specific entity, you’ll see what Google calls a “carousel”—a scrolling row of thumbnail images. To see an example of a carousel, go to Google and try a search term like “Ernest Hemingway books” or “Leonardo da Vinci paintings.” Thakur calls the carousel “a more elaborate answer to a more complex question” and says it’s often the best presentation when the results call for interactivity and exploration.

Shashi Thakur, a principal engineer at Google, leads the team integrating the Knowledge Graph into Google's core search product.

That’s all Google has rolled out publicly so far, but in the future, results derived from the Knowledge Graph will take many other forms, says Thakur. “There are lots of things people are doodling around with right now,” he says. “As the graph gets deeper and richer and broader, with more connections and more topics, that enables many more applications.”

As an example, Thakur says certain types of searches could yield “fancy infographics”—look up Saturn, and it might appear as part of an interactive graphic of the solar system. Or search “universities in California” and you might see all of them arrayed on a map. (Right now you see their seals or mascots in a carousel, which probably isn’t the most natural way to present this particular list.) The possibilities multiply when you consider how many Google searches are now
initiated from smartphones and tablets, which can accommodate more kinds of inputs and outputs, including speech.

Over time, Google users should expect to Knowledge Graph results popping up more and more often. “A user should learn to depend on the knowledge panel just existing,” says Thakur. “The biggest reason we trigger it some times and not other times is because we want to be really careful not to dilute the experience, and trigger it only in cases where we offer utility. But as the size of the graph grows, there will be increasing areas of the query space where we can show you that utility.”

And the Knowledge Graph isn’t just about search—it’s utility is already being felt in less overt ways in other Google products, Thakur says. If you have a television with Google TV software built in, you can call up video content with spoken commands; behind the scenes, it’s the Knowledge Graph that matches those commands with actual shows, channels, and Web content. Google Now, a Siri-like service that displays just-in-time information cards on Android phones, is also powered by the Knowledge Graph. And the applications will only multiply as more teams at Google figure out how their own products can benefit from the huge database, Thakur says. “The Knowledge Graph is general backbone for representing knowledge, and is a service to the entire company,” he says.

The Perfect Intelligent Assistant

As everyone knows, there’s a single revenue engine that drives Google’s whole universe of activities, from smartphones to robot cars to the $300 million it puts aside every year for Google Ventures, its in-house venture firm. That engine is AdWords, the program that lets advertisers bid to place cost-per-click text ads in the right-hand column of search result pages for specific keywords. AdWords accounts for about 70 percent of Google’s overall advertising revenue, while AdSense, which places ads on partners’ sites, brings in the other 30 percent.

In a situation like this, you might think the company would be allergic to any new idea that changes the way AdWords ads are displayed. But this is exactly what the new knowledge panels do. In fact, they sometimes push the ads so far down the right column that users must scroll or click to the next page to see them.

That has some AdWords advertisers displeased. “The Knowledge Graph released on us by Google seems to take up the right top ad space,” one advertiser complained in an official Google AdWords forum. “I’m paying a lot for AdWords and I think Google should have notified advertisers that search results pages were changing.”

But the Google search engineers I talked to say they have license to improve Google’s core search technology without regard to the potential effects on advertising. “We are building our dream search engine,” says Amit Singhal. “We are guiding the company toward the best experience for our users, and we are not really paying attention to whether people will click more on ads or less on ads.”
Lord knows there’s a good argument for letting Google engineers do their thing. Long before universal search and Google Instant and the Knowledge Graph, Google had superior ranking algorithms, which is how the company stole users away from older search engines and built a user base worth monetizing.

But it’s still breathtaking to visit a company that puts so much faith in innovation. The attitude seems to be: Yes, we might break something, but if we do, we’ll fix it so that it’s even better. After all, it’s possible that the Knowledge Graph could end up improving the algorithms Google uses to rank the relevancy of ads, perhaps leading to higher click-through rates and even greater revenues. “In today’s search system we see today’s design of ads, and tomorrow you will see tomorrow’s design of ads, and we don’t fully know what that is,” says Singhal.

Google’s main goal is to get users to the information they want faster so that they can carry on with their lives, Singhal says. Even with the Knowledge Graph, that isn’t happening fast enough for his taste. “It should be as simple as me saying, ‘Google, drive me to my dinner appointment,’” he says. “Which we can now literally do [thanks to the driverless car project]. But search still has a lot of friction between you and what you need, and it is not as magical as it should be. For it to get there, we will need to solve numerous small problems.”

These “small” problems include better speech recognition, spoken-word interfaces, computer vision, natural language understanding, machine translation, and contextual awareness—all tough computer-science challenges that legions of PhDs at Google are gradually chipping away at. The results will be seen not just in the classic desktop search scenario but also in Google’s mobile interfaces, such as Google Now, which anticipates your questions based on your location (automatically retrieving a boarding pass as you approach the airport, for instance).

“What you are starting to see is the emergence of this dual paradigm,” says Singhal. “Google is becoming the perfect intelligent assistant by your side, which can not only answer when you ask but proactively tell you things that are really important.”

Eventually, Singhal says, Google’s two brains will merge back into one, as the company figures out how to marry the Knowledge Graph with its vast keyword-based index of the Web. At a high level, the task will involve matching listings in the index with entities in the graph, so that Google’s whole picture of the world becomes more entity-based. “I can easily see that future” where there’s only one search going on under the hood at Google, rather than two, Singhal says. “These systems are certainly becoming more interdependent, and as you can imagine a lot of interdependency goes forward and forms integration.”

Singhal and many of Google’s other top engineers are old enough to have grown up watching Star Trek, and you can’t talk with them for very long without hearing a mention of the ship’s computer, which was seemingly ever-present and all-knowing. In fact, when Google Now was in development it went by the code name Majel, a reference to the late actress Majel Barrett, who supplied the voice of the computer systems in Star Trek: The Next Generation and other spinoffs. This science-fiction vision still grips Google today—and Singhal sees the acquisition of Metaweb, taking Google beyond a purely statistical understanding of the world, as a critical step toward achieving it.
“We all want the *Star Trek* computer, and what we are building is exactly in that direction, where you can ask it anything and it will proactively tell you things,” says Singhal. “That is where this is headed. The Knowledge Graph is one of those key components that are necessary but not sufficient to build it. Likewise with speech recognition. When you put the whole package together, that is how you arrive at the future—and we at Google are far ahead in all of those spaces.”