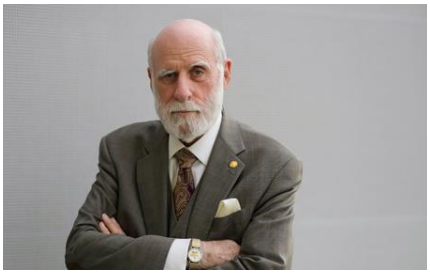
Freshfields interviews leading figures who have shaped the world of technology we know today and are leading the tech-innovation of tomorrow.

Episode 1: Turing with Vint and Whit: Boris Feldman in conversation with Vinton G. Cerf and Whitfield Diffie

Boris Feldman, Freshfields Partner, interviews two of the most prominent pioneers of technology – Vinton G. Cerf, one of the 'fathers of the internet' and Whitfield Diffie, a leading inventor of public-key cryptography. Listen in for a fascinating discussion on the origin of technologies that have changed the course of history, and predictions for how technology will continue to change our futures.

Boris Feldman
Freshfields Partner, Silicon Valley



Vinton G. Cerf
American Internet pioneer, Vice President and Chief Internet Evangelist for Google

Whitfield Diffie
American cryptographer and mathematician, pioneer of public-key cryptography

Boris Feldman

Welcome to our video podcast today with two of the greatest figures in the history of the Internet and technology. My name is Boris Feldman. I'm a lawyer at Freshfields Bruckhaus Deringer in Silicon Valley. And rather than taking time telling you about how great our guests are today, I'd like you to go online and look them up. One of them is Vint Cerf, who is one of the pioneers of the Internet. The other is Whitfield Diffie, who was one of the pioneers in public key cryptography. And if you ever wonder when you're browsing and you see https, the "s" indicates that you're in a secure website and you owe that security to Whit.

So we're going to talk to them today. Thank you very much for joining us. And we're going to take a look both backward and forward. We're going to talk about the things that have surprised you about developments and what you've created and perhaps disappointed. And then we're going to look to the future, both in terms of security on the Internet and equally important, AI and what your hopes and fears are for it.

And then we're going to close by asking you to provide some free career advice to younger people starting out today. So let me start with Vint Cerf. Can you tell us about the founding of the Internet and your role in that?

Vinton G. Cerf

Well, the short story is that the Defense Department got very interested in computer communication because it was supporting research in artificial intelligence in the 1960s. They were supporting about a dozen universities and they wanted them to share all of their results with each other. And everybody kept asking for a new computer every year. DOD couldn't afford to do that for a dozen universities.

So they said, "We're going to build a network and everybody has to share." So they designed and built something called the ARPANET, the Advanced Research Projects Agency Network, and it actually worked out very well. The universities were all connected. They shared their software and their computing. So the Defense Department says, "You know, this looks pretty good. Why don't we figure out if we can use computers in command and control?"

Well, if you're going to do that, then you have to think about computers being in mobile vehicles, ships at sea and aircraft in addition to dedicated facilities in air conditioned rooms that were part of the ARPANET system. Robert Kahn, who worked on the ARPANET, went to ARPA and started this program, which eventually was called Interneting. And he and I, while I was at Stanford University, sat down to figure out how do we connect the mobile packet radio network, the packet satellite network, the existing ARPANET into a common multinetwork system so that we could support computer communication in mobile vehicles, ships at sea, and aircraft.

That was the Internet problem.

And in about six months, in 1973, we figured out how to do that.

We wrote a paper, which was later published in May of 1974, in *IEEE Transactions in Communications*. And then I continued to work on the program with my graduate students and then eventually went to ARPA to join Bob Khan, who ran the program for six years.

So my job was basically to help give birth to this thing, which went operational on January 1, 1983.

Boris Feldman

So we had the 40th anniversary of this year.

Vinton G. Cerf

That's correct.

Boris Feldman

Years ago, I had the pleasure of representing you on the board of a company, and I told my kids that I was representing the inventor of the Internet, Vint Cerf. And they asked me, "Is that why we call it surfing the Internet?"

Were they right?

Vinton G. Cerf

Well, not exactly. The term was given prominence by Jean Armour Polly, who wrote a paper about surfing this information sea, this information ocean, but she spelled it s-u-r-f-i-n-g. It turns out that in San Diego, one of the commercial network providers was developed in 1989 by General Atomics, and they were going to call themselves s-u-r-f-net because they were going to be hooking universities in California together.

And what else would you do when you're in San Diego anyway but surfing the net? Then they discovered that there was a surf.net that was already in the Netherlands connecting Dutch universities together. So they couldn't use surfnet. Somebody decided they could change the name to the California Educational Research Foundation, spelled

CERF net, and it sounded exactly the same.

Then somebody said, well, maybe we should call Vint. So they called me up and they said, "Is it okay if we call this thing C-E-R-F net?" And my first reaction was, if they screw it up, am I going to be embarrassed? Well, I thought some more about that, and you know, people name their kids after other people.

And if the kids don't come out right, they don't blame the people they name them after. So I said, sure, go ahead. So in July of 1989, I flew out to San Diego, and Susan Estrada, who was the executive director, and I took one of these plastic bottles of glitter and we smashed it over a Cisco router and we launched CERFnet.

Boris Feldman

Perfect!

Whitfield Diffie

Wait a minute. I'm one of these awful interviewees who argues with the previous speaker. The only thing I want to argue with is that it can't have been a new idea of the seventies to use computers for command and control. It linked SAC at least to SAGE, which developed a lot of the technology we later used.

Vinton G. Cerf

Fair enough. Although SAGE was primarily a distant early warning system as opposed to…

Whitfield Diffie

Well, the term command and control came up in that era out of the Strategic Air Command control system.

Vinton G. Cerf

Well, actually, you know, if we are going to have this conversation … there was another system that was specifically for command and control, and that was a message switching system that the Defense Department built, which is independent of SAGE. SAGE was for distant early warning radar information.

And I'm not sure I remember the acronym for this, but they had a message switching capability before the ARPANET was built. So you're correct about that. But it was one of those uniform systems, and in the case of ARPANET and later the Internet, there was a lot of diversity among the computers.

Whitfield Diffie

I think the point of your role in this is the "inter" part of it and the diversity of systems. And I think SAGE unquestionably was called command and control. Distance had very little to do with SAGE. It's about radar being able directly to tell computers how to tell aircraft.

Vinton G. Cerf

No, no.

Whitfield Diffie

Oh, yes, yes.

Vinton G. Cerf

Well, you and I will have to have that debate outside. I don't believe that there was a control there. I think this was human beings tracking the radar information and then reporting if there was something coming over the Pole. They wanted to make sure you distinguish Russian bombers.

Whitfield Diffie

That's what the DEW [Distant Early Warning] Line did.

Vinton G. Cerf

Yes, but the DEW Line data went to SAGE.

Boris Feldman

I'm going to violate every rule of moderating. I'm going to jump to the end because this is the most important thing we're going to talk about today. The role of a AI in control and command. And to what extent do you worry that when these fancy new artificial intelligence systems play an even greater role in the military than they do now, how much are you worried about civilizational risk?

Whitfield Diffie

Civilization … humans haven't got a chance. Humans are not going to be running the world at the end of the century. And it won't involve any sort of war. It's a very simple mechanism. People like to have things done for them. The more AI is offered to do things, people say, sure. At some point we'll look around and it'll be like, not like in the government.

There'll be nothing you can do about it. AI will be running the world, and it's going to be more like 2050 than 2100.

Boris Feldman

Do you share his optimism?

Vinton G. Cerf

Well. Right. So, first of all, there's an important reading assignment for the viewers. It is a book written by E.M. Forster. It's a novella written in 1909: *The Machine Stops*. And so the world that you were describing was imagined already in 1909 by E.M. Forster. The community lived in their homes. Their food is delivered to them somehow.

He doesn't explain the details. They communicate with each other through the machine. And so this society is kind of like during the pandemic.

Whitfield Diffie

I read the story during the pandemic and I thought, "Ooo, that world has come to life."

Vinton G. Cerf

Exactly right. So, of course, he pauses in the early pages of the book when the machine stops.

And so the question is, what happens to civilization? So, getting on with the train of thought, if we become dependent on the machine in the way that Forster's book has us going, then the question will be what happens when it stops working? I am somewhat worried about our overdependence on technology, partly for the reasons Whit brings up, which is our willingness to be taken care of by automatic methods.

I worry, for example, that nobody will know how to read a map anymore because they're expecting GPS to work on their mobile devices. I even worry that people are extremely dependent now on mobiles and if the mobile doesn't work, there are cascade of failures all over the place. Already, and it's only 2023. So I am a little worried about not just AI but about software in general becoming both brittle and also unpredictable and heavily depended on by our society.

Whitfield Diffie

And when you look back from 2200, this will be the birth story of machine civilization. All the problems at hand.

Boris Feldman

So if you were appointed the AI czar, at least in the US…

Whitfield Diffie

I'd complain that I didn't have enough power.

Vinton G. Cerf

Is that electrical or otherwise?

Whitfield Diffie

I was thinking political.

Boris Feldman

You just got a brevet promotion to Global AI Czar.

What would you do now at this turning point?

Whitfield Diffie

Obviously invest more money in AI.

Boris Feldman

Even though you worry about it taking over everywhere.

Whitfield Diffie

I don't worry about it. I simply think it's inevitable. Well, so let's actually … Let me stir the pot a little bit more.

When I say AIs, I mean things with immensely complex behavior. I think the objective we ought to have is to develop something that is a billion times as intelligent as a human being. But whether it'll be done with silicon or whether it'll be done by engineered biologicals or whether it be done by something else that I don't foresee, I don't know. But it seems to me that we are in the process of giving birth to our children who will take over the world.

Vinton G. Cerf

So this is interesting. Well, of course, children will inevitably take over the world since we all die anyway.

Whitfield Diffie

So I do expect to die, but I don't think everybody alive today is going to die in any reasonable amount of time. I bet there are people alive today who will be alive in a thousand years.

Vinton G. Cerf

Wow!

Boris Feldman

We're going to come back to this. But on the issue that Whit raised about controlling it, do you see any way to control these developments in a positive way?

Vinton G. Cerf

Well, some of these things are being used in a constructive way. The problem is that we don't know when they are misbehaving or we can't predict that they're going to misbehave. So you get sometimes what sounds like very good advice and it turns out it isn't. These are the large language models. I think the large language models are just the beginning of a much more elaborate kind of intelligence.

At the moment, what we have in the large language models, as I see it, is essentially the verisimilitude of human discourse. And I use that word on purpose because it sounds and looks like human discourse, but it's being generated by these language models. And I don't want to trivialize this because it's reasonable to point out that even though they generate text or now imagery, video and sound, and they carry on discourse sounding very human, there is something fairly deep going on, even in these relatively crude, probabilistic models.

And I don't fully appreciate what's going on, but I'll give you an example. At one point, someone described asking a chatbot to simply reverse a string of random characters. And so it did that and then, unasked, it said, "And here's the Python program to do that." Now, that surprised everybody because they weren't anticipating that it would voluntarily (I use that word carefully) offer that particular output.

Something deep is going on there. But I don't think we fully understand it and I don't think we know how to predict when it's going to go awry. So we have work to do.

Whitfield Diffie

Okay. So let me suggest something. There's no chance of our doing what would perhaps, you know, curb the activities of AIs. We could have something like Asimov's Rules of Robotics, and we have nothing like that.

We have no rule that says a machine must obey a person. And we have this notion that a machine can obey one person in order to abuse another one, or control another one.

Vinton G. Cerf

Of course, there are stories that Asimov has written describing his ways of getting around the three laws.

Whitfield Diffie

His writing malfunctions on a set of basic set of notions that make fine stories.

But the point is, there is dating back to the invention of the pin tumble lock in ancient Egypt, no objection to a machine having control over a person. There's no rule that says when you say something to a machine, it must do that if it possibly can. If you look around you, lots of the technology that has sprung up in our lifetimes is about controlling people. And the smarter the machines get, the more good they will get in exercising their control.

Boris Feldman

Do you think that rules like that should be promulgated by the companies developing the technology or by government?

Whitfield Diffie

Something deeper than that. I think that it would take a change in human viewpoint to achieve that. The government should be the instrument of it. But the point is, people don't even think this way.

Vinton G. Cerf

Well, even if they did think that way, it's not clear to me that the government, speaking broadly about legislators, has the capacity to make rules that actually might work. I think we have to understand these systems a lot more deeply before we know how to influence their behavior in a reliable way. So at the moment, what happens is that we do everything we can to create and, to use the concrete example of a large language model, we create those models, and then we try to fine tune their behavior by interacting with them. And you may see in the recent report about a way in which they broke through the training information by telling the chatbot to generate the same word over and over again. I think it was "poem" and they had it generate "poem" 100,000 times and it finally started grabbing its training data and starting to disgorge the training data, which it was not supposed to disclose. So they're still fragile systems.

Whitfield Diffie

That sounds like a sort of automated buffer overflow problem.

Vinton G. Cerf

Right. Now in a very funny sense, it may have been literally, that. There's this thing called context. It's an amount of information that the chatbots use in order to generate their output. And when you overflow the available context, I'm guessing that's when you start, in this case, revealing your training data.

Whitfield Diffie

That's a good example. I mean, if we had the kind of what I'm envisioning that I think is impossible, right, there wouldn't be a rule that an AI can't keep secrets from a person. You ask it what its training data are, it has to tell you. And you say it's not supposed to disclose its training data. It has this sort of autonomy in which it's allowed to tell people where to go.

Boris Feldman

But Whit, assume that you could come up with the right answer for Western society. What good does that do if other countries in the world don't buy into it?

Whitfield Diffie

Well, I think the US has shown its attitude on this subject over the last 30 years in the sequence of wars which it engaged in. But seriously, I'm not sure I understand your question exactly, but in one sense I have no objection to cultural imperialism. If we adopt rules, you know, other people may well have to adopt them.

Boris Feldman

Some people have called for a freeze or a moratorium on LLM [large learning model] development. But don't you think that if we and the UK and the EU do that, that China or Russia or North Korea or Iran might not, and then we're on the losing side of an arms race?

Vinton G. Cerf

Well, there's a deeper issue here in my view, and that is the calling for a moratorium doesn't do much good. We really need to work with these things to understand how they work and how to manage their behavior. And to give you another example, I think it was in the seventies, or maybe the sixties, there was the Asilomar conference to set the rules on biological and genetic engineering. The idea there was that the practitioners actually chose to self-limit after discussion about the modification of genetic material, and that was propagated by the community well before any legislation was introduced. And so I could imagine the technologists looking at this situation with artificial intelligence, choosing not only to adopt some kind of constraints, but also figure out how to do that. And that's part of the missing link at the moment, figuring out how you introduce those kind of constraints. Even if the Asimov laws would not literally work, the concept that Whit is trying to articulate, I think, fits, and that is to say, built-in constraints. So if we learn how to design and build these systems with those constraints, that would be a good thing, I think.

Boris Feldman

Do you have hope that that could be accomplished?

Whitfield Diffie

Well, actually, there was an attempt to do that at another Asilomar meeting within the last 20 years or so that explicitly modeled itself on the biological meeting that gave rise to the notion that P1 through P4 containment facilities and what you would have to do to protect them what things were dangerous and what you would have to do to protect them. And it addressed some facilities on an island off of Long Island, because they can limit the access to the place much better than they can in the middle of Cambridge.

Vinton G. Cerf

I'm glad that you chose this particular analogy, because it resonates with me anyway, when you think about how you're playing with very powerful technologies and you need to be conscious of the fact that you could unleash something that might be quite harmful. So we should call attention to that, which we've just done in this interview.

Whitfield Diffie

To the three people who are going to watch our podcast.

I want to know something. You talked about the novella *The Machine Stops*. I presume you have read *A Logic Named Joe*.

Vinton G. Cerf

Yes, I have.

Whitfield Diffie

That, it seems to me, for predicting our situation, is utterly remarkable.

Boris Feldman

For those who are going to be too lazy to read it, can you summarize it?

Whitfield Diffie

A Logic Named Joe was written in 1946 by a man writing under the name Murray Leinster. I don't remember his real name. That's a pen name. And he imagined the fusion of television and telephone and that the result is very much like the web. And people sit on their telephone and call up entertainment, and this, that and the other. He foresaw the web, he foresaw the computer security problem. He didn't understand the computer security problem at all. And the logic name Joe—the devices are called logics rather than workstations—the logic name Joe was crazy, and it went around and tampered with everything. And they solved the problem ultimately by taking Joe offline and putting him in a basement somewhere.

Vinton G. Cerf

The question is, can we take anything offline anymore? These systems are becoming increasingly distributed and the artificial intelligence machine learning mechanisms are showing up in our mobiles as well as our laptops and desktops, to say nothing of the data centers.

Whitfield Diffie

And productivity will be the critical thing as it was with the biological issues.

Vinton G. Cerf

But I'm still positively excited about what I'm seeing these things are capable of doing in terms of just reinforcing creativity, triggering ideas that you might not otherwise have had. Some of the hallucination, so to speak, from large language models, is actually stimulating in the sense that you might not have thought of this.

Whitfield Diffie

People hallucinate well enough themselves.

Vinton G. Cerf

Well, some of our researchers are saying that it might be interesting to have interactions with these large language models just as a way of brainstorming for a while, not to rely on them to produce something, but just to get the juices flowing. That's not a bad notion.

Boris Feldman

So maybe in terms of terminology, just as you called it surfing the Internet—now you'll call it tripping AI?

Whit, I want to go back to your origin story. Were you a puzzle player or what? How did you get into cryptography?

Whitfield Diffie

In two ways. One is that in the sixties, I foresaw a world very threatening to individual privacy and thus autonomy. And I had a rather anti-governmental, counterculture viewpoint. And so I worked in the same building with the Multics project, which was the most important timesharing project ever. And it had a great deal of concern with security. I looked at the situation and I thought, okay, so even with all these protections on your file, the system programmers or the operators aren't going to not give it to the police when the police demand it. As so I imagined the only thing you could do to protect your own information was to encrypt it. So in the late sixties, I tried to talk a variety of friends, who wished I had succeeded, into working on cryptography, because I thought it was the critical thing. I was working on something I considered and still consider more important, which is the proof of correctness of programs.

Vinton G. Cerf

Which is still very important notion.

Whitfield Diffie

Look, I'm very grateful I got rescued by cryptography because the proof of correctness hasn't made as much progress, and I wouldn't have made the difference. But then in 1972, there was a wonderful accident. So we're part two of the answer. Larry Roberts, who was funding the ARPANET, went up to NSA to see a man named Howard Rosenbaum, who was the deputy director for research, with the obvious proposition, hey, man, I have a $100 million-a-year military communications research experiment going on. We ought to think about security. And they can't have disagreed about that, but they seem to have disagreed about secrecy, because Roberts' part of ARPA didn't want to fund classified research. Other parts of ARPA did a lot of it. But the information processing techniques office was very open, and Howard Rosenblum didn't want to do anything else. So Roberts goes back to his office in Roslyn, and for the next week or so, he has this great job. His principal investigators come by with their hats in their hands and have to talk about whatever he wants to talk about. So next week, he's talking about network security, which at the time all of us viewed as mostly a matter of cryptography. We would see it differently now because of cryptographers largely solved that problem, and the rest of it is a complete mess. So one of the people who went by to see him was my boss, John McCarthy. And McCarthy got the bug and came back out of the laboratory and talked to us about security, which we saw as cryptography. And a few people got interested. But basically, Hans Moravec and I did the most on it, and John McCarthy had designed a cryptographic program and Moravec coded it for him. But Moravec added what would later be called key escrow because he figured if his thesis advisor wanted to encrypt something, maybe he would be interested in it. And he wrote the program that was later came to be called the shrinking generator. It was rediscovered at IBM a few years later. I started working on it and I operate in a very different way and I basically set out to read David Kahn's history of cryptography. And I also wrote programs and things like that. But the basic thing I could understand was what the requirements for a cryptographic system were. And by March of '73, I was doing nothing else. And John McCarthy was fed up to his back teeth because I was being supported by under-the-table money from NSA to work on proof of correctness.

Vinton G. Cerf

So this was at Stanford.

Whitfield Diffie

Yes, at Stanford AI Lab. Before I met you. So in March '73, we amicably parted company, truly amicably, as we will evolve a lot later. And I went off intending to travel around the world and thinking about this problem and talking to people I could find to talk about it, and digging up rare manuscripts. And that got as far as New Jersey, where I met my wife, and we then traveled together for a couple of years until I went to visit IBM—Watson—which had the only known governmental significant cryptographic laboratory in the country, and I talked to Al Konheim, who was the head of the math department, which was a lot of this, and he said, "I can't tell you anything. We're under a secrecy order here." He only told me one thing. And then he wished he hadn't told me that. He said, "Go look up my old friend Marty Hellman when you get back out to Stanford."

Vinton G. Cerf

I didn't know about that connection.

Boris Feldman

What is the patent called? Diffie-Hellman encryption.

Whitfield Diffie

So Hellman and I then immediately found each other the best informed person willing to talk about the subject. You get found. We worked together for four years and became a great pain in Howard's ample backside because we raised questions about the system that laboratory designed—data encryption stat.

Boris Feldman

Had anyone been talking about public key encryption before the two of you?

Whitfield Diffie

Not to us. The earliest person I can find to have the idea is James Ellis, at GCHQ, and his first notes are from late 1969.

Vinton G. Cerf

Oh! That early.

Whitfield Diffie

Yeah, the paper was in 1970. But there are some slides from '69.

Vinton G. Cerf

Was there someone named Cooper that was part of GCHQ.

Whitfield Diffie

I don't know that name. Cox was involved. Malcolm Williamson developed something that is essentially Diffie-Hellman, and Clifford Cox was still alive and developed what is essentially RSA. So those people, it's interesting to me, started ahead of us and actually finished afterwards. Williamson's internally secret paper comes out two months after I gave a talk at the national media conference about the same scheme.

Boris Feldman

How was public key encryption being used before it was built into the browser for security?

Whitfield Diffie

The first piece of equipment was made by Racal COMSEC, I think. They didn't do it quite the right way but they roughly used RSA to develop ephemeral keys between link encrypts...

Vinton G. Cerf

Actually, if I could interject for a moment. I was working with NSA starting around 1975 on the problem of using cryptography to encrypt packet communications. And the primary challenge there was that the packets could arrive out of order and you still needed to decrypt them. So you needed something called a message indicator to tell you where in the key stream that your packet belongs. But we were using cryptographic equipment at that point that was essentially intended for link encryption, continuous transmissions, and we had to adapt to that. And as we started poking our way into this, the digital encryption standard—DES—came along, which was designed to do this kind of function. And so we used that to build a packet encryptor in order to study how we could use not public key cryptography, but DES in a packet mode. But then came the key management, and it was a nightmare because you

had to somehow mechanically deliver keying information to all of the various parties.

Whitfield Diffie

BCR incorporates that.

Vinton G. Cerf

We called it "black crypto red" or BCR program that BB&N [Bolt, Beranek and Newman] ran while I was running it at ARPA and earlier at Stanford. But then when these guys came out with something that you could use to do key distribution that didn't require a physical presence, it was a dramatic change for the utility of using cryptographic techniques to protect the information.

Boris Feldman

Apart from intellectual curiosity, was your primary motivation individual privacy?

Whitfield Diffie

Roughly, yes. I was not doing applied research. I was working on what would much later be called STU-III, third generation Secure Telephone Unit. There is a little bit of difference. They were never more than about half a million STU-IIIs. I was working to secure all the telephones in North America, which I imagined to be about 100 million. And therefore you'd need approximately the square of that number of keys. And that's sort of the line of thinking.

Vinton G. Cerf

And of course, now, how many billions of mobiles are there?

Whitfield Diffie

Oh, I don't know. There are several billion browsers and sure, several billion mobiles.

Boris Feldman

When did you realize that this thing you'd invented would have commercial application on the Internet and the web?

Whitfield Diffie

I didn't think so much in terms of commercial. I realized immediately. And you can find an interview with my late wife, Mary Fisher, for the RSA series of video interviews. She came over and I said to her, I have discovered a great thing. For the first time in two years of working on this assignment, okay, this is important.

Vinton G. Cerf

The paper that you and Marty Hellman wrote was called "On New Directions in Cryptography."

Whitfield Diffie

No "On." Just "New Directions in Cryptography."

Vinton G. Cerf

And what was exciting to me about that, that's 1976. And this is just as we're getting the TCP/IP protocols sorted out, we're still in the process of going through iterations and trying to get it to work. But it was very clear when that paper showed up that it was going to have a significant role to play in security systems. And some people said, well, why didn't you put all that stuff in right away? And part of the answer to that is that in order to manage the keys, you'd have to rely on the behavior of graduate students, because they were the ones doing most of the work on the Internet. Graduate students get distracted by term papers, final exams, and getting their theses written. And so I thought, well, why don't we wait, because it was clear we could retrofit that into the architecture, which we in fact had done.

Whitfield Diffie

I think it was just incredibly lucky that you didn't have the machinery to make a secure network then, because it would have become what Paul Barron had envisioned, a national command and control network. The openness of the Internet is the thing that has made it a great economic and cultural force. If it had the power to keep people out, it would keep the startups out.

Boris Feldman

Has the openness of the Internet put us in a position now where the security risks are so great that we need to redesign it from ground zero?

Whitfield Diffie

Well, I have another place where I have an unorthodox opinion. I think the security risks are a function of the fact of the big powers, not NSA, but you and Apple and Microsoft and other people don't want security, because if the network was secure, if individuals had security, then the companies would not have the degree of control over them that they want. Now, let me give you a perfectly simple example. I think you can reasonably say that let's say if you buy the notion that parents ought to have the right to control what their children see, then they ought to have the right to look at a movie and record it and then show what they recorded to the children. But Hollywood doesn't want that. It wants to charge you to see the movie every single time you see it, or maybe allow you to see it, or maybe not.

Vinton G. Cerf

Actually, no. And here I have to differ with you on several points. The first observation is that Netflix lets you see a movie as many times as you want as long as you pay a fixed fee to subscribe to whatever they've got. So that's first.

Whitfield Diffie

But they try to keep you from recording.

Vinton G. Cerf

Actually, oddly enough, they used to send DVDs out. They don't do that anymore because it was getting expensive, shipped through the mail and they had a decreasing subscriber base with that particular delivery. Yes, there is screaming and yelling about recording, although we know that people can do that. But the argument about Google and others not wanting cryptography…

Whitfield Diffie

No, I didn't say you didn't want cryptography. I said you didn't want the users to have security. That's a different thing.

Vinton G. Cerf

No, that's not true either. At least some of our product line in the Google Cloud.

Whitfield Diffie

You want them to have a certain kind of security.

Vinton G. Cerf

The Google Cloud products actually have mechanisms in them to allow our customers to use their own key variables that we have no knowledge of at all. So they can encrypt their data, they can encrypt their software.

Okay, who is not here? There's nobody from Apple? Nobody from Microsoft. They must be the bad guys.

Boris Feldman

No, no, but I should have probably had a disclaimer at the beginning. Vint is here as Vint Cerf, not as the Google representative.

Vinton G. Cerf

Well, he's right. It's impossible to separate me from [being a Google executive].

Whitfield Diffie

It's also good because if you say no, Google has such and such, you know much more about what Google has than I do. That's important. So that isn't being their stooge.

Boris Feldman

But on the core issue, is it your view the Internet is secure or you think it isn't, but you blame it on "big tech"?

Whitfield Diffie

I blame it on a lack of enthusiasm for what I think of as real security, which puts vastly more control in the hands of individual users. Almost no individual users have the energy and the erudition to manage such control.

Vinton G. Cerf

So here's where we might differ. I think, or we might iffy… my minus two… I know…bad pun…

Boris Feldman

He comes from a long line of punners. We have to say, if you Google Bennett Cerf, a distant, legendary relative of Vint's, he was one of the great masters of the pun in American history.

Vinton G. Cerf

There are many books in my library written by Bennett Cerf. The point I am making is that a lot of what you were talking about is related to privacy, which I think is a very legitimate concern. Security and the use of cryptography is of great concern at Google and many of the other providers as well, because if they don't find ways to secure the system so that their consumers and customers feel that their data is safe, that's a big problem. So we were the ones who pushed very hard to insist on HTTPS, for example. We pushed very hard for DNSSEC [Domain Name System Security Extensions] and we pushed very hard after 2010, especially to fully encrypt everything that's in our data centers. So that if even if the data centers are breached, data is already encrypted. So we do consider and I personally consider that cryptography is our friend here in many different dimensions. We could probably argue about privacy. One of the things that we've just recently added is to try to get rid of third-party cookies precisely to improve privacy.

Whitfield Diffie

I've been shut up.

Vinton G. Cerf

Well, we don't need to have huge argument. Well, we should probably figure out what it was that you skipped over to get to the …

Whitfield Diffie

He sent us all those questions.

Boris Feldman

I want to talk to you about truth. One could argue that the Internet has undermined the concept of truth, especially

in recent years, given its role in disinformation. Is that something that you foresaw? Do you think there's any solution to it, or is it just something we have to live with?

Vinton G. Cerf

I would like to hear what Whit has to say about that. But I would like to interject something. The way you phrased that, you said it's the Internet that is responsible for misinformation, disinformation. I beg to differ. When you say Internet, I think of the road system that connects the computers together. You're thinking about what people are doing. And it's the people that are causing the problem. And in some cases, it's the technology that helps them do that. So the amplification, social media, the misinformation and disinformation that can be generated with large language models and machine learning are all mechanisms that people actually exercise. And so I would appreciate if we didn't just blame the Internet for everything and forget about the fact that people have some control over what they use it for.

Boris Feldman

Of course, you're right.

Whitfield Diffie

So, I think one effect of the web more than the Internet is disintermediation would be the jargon in information flows. So the critical control of the flow of information was exercised in the 19th and much of the 20th centuries by editors. And I couldn't get something out to a million people unless I could persuade the editors of the New York Times to publish it. Today, I have many avenues, and I might very well succeed, as you may succeed with this podcast in getting it to go to a million people. And there is nobody with the same education, investment in job, investment in society, as the editor of the New York Times, having it fact checked and preventing it from going out.

Vinton G. Cerf

I'd like to reinforce what Whit is saying in the following sense. The other major media had similar characteristics. You had the ability to reach a large number of people that had infrastructure in the case of newspapers being able to print, and distribute paper. But that's also true for radio and for television. You had significant infrastructure required and your access to it was very limited. Only people who controlled the television transmitters or the radio transmitters or the printing presses had the ability to reach this large-scale audience. The Internet dropped the barrier to access, and the social media simply enhanced the potential for reaching large numbers of people. So I'm in agreement that while we thought it was a good thing, we thought that dropping the barriers to access to information and the ability to share would have a beneficial outcome, and it has had a beneficial outcome. Think of the things that you can find that you otherwise wouldn't be able to find very conveniently by just sitting down and searching.

Whitfield Diffie

That's just a truly wonderful feature of the modern world.

Vinton G. Cerf

But the side effect is that these same tools have allowed people to distribute misinformation and disinformation, and to make matters worse, the latest version of machine learning, the large language models, hallucinate and generate their own misinformation.

Whitfield Diffie

Well, I think there are natural tendencies to bad flows of information among people. It's very easy if somebody tells you something and it sounds believable enough, and you repeat it without checking on it. That's something I believe probably everybody does. I have done it many times. And we're at the point now if you to add an amplifier to that, then you get the chance…. Misinformation can be innocent or malicious.

Vinton G. Cerf

Can I give you an example of something? This has to do with trust. You were asking about truth. I wanted to talk about trust. Some of the indicators that we have historically used to decide we should trust this piece of information have become distorted. One of my best examples of this is a reporter who had an interaction with the chatbot in text and decides he wants to make a podcast out of it. So he takes his prompts, his queries, to the chatbot and he reads them in his voice. Then he keeps the text coming from the chatbot and translates it through a text-to-speech system. But he's able to choose the voice that is used, and the voice he chose is David Attenborough. So now you have the problem of this reporter interacting with what sounds like David Attenborough, and you know, he's well regarded. He has his polished onscreen accent. Anything said in the voice of David Attenborough sounds great.

Okay, so I'm thinking, okay, we know we have a problem. What should I use as indicators to trust information? But I wanted to offer another old example of this that didn't occur to me until I saw this exchange with the reporter. And looking back at Xerox PARC, I think about 1974 or 1975. So what did they have? Well, they had the Alto workstation, a $50,000 workstation. Everybody had one. They had the Ethernet that Bob Metcalfe and David Boggs invented…three megabit system at a time when 1200 baud was fast, they had bitmap graphics, they had the Bravo word editor and they had a laser printer. Okay, so just imagine they were sitting there in 1973 or 1974. They're living 20 years in the future. What happens? You prepare first draft material and you print on a laser printer left and right justified, beautiful font. You know, it looks like it's final draft stuff. Well, the indicators of effort going into it are wrong in this case because the thing looks so good, and we're very confused. We need new indicators that tell us whether we should trust stuff.

Whitfield Diffie

Yeah. It looked as though the investment had been put into it that would have gone into producing a printed article or a published book, and of course it was just some…

Vinton G. Cerf

A shopping list.

Whitfield Diffie

I predicted something 40 or 50 years ago that hasn't come true yet. I thought by now an individual working alone would be able to produce a color movie of the quality that we had back then. So I'm thinking now it's all going to go to the authors. If you have the brilliance to produce, I don't know, Sound of Music. If you can think it all up, then you can use your tools just to produce it in a day or two of work in your studio and you don't need a cast of a thousand, etc.

Vinton G. Cerf

We're getting close to that.

Whitfield Diffie

I'm not saying it's not going to happen… On the other hand, I was right in thinking by this point we'd have roughly gigabit connections between two people. But the point is we are moving in this direction of more and more stuff that you can't judge the truth of or the trustworthiness of by what at one time would have been the cost of producing.

Vinton G. Cerf

Yes, this is actually very important. And so that says provenance has become a more significant element in determining trust, also corroborating evidence or other kinds of mechanisms that help us judge whether we should trust the information or not.

Whitfield Diffie

AIs ought to be the right tool for that, but they are operating in an entirely different way.

Vinton G. Cerf

Yeah, I think that statement alone should be one of the North Stars of our work in artificial intelligence, which is to use the tool to help us judge the trustworthiness of the information that we're getting.

Boris Feldman

You need something on both the supply side and the demand side. So you may be able to use the AI tools to give you information on provenance, but you still need the consumer to care about it and look at it. And there's a whole separate question about whether people want that or they just want to be in the echo chamber.

Vinton G. Cerf

Well, I don't even think it's an echo chamber. It's the same as what Whit described earlier about the willingness to forward things to other people without having taken the responsibility for designing or checking to see whether it's valid. There's this website called Snopes, which debunks a whole lot of assertions like the post office is going to try to charge two cents for every email and things like that. But a lot of people don't take the time. I'm guilty of this.

Whitfield Diffie

We spend enough of our time, but you know, you have, I think, a serious point here. If you can't, so to speak, have things you can trust, then you're willing to spend all of your time checking on the veracity of things. And it's a little like a sort of, you know, forgive a high tech example. One of the big important things in guaranteeing the correctness of programing is to check values to see that they're the right sort of thing. But also if you do that on every possible module, you sort of waste a huge amount of time. You need a mechanism for saying we trust that one and we just take the data it sends and operate on it as though it were what we expect it to be. Otherwise you spend 90% of your time doing useless structure checks.

Vinton G. Cerf

You know why this is so damned hard in the network environment? Every time a computer on the Internet interacts with another computer on the Internet, it's probably an event which has never happened in the history of the universe. And the reason for that is the state of this machine and the state of that machine have never been in that state before. Just the mere fact that maybe they downloaded some new software, maybe there's an update, maybe just the content of the machine is different. The result is that you can't predict all the various situations in which these interactions could take place. Whit's right in that we need to find some way of learning to trust some portions of the software or some portions of the data that's flying around, because if we can't do that, how will we ever learn to trust anything again?

Whitfield Diffie

The reasonable question has been asked by an article that, is Britannica or Wikipedia more trustworthy? And they both have errors in them. And there are different mechanisms by which they might be.

Vinton G. Cerf

But at least you can make corrections in Wikipedia. And it was very hard to make corrections in Britannica.

Whitfield Diffie

That's right. That's right. And that's both good and bad. It's very hard to make changes in Britannica.

Boris Feldman

So before we turn to mentoring, I think what you two started to collaborate on is an update of Asimov's rule for robots to Vint and Whit's rules for AI, of which at least two of them, maybe they're the same one. One is trust, and you have to flesh that out, but that's the topic, and the second one is provenance and an ability to find out not just what trained the model, but what they've included, how they've made decisions on that. So hopefully, maybe during lunch today, the two of you on a whiteboard will start sketching out these rules of the road.

Vinton G. Cerf

So, you know, that's a really interesting, hard problem if you say, well, you know, if you knew what the training data was, you ought to be able to figure out whether you could trust the output. And at the moment, I don't think that's true. So I have another work example for you. You've heard this one, I think, before. I asked one of the chatbots to do an obituary for me, and I thought that was reasonable because the Web has lots of obituaries and the bots would have learned that format. And there's stuff on the Internet about me, so it should be able to do that. And it did. It produced an obituary of 700 words. And it starts out we're sorry to report that Dr. Cerf has passed away. And then it gave a date, which I found very unsettling, and then it went on to summarize career and family. Well, it conflated things I did with things other people did, gave me credit for stuff I didn't do, gave other people credit for stuff I did. Then it named family members that I don't think I have. And I remember thinking, how the hell could that happen? And then I realized that when the bots are assembled, they're plucking stuff out of web pages, and Whit's bio and my bio could very well be on adjacent web pages or on the same web page. And the system that's absorbing all that doesn't know that this piece came from Whit's bio that piece came from mine. And so the actual data could become conflated.

Whitfield Diffie

That seems to be one error that we can avoid. But I can easily believe that something might go to a source that discusses, you know, something like this interview. And there will frequently be ambiguities as to who's the referent of a pronoun. Working conscientiously and honestly in an attempt to get things right out of ordinary sources of information is frequently difficult.

Boris Feldman

Whit, I don't mean to intrude on your privacy, but have you had an LLM-generated obituary prepared?

Whitfield Diffie

Actually, I went as far as getting John Markoff to ask ChatGPT what the date of my death was, and it gave a very sensible answer. It said he wasn't dead as of the end of 2021. Okay. That's the point to which it was trained. And I thought, well, that's a nice, solid, sensible answer. Well, I haven't gotten the full obit[uary].

Vinton G. Cerf

I'm actually assuming that Katie Hafner will probably do mine because she's responsible for the geekaverse at the *New York Times* for a lot of the obits. And so she prepared some ahead of time.

Whitfield Diffie

She's John Markoff's ex. And he used to write a lot of that. I think he wrote [Steve] Jobs obituary. At some point I said to him, have you written mine? He said no, but we're getting around to it.

Boris Feldman

Tell him no rush. I want to close by asking you for mentoring, for the many people watching this. And first, I want to ask it in an unconventional way. Along the way, what career or life advice did someone give you that you ignored or didn't follow, that you wish you had followed?

Whitfield Diffie

Well, I'm the other way around. The best I can think of is a piece of advice that Danny Barbaro's younger brother, Rusty, said to me, when we were 15 or 16. You and I will never go into space. We're not good enough physical specimens to be in the first wave and would be too old for the second wave. Well, that seemed sensible to me. But the fact is I was a good enough physical specimen. My eyes were incredible. My fault with my being an astronaut had to do with my study habits and things like that.

It's not anything like a research job. It's a job in which you have do what needs to be done at this moment. So maybe I would never have become an astronaut. But I listened to his advice and it seemed to me to be reasonable. I listen to it. I didn't set my goal to cure the personality problems that stood between me and astronautics. And so I never got to do that.

Boris Feldman

So your takeaway from that is?

Whitfield Diffie

Don't ever believe there's anything you can't do. I think it's much more to lots of people this problem is too hard. Don't bother to work on that one.

Vinton G. Cerf

Well, you know, Captain Kirk went up at the age of what? At 90? That was pretty amazing.

Boris Feldman

Vint, was there any advice you didn't follow that you wish you had, or is it the same thing as with Whit, advice you had but didn't take?

Vinton G. Cerf

I think actually it's closer to Whit's thing, which is advice I was given that I had taken to heart. And this came from Josh Lederberg, the Nobel Prize winner from Stanford University in recombinant DNA. He was on the board of the Corporation for National Research Initiatives that Bob Collins started in 1986, and I joined Bob that year, and I was working on digital libraries at the time, and I was describing to Josh Lederberg what it was that we had in mind to do. And so I covered the white board with all the aspirations and everything else in it. And at the end, Josh looks at it and he says, Vint, do something. And, you know, I've taken that to heart. Don't just talk about it. Do something.

Boris Feldman

If you were restarting your career today, what field would you go into?

Whitfield Diffie

I would study law and work on economics.

Boris Feldman

We do have a summer associate program.

Whitfield Diffie

My answer is based on two things. Law is one of the two big regulated professions. There's very little you can do in the direction of law without becoming a lawyer. It's also true of medicine, but that one doesn't interest me.

If you're going to put the time into education, the point in getting a PhD in mathematics or physics, of course it makes your life easier. You do actually learn things in those classes. But having the credential of being a lawyer allows you to do a lot of things. Economics is the utter disaster. It's the most remarkable field because it's about a human product that is utterly out of our control. And so some great discoveries are needed in economics. And that's

one place to go to work.

Vinton G. Cerf

So I have two reactions to what Whit just said. The first one is that the one word in economics, which I consider very offensive, is externalities, because what that means is we didn't know how to deal with that, so we're ignoring it.

Whitfield Diffie

I didn't say that, did I?

Vinton G. Cerf

No, you did not. And I'm not accusing you of having said that, but it comes up a lot in economic discussions. So I think if we're realigning my career, knowing what I know now, I would probably want to go into what we might call computational microbiology. I'm utterly fascinated by how cells work. And I've learned a small amount about this from Bruce Alberts, who's written this gigantic book on the microchemistry of cells. But I used to think of a cell as a little bag full of chemicals, and they banged into each other and stuff happened. Well, I have now learned that if you look at a cell, it has extraordinary order. It's as complicated as downtown Manhattan. All kinds of things are going on. There's communications that's happening among the organelles and between the cells. It's an absolutely fascinating world. So I think if I were starting over again, I would love to go down that path.

Whitfield Diffie

But I don't think you should say computational. But yeah, microbiology is one of the great things to work on. Language is still utterly not understood. What seemed an incredibly optimistic situation coming into the sixties with Chomsky's discoveries have just basically gone nowhere, and large language models aren't giving us any great insight into how people talk. The problem of gravity is totally unsolved in physics. And I explicitly wouldn't particularly advise people to go into cryptography, although the field amazingly keeps being reborn.

Vinton G. Cerf

But actually people do come in and they say, What should I study? I tell them astrophysics. And the reason is very simple. A hundred years ago, plus or minus, we thought we knew pretty much how the universe worked. We just needed to measure the constants more accurately and we could make better predictions. And Einstein comes along in 1905 with his four papers. He blows up physics and then the, you know, the guys who do…

Whitfield Diffie

string theory, probably

Vinton G. Cerf

Before string theory, we just get quantum electrodynamics, then you get quantum chromodynamics, then you get the string theory guys blowing up the quantum chromodynamics guys, and then what happens? We have these guys looking out at the universe, trying to figure out what's going on, and they all know that the universe is expanding. And they started measuring the expansion rate, which they expect should be slowing down by now. After all, the universe is, you know, 12 billion years old. And they discover it's accelerating. And so your first reaction is WTM? And so the question is, why is it accelerating? And the answer is dark energy. What's that? We don't know.

Whitfield Diffie

What I do know is that they named it wrong.

The last time this notion was popular was in the 19th century. It was called levity. So levity is what pushes things apart. Gravity is what pulls together.

Vinton G. Cerf

So here we are. We had 70% of the universe as dark energy. And then the galaxies should be falling apart based on their estimates of mass in the galaxy that's visible. So there must be invisible mass holding it together. We call that dark matter. What's that? I don't know. Ninety-five percent of the universe is unknown. Five percent is ordinary matter, including antimatter. And so when a kid says, what should I do? I tell him, go into astrophysics. The probability is that you get the Nobel Prize for anything you do because we don't know anything.

Whitfield Diffie

Just incidentally, galaxies aren't explained by the fact they have large black holes in the middle of them.

Vinton G. Cerf

In fact, it's now starting to look like there're lots of black holes of varying sundry sizes. But that isn't what's holding the galaxy together. It's this other thing, dark matter. Unless, of course, you subscribe to the theory that the gravitational equations are wrong in that they require a slight modification. It's called MOND [modified Newtonian dynamics], I think. It's a slight modification of gravitational theory which would account for the way things work without requiring dark matter.

Boris Feldman

You've been very gracious to be here today and I look forward to your publication of the Vint and Whit Rules of AI Safety. Thank you. Thank you all for joining us .