TYPING SOUND] [MUSIC PLAYING] NARRATOR: This is Louis XIV, also known as Louis the Great,] Louis the Grand Monarch and Louis the Sun King.BEN GOMES: But I say "Gohms." It's a Portuguese name. NARRATOR: This is Ben Gomes. He knows a few things about search. Uh, that search. Anyway, he's kind of a big deal, even though he'd try to convince you otherwise. Ben worked on Search for more than 20 years. But that's now where his story started. BEN GOMES: So I was born in Dar Es Salaam in Tanzania. But at a very early age, my parents moved back to India to Bengaluru. And there was a few books at home from my elder siblings. And that's the information I had access to, including I remember one torn encyclopedia that I think my grandfather had given my mom. So it was really out of date. In 5th grade, I got two presents-- a bike, which my parents thought I'd be very excited about, and a much better encyclopedia. And I was actually much more excited about the encyclopedia-- this is where geeks come from-- than the bicycle. And my parents didn't know what to do with this. [MUSIC PLAYING] When I look back at how we found information, it was so dramatically different from today. When my mother was growing up, where there was not even access to a good library, you would have just accepted the fact that you didn't have the information, and that's the way it was going to be. When I was growing up, for some kinds of information, there was a decent library. But you still had to take this bus. It took about an hour. You had to look things up in a card catalog. That took time. Now today, we measure in fractions of a second the time it takes for you to get information. I think that reduction in friction is absolutely dramatic, because it can enable people around the world to have equal access to information. It's not just that people in some places who have access to the best libraries. Everybody should have access to the highest quality information. So that combination of a deep technical problem and I think a fundamental human need to understand the world around us, to know more about the world around us, is the heart of Google Search, and what keeps me coming to work still so excited 20 years later. So in the early days, I wondered whether the company had the infrastructure to be a real company. Because when I had come for my interview actually, there was not even a sign indicating that this was Google. So I was not sure I'd come to the right place. But halfway up the staircase, there was a small neon sign that said Google. So that's when I knew. [MUSIC PLAYING] And it generally felt completely chaotic. And Jeff was there. Jeff is also brilliant. JEFF DEAN: Yeah, we were a very small company. We were maybe about 25 people. We were all kind of wedged into this second floor area in downtown Palo Alto. I was in an office with Urs Holzle. BEN GOMES: Urs was in charge of all of engineering. And at the time, I don't think I knew how to pronounce his name. But he put the three of us named Ben in one office, just so people would walk by and say, hey, Ben. URS HOLZLE: Yes, we had the Ben Pen. I think it was pure coincidence, actually. My first reaction to Google was, I have no idea what Search is, so it's probably not for me. But then I was intrigued by the problem. It was clear that there was some real value there. Because without really good ranking, all that growth of the web would be wasted if nobody could actually find the things that were there. BEN GOMES: So one of the core aspects of Search is, how do we rank results and how do we find the most relevant information. So a lot of people work on that. You'll get really good stuff on this from Pandu, actually. CREW: Pandu? OK. [MUSIC PLAYING] NARRATOR: This is Pandu Nayak-- PANDU NAYAK: Hi, I'm Pandu. NARRATOR: --head of Search Ranking. His personal motto-- PANDU NAYAK: No query left behind. NARRATOR: Before working at Google, Pandu

worked at an artificial intelligence lab at NASA. PANDU NAYAK: Yeah, we built an autonomous system that provided high-level control to a spacecraft called Deep Space 1, really the most exciting thing that has ever happened in my life-- in my professional life, I guess. NARRATOR: After doing that, he wanted a new challenge. PANDU NAYAK: I oversee the Ranking team. So ranking is important because if we simply return the million pages that match your search query, that's not particularly helpful. And so we need to rank the pages that you might find useful. Hopefully, these are at the top of the results. We're really trying to bring information to the world at large and make it useful so people can improve their day-to-day lives. And I feel really lucky to have the opportunity to work on this mission. [MUSIC PLAYING] NARRATOR: Let's go back a bit. Summer, 1999, room 300 and something in the Gates building at Stanford. And these two guys, Larry and Sergey, who were about to announce something so big it merited matching polo shirts. LARRY PAGE: OK, maybe we should get started. So what is our mission? So how is Google different? Basically, we want to organize the world's information and make it universally accessible and useful. NARRATOR: 20 years later, bigger stage, same deal. [CHEERING] SPEAKER 2: And today, our mission feels as relevant as ever. [MUSIC PLAYING] NARRATOR: So what does this actually mean? Here are a few takes. CATHY EDWARDS: I think if we weigh up the various parts of the mission, to me the most important piece is organizing. There are hundreds of billions of web pages that are out there. Our job is to filter through that and to really give you what you are looking for at that moment in time. NICK FOX: And then the next part is the world's information. So information means really anything. It started out for Google with web pages, but it's so much more than that. DAVID BESBRIS: Whether it's physical books that we need to scan or maps that we build of every place on Earth, that's information, too. And it's not web pages. It's the kind of stuff that we organize today. TULSEE DOSHI: And then I think that word universal is important, because universal means for everyone. NICK FOX: Whether it's someone that can't see, whether it's someone that can't hear, people that speak different languages, really make it accessible to as broad a set of people as possible. DAVID BESBRIS: We might be goofy people who come to work in T-shirts and desperately need haircuts and things like that. We may not look super serious, but we know how much people rely on this. We take that mission and really, really seriously. SPEAKER 3: 1.0. NARRATOR: So it sounds like the mission is pretty important to these folks. But here's another important question. CREW: So how would you explain how Search works? BEN GOMES: Right. Yeah, so how does Search work? TULSEE DOSHI: How Search works? [MUSIC PLAYING] PANDU NAYAK: How Search works, in a nutshell. NARRATOR: This is server rack 3349b. It lives here in Ballybane, Ireland, along with cows, a golf course, and Kavanagh's Auto Accident Repair Center. This is one of the places where Search happens. Search is a big piece of software that takes the words you type in here and looks for them here, on the worldwide web. It can do that because first it downloads a copy of the entire web, scans it, and makes a list of all the words and lists of all the pages each word appears on. It's like the index of a book, except 10 trillion times longer. Lasagna appears on 59 million of those pages. When you search for lasagna, the software puts these pages in order with what it hopes are the most useful at the top and less useful at the bottom. Most people searching for lasagna want a recipe for lasagna. COOK: Look at how delicious that looks. NARRATOR: Some people want nutrition facts for lasagna. And a few people want to learn about the life

and research of Louis C. Lasagna, MD. They call him the father of modern pharmacology. The software living on server rack 3349b helps rank those pages, depending on where you live, whether the page was updated recently, [OVERLAPPING] how many other pages link to that page, how many times the word lasagna appears on the page, is lasagna in the title, is lasagna bolded, are there pictures of the lasagna? It does all this in less than one second, billions of times a day every day, mostly for things that are tougher to figure out than lasagna. BEN GOMES: So behind the scenes of Google Search, there are many kinds of engineers and many different teams that come together to bring you to such experience you see, teams around the world, in many other countries-- Zurich, London, India, Japan, so on. You have teams that are working on the interface by which we present this information, teams working on the evaluation processes, processes that sure that the changes that are happening are good changes. And then there are teams of engineers who work on ranking. They might examine the kinds of queries where we are not doing well today, and think about, what are the kinds of techniques we could use to enable us to do better in the future? NARRATOR: Like the team that's about to enter this meeting. ELIZABETH TUCKER: Anything we need to know? CREW: Don't look at the lens of the camera. ELIZABETH TUCKER: OK. All right, let's do it. NARRATOR: Despite their lack of on-camera experience, they're working on what could be the biggest change to Search in over a decade. SUNDEEP TIRUMALAREDDY: Things are getting exposed. SPEAKER 4: Part of that is building-- NARRATOR: But we'll get back to them later. SPEAKER 4: But we will actually see some-- [DOOR SLAMS] BEN GOMES: So Search is a pretty complex product. It's a big effort to actually make these things work, to take all of these different pieces of the system, using a lot of mathematics, and then trying to bring them together into something more real, into something that can actually be turned into an algorithm. NARRATOR: All right, so behind the scenes, people at Google are working on algorithms. [MUSIC PLAYING] Let's dig into that for a minute. At its most basic, an algorithm is just a set of mathematical instructions that a computer follows, kind of like a recipe. Just like there are different recipes for different dishes, there are different algorithms for different jobs. Some make elevators go up and down. Some predict subway delays. Some help cars parked themselves. The Google Search algorithms exist to return high-quality information based on a user's query, stuff like all of the text, pictures, videos, and ideas that people have taken the time to put on the open web, stuff they want other people to find and read and watch and look at and learn from. PRESENTER: Hey, guys! LUCAS: Lucas here. ROOFER: In today's video, I want to show you-- TEACHER: --how to simplify a rational expression. NARRATOR: We're talking about the angle of the Leaning Tower of Pisa, how to hit a 7-10 split, whatever this thing is. This is the information that Google tries to organize and make universally accessible and useful, because this is the kind of information that people are out there looking for. But you know what they're not looking for? VOICE: Act now! VOICE: We will be with you shortly. VOICE: A whiter, brighter smile! VOICE: Hey! NARRATOR: Spam. Not the delicious kind, the bad kind. CATHY EDWARDS: Yeah, so let me just talk about spam for a minute, because spam is one of the biggest problems that we face. NARRATOR: This is Cathy Edwards, head of User Trust for Search, which basically means she deals with a lot of crap so the rest of us never have to. CATHY EDWARDS: Broadly, spam is what we consider a low-quality page that is artificially boosted in our results. NARRATOR: She's talking about

pages that use AI-generated nonsense text, hidden keywords, and hijacked URLs to trick their way into people's Search results, pages like fastcashonline.org, topicalarticles.info, the kind of websites that, when you end up on them, you hit the Back button as quickly as possible, because they're [BLEEP]. Because they're spam. DAVID BESBRIS: There's a wide variety of motivations why people do this. Sometimes it's commercial interests. CATHY EDWARDS: Spam, where they're trying to sell things that are a little bit dubious, right? Or sometimes it can just be to capture more of the user's clicks. And that's not right. That site is not getting those links organically. It dilutes the value of that signal. It makes it even harder for us and it makes it harder for users to find great information. DAVID BESBRIS: It's a very, very hard problem, because people on the other side are very motivated to succeed. And they're smart, too. And they have resources, and they're working on it also. We solve one part of it, and they adapt and they do something else. CATHY EDWARDS: And that's the reason that we keep Google's Search algorithm a very closely guarded secret, recipe-for-Coke-level guarded secret. DAVID BESBRIS: Because if we talk about our [? Search ?] [? signals ?] too much, then people will manipulate them. And that breaks Search entirely. Fighting spam is a cat and mouse game. It's not something that I think will ever be solvable. CATHY EDWARDS: As an example, 40% of pages that we crawled in the last year in Europe were spam pages. This is a war that we're fighting, basically. NARRATOR: So yeah, people at Google hate spam, which is one of the reasons they're always making changes to Search, to keep spam out of your results and to keep high-quality information in. [MUSIC PLAYING] BEN GOMES: OK, so you've got the Search engine and it's working. And by all accounts, it's working better than any other search engine has worked before. And every day, you see millions of queries. And clearly users are happy. But as an engineer, you ask yourself, how can I make this better? You see many ways in which we are still failing. And you see a ton of opportunity for us to make it even better. And over a period of time, the developments we've made in the Search Engine have had a dramatic impact on how well it actually works for users. PANDU NAYAK: No, no, I don't think we had that particular problem. Even though we've launched a whole series of changes over the years that have, I think, meaningfully and materially improved the Search result sets, I'm here to tell you that Search is far from a solved problem. In fact-- BEN GOMES: There's actually no end in sight, in terms of when this will actually be solved. Because the world keeps evolving. We're coming up with new devices. We're coming up with new ways of interacting with information. We're coming up with new information sources, like videos and so on, that are adding in new opportunities, as well as new challenges. CATHY EDWARDS: The content on the web has changed. Users have changed what they're searching for and how they search. For example, 15% of the queries-- PANDU NAYAK: 15% of the queries-- BEN GOMES: 15% of queries we see every day-- CATHY EDWARDS: --we have never seen before. That's just going to keep happening, and we're going to need to constantly evolve to keep up. It's a little bit like the Red Queen says to Alice in "Alice in Wonderland," you need to run as fast as you can to stay where you are.And so what was happening on the "did the Holocaust happen" type of queries is that the relevant signals were overpowering the quality signals to a degree that was resulting in low-quality results for users. PANDU NAYAK: We have long recognized that there's a certain class of queries, like medical queries, like finance queries, in all of these cases, authoritative sources are incredibly important. And so we

emphasize expertise over relevance in those cases. So we try to get you results from authoritative sources in a more significant way. MEG AYCINENA LIPPOW: And by authoritative, we mean that it comes from trustworthy sources, that the sources themselves are reputable, that they are upfront about who they are, where the information has come from, that they themselves are citing sources. PANDU NAYAK: And so the change we have made in the case of misinformation is to change the ranking function to emphasize authority a lot more, and this has made all the difference. SPEAKER: Actually, not these. NARRATOR: Misinformation is one of the challenges that comes with helping people find what they're looking for. But it's not the only one. Launched in 2010, the Autocomplete feature has saved millions of hours in people's time by guessing what they're searching for before they finish typing. But when those guesses have been wrong, it's led to some pretty disturbing predictions. REESE PECOT: A few years back, we started hearing from people that sometimes folks were typing things into Autocomplete and they would be shocked by some of the predictions that they were getting. Autocomplete was designed to help people complete their searches faster. Instead, we were actually returning them information that they weren't searching for. When we provide you with something that's shocking, that's not relevant, we've really at that point not stood up to our core principles. PANDU NAYAK: I think I and all the members of the team felt a deep personal responsibility to try and develop the systems to minimize these kinds of occurrences as much as possible. First, we developed a set of policies that say what kind of predictions that we would not want to offer to users. REESE PECOT: Things like violent content, sexually explicit content, hate speech. But we also publish those policies. That way people can see where we stand, and then that gives us some accountability. PANDU NAYAK: With these Autocomplete algorithms, we try not to surface predictions that violate the policies. Now, these algorithms are very good at what they do, but they're not perfect. And every so often, we'll get some predictions that in fact violate them. REESE PECOT: So you can report if you've seen a prediction that violates those policies. And every day we get flags from our users out there to tell us where we might be seeing problems in the product. PANDU NAYAK: We use those reports to improve our algorithms to try and see whether we can address the whole class of problems that the report might be just pointing towards. But one thing that I would like to emphasize is that this in no way prevents users from searching for whatever it is that they want. They're absolutely free to do that. NARRATOR: Think about it this way. Search is like a door that leads to the web. With Autocomplete, it's the kind of door that senses you walking towards it and opens for you. But if you're typing a query that violates its policies, the automatic part stops. The content of the web is still behind the door, but you won't see any results until you complete the query yourself. NICK FOX: Search isn't perfect. We do make mistakes. We make more mistakes than we would like to make. But we need to learn from them. We need to get better. And we need to continue to improve to avoid those cases in the future. Each time that something happens where we become aware of a bad result, we use that as learning. We use all that feedback to continue to improve it and make sure that Google one day from now, five days from now, 10 days from now, 10 years from now, is continuing to get better. BEN GOMES: Many people tend to think that Search is really easy. You type in a few words, you get a few documents, and the process feels very easy. And in many ways, that's what we want to achieve. We want Search to be very easy for people. But behind that is an

extremely hard technical problem of actually understanding what people mean when they type in a query, not just matching words, but actually understanding language much better over time so that we can match the thing you asked to the concept that you were really looking for in the documents, and we can bring these two things together. It's an absolutely fascinating problem to work on, because it lies at the frontiers of what computers and computer science can do and our understanding of basic aspects of how we wish to interact with computers as human beings. NARRATOR: As long as there have been machines, humans have tried to get those machines to do more. Of course, for most of history, the machines couldn't speak human. So humans had to come up with new ways to tell machines what to do. Joseph Jacquard used cards with holes punched in them to tell his loom, put the thread here and here and here. It made weaving complex patterns easier. Punch cards were a big idea. They're how early computers took instruction, did math, solved equations. NARRATOR 2: Holes punched in the card represent data to be placed in the computer. NARRATOR: Then computers got screens and keyboards. But you still couldn't talk to it like you'd talk to a human. You had to write it in code. C colon, slash carat smartdrv dot exe. Once Search came along, things got a little easier. You just put in the words you were looking for and Google came back with websites. But you were still writing in code-- "ice cream shop 27705," when really you meant, "where can I get some ice cream around here?"And Google only gets paid if the person doing the search, maybe your neighbor or your brother-in-law, clicks on Bart's ad, which is always labeled "Ad." It helps people find mowers to buy, and it helps Bart and the store get business. BART: Have a nice day. NARRATOR: And it helps pay for all the stuff that keeps Search and Maps and Docs working and free. That's why there are ads. PANDU NAYAK: Since I've been at Google and worked on Search for the last 14 years, I have to say that no one, absolutely no one, comes to me and says, you know, I did this search and the results were great. Nobody says this. They only call to complain that they did something and it didn't work. NARRATOR: And the name of the man who's been collecting Google's dumbest Search mistakes for the last 14 years? [CHEERING] Senior Software Engineer Eric Lehman. CREW: Eric L, take 1, mark it. ERIC LEHMAN: Over the years, I've been gathering some of my favorite bloopers. I'll walk you through some of those. So how far from the coast is Cambridge, Massachusetts? It's actually a little over 3,000 miles from the West Coast. How many calories in 330 tons of butter? So this caused an overflow error, and we said about minus 2 billion. Mm-hmm. What color is green? That's a tough one. Blue? Sure. For the search "meat nutrition facts," we brought up all kinds of detailed information.When DeepRank was tested on this query, it understood that the result was about cookies, reducing the prominence of the incorrect recipe, and instead elevating useful, relevant information about cooking fish.SPEAKER: The distilled model's pretty quick [INAUDIBLE] SUNDEEP TIRUMALAREDDY: Yeah, I think 10 milliseconds or so. BEN GOMES: It .seems like a reasonable trade-off for this level of win.NARRATOR: This is Ben Gomes