George W. Bush isn’t the only “W” in the world. People who worked with Unix in its early days know Peter Weinberger as the “W” in the AWK programming language. Named for its creators, Alfred Aho, Peter Weinberger, and Brian Kernighan, AWK arrived in 1977 and became part of Version 7 Unix, a key release by Bell Laboratories. AWK gained notice for helping people manipulate text-based data quickly. Since then, AWK has morphed significantly and has been blended into many flavors of Unix. GNU awk, also known as “gawk,” is one of several free versions still in use.

Today, Weinberger works for the king of search companies, Google, where he helps create tools to better manage the mounds of information created by employees.

Weinberger spoke with S&P about Google’s technology and security challenges, the future of search, and related privacy concerns.

By the way, if you Google “Peter Weinberger,” you will find another legacy of his from Bell Labs. A colleague created an image that morphed Weinberger’s photo with an AT&T logo, and it became a running joke inside the company. The image remains tucked away in the darndest places, both inside software and offices. (See http://spinroot.com/pico/pjw.html for a peek at the image and its history.)

S&P: Tell us about your current role at Google.

Weinberger: I’m a software engineer. I write programs, infrastructure stuff mostly. I’m working on tools for processing various kinds of logs created by internal software. I’m not working on security directly, but a lot of what we do is affected by security and privacy concerns. As a result of the Sarbanes-Oxley Act, most companies now create a lot of logs to verify what code they’re running, and [show] how people work with internal data.

Also, I spend a lot of time reading code. At Google, you can’t check in code without it being reviewed and approved by a peer, so if you work with productive people, there’s a lot of code to review.

S&P: What are the biggest technology challenges for Google today?

Weinberger: Scale is the problem. Our business grows rapidly. That means every year, a lot of the technology decisions made a year ago don’t look so good any more. Exponential growth is a very pleasant problem but requires a lot of work.

S&P: Is Google running on Linux today, and is that helping you manage the scale challenge?

Weinberger: The servers are Linux. My desktop is Linux, but we, like most companies, have a lot of Windows desktops. And some engineers use Mac laptops. I think Linux is helping as well as anything could with the scale. It does make the boxes less expensive.

S&P: Localized search has been a big effort in the industry for a while now. Why is this so hard to pull off, technology-wise?

Weinberger: My uninformed opinion is that it’s mostly a data problem. For example, walk down the street and you see that local stores turn over all the time. Who’s going to keep track of all that information? I don’t know a way of keeping up with it. It’s intrinsically complicated today.

In the old days, when the phone company was a monopoly and everyone had a phone, the phone company knew where everyone was, but only told you once a year. That’s not so useful today.

S&P: What are the main security and privacy challenges that worry you?

Weinberger: It’s interesting, many years ago when I started at Bell Labs, how things were different. We could break the Unix file encryption and read the encrypted password file. Encryption is perfectly satisfactory these days. The amount of computing power it takes to break even DES [Data Encryption...}
About Peter Weinberger

Peter Weinberger received a PhD in mathematics (number theory) from the University of California, Berkeley. After teaching mathematics at the University of Michigan, Ann Arbor, he moved to Bell Laboratories, where he worked on Unix and did research on topics including operating systems, compilers, network file systems, and security. He then moved into research management, ending up as information sciences research vice president, responsible for computer science research, math and statistics, and speech.

After Lucent and AT&T split apart, he moved to Renaissance Technologies, a technical trading hedge fund; as head of technology, he was responsible for computing and security. In 2003, he moved to Google in New York, where he now works on software design and implementation. (In the Google style, he admits to no title.) Weinberger also serves on various groups giving technical advice to the US government.

Standard, which is quite old by now, is very large. As for attacking computers, the early Unix attacks were escalation-of-privilege attacks. But today you don’t need to escalate privilege to get control of a user’s machine. You have phishing and spyware attacks that are much harder to fight. That’s a substantial and annoying change. There are fewer technical ways of defending against these attacks. And I think everyone is taking identity theft seriously. A lot of the mechanisms [we rely] on for authentication and digital commerce need the endpoints to be secure, but if your laptop is compromised, then potentially it’s all for naught.

S&P: Are there privacy challenges related to the work you’re doing with the internal software log data?

Weinberger: Yes, in principle, but the execution logs I’m looking at are associated with access to protected data. You’re not trying to notice what individuals are doing; you’re trying to understand why the response time for their getting results is mediocre. It’s a different situation than in a financial company that is worried about leaks of inside information.

S&P: Do you have any security concerns regarding Google Maps? As these maps become more common, what are the security issues?

Weinberger: Me personally? Not really. One of my friends was looking at Google Earth, and [that] observed they could date the picture by where their car was parked, but it was just a tiny blob. Until the resolution gets hugely better, there’s not much there that I worry about.

S&P: How do you personally envision Web search being different a few years from now?

Weinberger: Google maps has really made a difference—combining street addresses with satellite photos, that’s not something I would have thought of before it came out. But a lot of work is being done with this, not just by Google, and it’s clear that in the very near future it’s going to be easy to put nicely labeled satellite photos on your Web site.

What I’d like is more of what I call real information. Web search today finds Web pages that match a description you’ve given, but the quality of the information depends on which Web pages are authoritative and which are not. For example, if a lot of Web pages decided my birthday was April 1, pretty soon when you Googled on my birthday, that would be the answer, even if it’s wrong. Today it can be hard to [recognize] authoritative Web pages. To some extent, we depend on good will.

Also, public database information is very hard to get from regular search. There’s a lot of information in those databases today. For example, if you wanted to know who the oldest person on your street was, that’s not an answer you can get from search engines today. And it would be data from the last census anyway. That’s a lightweight example, but there’s a whole class of these things.

S&P: What else do you want from search?

Weinberger: The other thing I’m looking forward to would be good search from my cell phone.

And one question that keeps coming to mind when I search and come across a foreign language site is how a search engine could understand it better. Let’s say I search on “Athens population.” For several reasons, I’m unlikely to get a translation of an official Greek government site, which might have the most authoritative information. It’s a hard problem to solve. It’s not clear that it’s impossible, it’s just not clear how good a job you could do. It’s not going to happen within the next few years. The problem is really serious the other way, for people who don’t read one of the Web’s major languages.

S&P: Earlier in your career, what inspired you to work on the AWK programming language?

Weinberger: At the time, I was in a group that worked on database projects, and the research guys were looking to build a Unix utility that did some database work. We decided to do some kind of report generator. We produced AWK to break each line of code into fields. A design goal was that we wouldn’t have to explain the syntax to anyone.

At roughly the same time we...
were doing AWK, there was a program at Xerox PARC [Palo Alto Research Center] that had quite similar goals. But they wanted not fellow programmers, but people like secretaries to be able to generate better reports. It seemed very usable, and nicely designed. What’s striking is [that] AWK survived, and their project (whose name I have forgotten) didn’t. It’s not clear that the best technology will always win. In our case, Unix was being given away for free to universities, and AWK went along for the ride.

**S&P:** During your time at Bell Labs, did you work on any security tools?

**Weinberger:** We worked on security-breaking tools. There was sort of a cottage industry trying to break file encryption. Security’s hard. It’s much easier to look for flaws.

**S&P:** Given your experience working for Renaissance Technologies, a technical trading hedge fund where you were responsible for computing and security, how useful did you find commercial products in the security space?

**Weinberger:** Firewalls come right to mind. A lot of work went into firewalls, and they are strong today, but with antivirus software and spam filtering, you have to do a lot of the work yourself. As long as you’re running Windows, your machine is intrinsically insecure. These horrid email-borne viruses show up in all kinds of unexpected places. There’s a real, three-way conflict between running secure systems, getting your organization’s work done, and letting people use their PCs.

**S&P:** What are your technology pet peeves?

**Weinberger:** I’m sure I’ve got dozens. Password management. Sometimes people, like auditors, think users should regularly change passwords. I think this just encourages people to use crummy passwords—because either they have to write the passwords down, or they need a sequence of words that are very similar.

One of my other pet peeves here is curious error messages. A lot of them are still awful. They start with the word “abort” but the program continues. We have a lot of legacy software, and you end up with these situations where an error message pops up, but the computer keeps on going. Deep in the software something went wrong, and reports it, but you’re seven levels of abstraction away, and it’s just useless noise.

**S&P:** You’ve witnessed several generations of computer technology and business models come and go. What are the most important ways that security technology has improved during that time?

**Weinberger:** There are much better underlying mechanisms for security. When you go to a Web site and type in something confidential, all that stuff works fairly well, assuming people on the other end treat the data properly. VPNs [virtual private networks], SSL [Secure Sockets Layer], all those certificates your browser checks automatically … these are big advantages.

**S&P:** Some of your Bell Labs colleagues merged a picture of your face with the AT&T logo, and the resulting image showed up all over the building. What’s the most surprising place you ever spied the image?

**Weinberger:** It all started because Tom Duff decided the part of my hair matched the white spot on the AT&T logo. The first time I saw the picture in public was pretty striking. It was done around a time when I had just been promoted and was a little nervous. I was having my first meeting with someone outside of research, and I looked out the window and there it was on one of the water towers. Another time, we had a meeting in the auditorium in Bell Labs. At some crucial point, a large number of people held up paper masks with the picture on them. What can you do at a time like this? Take it well and enjoy it. Most of the time I found it funny. I’m pretty sure that a picture made with little magnets is still in the building.

**Laurianne McLaughlin** is a freelance technology writer based in Massachusetts.

### With antivirus software and spam filtering, you have to do a lot of the work yourself.

The aim of the journal is to provide a forum for research and applications for multimedia and information security topics including, but not limited to:

- Multimodal interfaces
- Coding and transcoding of Multimedia signals
- Multimodal Biometry
- Digital watermarking, new methods and benchmarking
- Ciphering of Multimedia flows
- Mobile Multimedia