The internet at 40
A brief history of the future

Penny Boreham
Hello, I'm Penny Boreham, and I'm here with John Naughton, professor of the public understanding of technology at the Open University. He's also author of 'A brief history of the future – the origins of the internet'. John, 2009 could be seen as the fortieth anniversary of the internet, which you've described as one of the twentieth century's greatest inventions. What's so special about it, and how did it all come about?

John Naughton
Technically the internet is a network of computer networks. It's not just a computer network, it's a network of networks, and that's quite significant. What sets it apart from anything we've had before is the fact that it's the first serious communications network we've ever had in history, which is not owned or controlled by anyone, and which anyone can use so long as their computer speaks the required technical language. We've never had anything like that before.

It emerged originally from a project funded by the US Department of Defense.
The Pentagon in the 1960's was paying for a lot of expensive, incompatible mainframe computers which couldn't talk to one another, and the original idea was to devise some network which would enable these very diverse, very expensive, incompatible machines to talk to one another and from that, the basic technology which became the internet evolved. That project began in the mid-sixties, the first of these networks, the ARPAnet, started in 1969, so we're heading into in a way the fortieth anniversary of this kind of networking technology.

Penny Boreham
So what was the ARPAnet actually meant to do?

John Naughton
The idea was that a researcher in North Carolina who's paid for by ARPA would be able to log into a machine at MIT, or a machine in Caltech in California, or whatever, and use it. That was the original idea, so it was to make better use of very expensive and scarce resources, and everybody thought at the beginning that's what it would be used for, but the astonishing thing they discovered very early on was that actually there was very little of that intended use, and instead of that people who were on the ARPANET were using it for what many of the original funders thought was a rather frivolous purpose, they were sending messages to one another across this network, they were using this amazingly expensive facility designed to enable people to log into mainframe computers, and instead of that they were sending what became email, electronic mail, and that really did take the designers by surprise. If you know anything about human beings, and of course the problem with many engineers is that they don't, and I'm an engineer so I can speak with authority on that subject - they underestimated the fact that human beings are fundamentally compulsively social individuals and if you give them a really good communications tool they will use it for communicating with one another. In retrospect it ought to have been obvious, but for a time it wasn't.

Penny Boreham
Was anyone else working on networks forty years ago?

John Naughton
The development of the network went in stages. The first stage was the ARPANET. That was built between '69 and '72, and by the end of 1972 it was effectively complete. While it was being built a number of other networking projects were under way in different parts of the world – there was one in Hawaii, there was one in France, and there was one most

interestingly in Britain in the National Physical Laboratory, led by Donald Davies and his colleagues.

The point about these networks is that although they all had a family resemblance to the original ARPAnet in the sense that they used the same underlying technology, they were all radically different in some respects and, more importantly, they were not owned and controlled by the US Government. So when the ARPANET project was completed in '72 the obvious question was well, what do we do next? And the decision was we'll try and seamlessly network these other things, in other words we'll inter-network them, that was called the Inter-networking Project originally, and it was led by Vint Cerf and Bob Kahn and it started to work in 1973. The basic idea they had was we have to find a way of seamlessly linking networks that we don't control, and how do we do that?

They also faced a very difficult design decision which is how do you design a network that is future-proof, how do you design something that will work for applications that nobody's ever dreamed of yet? That's a fantastically interesting question, it's a fantastically difficult question, and that was the design challenge that they faced. And they came up with the solution which was brilliant and simple, and the solution basically said that the network we'd build would have two main characteristics. The first is that nobody will own and control it, it will be in that sense ownerless, and that's a big change from the past because up till then all our communications networks had been owned and controlled by somebody, usually governments, or state monopolies of some kind. The second thing is that we will not design it to be optimised for any particular application, we'll make the network simple and effectively stupid. It will do only one thing, it will take in data packets from one end and it'll do its best to deliver them to the other end, but it won't care what's in the data packets, it won't even ask, it'll just do its best. And in that way you build a dumb network and you leave all the intelligence to the end, to the extremes of the network, the idea being that essentially if somebody was bright enough to have a clever idea that could be achieved by using data packets, then the internet would do it for them, no questions asked, and that was the big change, that's what led to the explosion of innovation that we've seen.

It led to a whole list of surprises, some of them very pleasant surprises like the World Wide Web, and some of them unpleasant surprises like malicious software and malware, and some which have good and bad sides like for example, file-sharing, music file sharing, which is what has undermined and taken the music industry somewhat by surprise. All the ingenuity was left to the edges of the network. Anybody could play, anybody could use it, for whatever purpose they thought of, so long as it could be done using data packets. And that's the key to understanding the network, that it became a huge global machine for springing surprises, and that's why it has sparked such a wave of innovation, sometimes very disruptive innovation, innovation that has distressed some industries and may yet destroy others, and created whole industries that never existed before. All of that is possible because of that original design decision to have a network that had no central control, and which did something very simple, wasn't optimised for any particular application.

Penny Boreham
So, is that why there still seems to be some confusion about what the internet actually does?

John Naughton
The interesting thing for me is the extent to which the internet is still fundamentally misunderstood, often by people who ought to know better, people who have serious corporate or political power in our societies. The big mistake most people make is they think that the Web and the internet are the same thing, and nothing could be further from the truth. The best way of thinking about it is to use a simple metaphor or analogy. If you think about railway systems, then the internet is the tracks and signalling on which different kinds of traffic run. One important kind of traffic is web pages going from computer to computer; another kind of important traffic is email; another one is instant messaging; another one is voice-over IP, telephone conversations; another one is streaming media like the BBC iPlayer or what you get from iTunes, and so on. But these are all different kinds of traffic and they run on this underlying architecture, this underlying kind of infrastructure. The internet's far more important than anything that runs on it.

Penny Boreham
What sorts of traffic ran on the network in the early days?

John Naughton
The internet first of all came into being in January 1983, and from '83 onwards there was actually an awful lot of stuff on the internet, there were an awful lot of files, there were a lot of publications, there was even music, all over the place, and wherever there was a research lab with an internet connection there was a server containing large depositories of various things, including data from CERN, for example. Although all that stuff was out there, to find it an access you really had to be a geek, you had to understand the incomprehensible gobbledygook that you have to type on a command line in order to locate it, you had to know the address of the computer where it was held, you had to know the name of the path to the directory in which the stuff was held, and all that kind of stuff, and then you had to be able to organise file transfer from one to the other, and all the rest of it, so it was not a system suitable for ordinary human beings. It was fine for computer scientists, fine for engineers, fine for geeks, but not suitable for human beings.

Penny Boreham
Is this where Tim Berners-Lee comes into the story?

John Naughton
Tim Berners-Lee is a very interesting case because first of all he's not a computer scientist originally, he's a physicist, and secondly he has a very bad memory. And many great computer inventions have been created by people who are trying to solve their own problems. In Tim's case he was working for CERN, the big part of the research laboratory in Geneva, and CERN's a very complex place, it has an awful lot of experimenters coming all the time, they come for short periods, they do very complex experiments, then they go away to their universities in other countries, or their host institutions, or whatever, and so there's a constant flow of people and a constant flow of data, huge experiments going on and they produce large amounts of data, the data's stored on the computers. At any given moment CERN has the same kind of memory problem as Tim Berners-Lee had himself.

His idea was, can we use software to help us find, locate, and retrieve information, and he had a shot at doing that in the late eighties, and then in I think '89 he made a proposal to his bosses that he, and a very small group of colleagues which included Robert Cailliau, would essentially create a huge global document location and retrieval system, which would run on the internet, and for want of a better name, they called it the World Wide Web. The intriguing thing about it is that the Web is probably now the largest transformation in our communications' environment since the invention of print. But it was almost exclusively the creation of one man and a small team of colleagues, and they did it without asking anybody's permission other than their bosses. They asked permission to devote some time and resources to the project, they were allowed to buy a couple of fancy computers to work on it, but when it was finished, when they wrote the software and the architecture for this thing they called the Web, they released it on the internet and they didn't have to ask anybody's permission. They just put it on servers and it went from there. It's really significant in the history of the world to think of a very small group of people led by a single individual which could have such a gigantic impact in such a short time.

Penny Boreham
How does Tim's big idea work?

He said, what if we could find a way of providing normal human beings with a kind of a window onto this that made sense to them? That was the original idea behind the Web. What a browser is is effectively a kind of window - when you look through it you see files on servers stored in some location. You don't have to know anything about that location or anything else, you just have to see that this file is here, and you just click on it, and all the computer gobbledygook that you need in order to make that transfer happen, happens behinds the scenes, you don't see it. If you're somebody like me who used the internet in the '80's, and who knew how hard it was to find and locate stuff, it's a miraculous breakthrough, it was a wonderful idea. And he also took a set of older ideas, which were called hypertext, in order to

devise a way of organising this way of looking at the internet and its resources. It's very sobering when you think that the first webpage probably was published in 1991, and I'm sitting here today and nobody knows twenty years on how big this system is, you know, there are educated guesses which talk about 700 billion pages which is mind-blowing, absolutely mind-blowing and it all happened because he had this great idea and he and his colleagues were able to implement it, and they were able to put it on the internet, and not have to ask anybody's permission.

John Naughton
There's no doubt it's been a fantastic success, but are there elements of it that have not yet matched up with Tim Berners - Lee's original dream, do you think?

Once it became clear in the early 1990's that the Web was going to be something big, that people were going to publish a lot of stuff on it and therefore other people would want to find and use it, it became absolutely imperative to find a way for locating stuff. It became clear pretty early on, I think it was clear to Tim and his colleagues, that the usefulness of this World Wide Web that they were creating would depend on the ability of people to find what they were looking for, and that meant the evolution of a new kind of computer program, a new kind of automated software system, which we now call search engines. There had been a number of early search engines on the old internet, the pre-Web internet, some of them had funny names like 'Gofer' and 'WAIS', and so on. They were a big deal for a while but once the Web arrived it was clear that they weren't up to the task and that a different kind of technology was needed. The basic idea was of creating a software robot, sometimes called a 'bot', which did nothing except crawl the Web 24 hours a day, 365 days a year finding pages, indexing them, creating us a huge database which could then answer queries, that's the origins of the search engine idea, and an awful lot of the development that we've seen on the internet and on the Web since then has been in one way or another connected with search because search is the big problem.
When you have a universe of information as big as the Web its usefulness is limited to humanity if people can't find what they want. Tim Berners-Lee from the beginning, I think, had a vision about this which was that the Web pages would have to be more informative than they initially were. His idea was that it should be possible for a software robot, a bot, to look at a Web page and be able to make some intelligent inferences about it. So, for example, a brute force search engine will not be able to distinguish between a Web page which is about the film *'Casablanca'* or a Web page which is about the city of Casablanca. Tim had this vision of what he called the semantic web, which was web pages structured in such a way that intelligent inferences that would distinguish the city from the film would be possible. That dream has not been realised, although we have moved towards it to some extent, but that remains a vision for the future, for the time being. What has replaced it is much better search engine technology, largely driven by Google.

Penny Boreham
Have all the surprises that the internet has generated been as beneficial, would you say, as the world wide web?

John Naughton
The problem is that innovation is no respecter of character so you have good innovators like Tim Berners-Lee and his colleagues who created this wonderful thing called the Web, but the internet is also hospitable to people who are clever, whose intentions may not be benign and very early on it became clear to people who could do the programming that this system was also capable of being used for other purposes, and so the internet proved to be very hospitable to what we now call malware, malicious software, that has grown organically, just as everything else that runs on the internet has grown organically, to the point where it's a major pest and sometimes maybe a threat to our safety and security.

Just to give one example – the internet was designed by researchers, often graduate students but certainly engineers and computer scientists. Most of whom knew one another, trusted one another and they worked co-operatively together, so when they were designing parts of the net they did so without being suspicious, and a classic case in point was that when they eventually realised that email, was going to be a very big use of this network, they designed

an email system which was very efficient at transferring email from one place to the next, but which was not suspicious about the origins of where it was coming from. The original email system that was devised for the internet, an email server would never demand authentication from a sender, it would never say, I'm not going to pass on your message unless I know who you are for sure, it didn't do that, that wasn't built in, the authentication was not built into the system, and of course that loophole is what was exploited by spammers, so spam became very rapidly a colossal problem because of that intrinsic design lacuna in the centre of the network. The hole in the network existed for a good reason, it existed because it was designed by people who trusted one another, but they were dealing in a world where people are not trustworthy. Spam has proliferated to the point where it's a real menace, and it takes up a lot of the bandwidth of the network. It's particularly obnoxious because it strikes at poor people more than rich people. If you are in countries in the world which have very limited internet bandwidth, internet bandwidth is very precious when you're poor especially, and you find then that 80 or 90% of that precious bandwidth is actually being choked by spam, it's horrible. We in the West can deal with it in different ways. We have all kinds of elaborate schemes and filters, and blocking systems, and the rest of it, so most of us probably don't see anything like the amount of spam that goes on. But that's an example of the way in which malicious software took advantage of the openness of the net to innovation, it happens to be a malicious, a very bad innovation but, nevertheless, the internet did it, you could do it with data packets so the internet does it for you, whether you're a spammer or whether you're Tim Berners-Lee.

Penny Boreham
Are there other sorts of vulnerabilities are we exposed to when we log on to the internet?

John Naughton
If somebody asked you is there a sure-fire way of keeping my computer safe from malicious software and from attack - there is a sure-fire way: never connect it to the internet, ever, that's the only sure-fire way of doing it. But basically if you connect it to the internet it is going to be vulnerable in some way, potentially vulnerable in some way, because of the growth of the malware industry and the intrinsic problems of connectivity. One of the big difficulties we have had is that most people rely on software, and software by definition has bugs in it, there's no such thing as bug-free software. Likewise most communication software will have somewhere in it security vulnerabilities which can be found and exploited by clever and ingenious people. That's true for every bit of communication software ever written, I think. Now the problem is that most of the companies, all of the companies which make software, and especially the companies that sell software in huge quantities, connectivity software, various kinds, anything that connects to the net and the rest of it, they have strangely been allowed to escape legal liability for the security deficiencies in their products, so that if a very large American company based in Washington State, say, which I will not name, if it produces software that runs on 90% of the world's computers, and that software has a security hole in it, and your business or your family is in some way damaged by that flaw, by some unscrupulous person exploiting that flaw, you have no legal recourse. If another very large American company, or indeed British company, sold you a car and it had a known deficiency in it which caused the brakes on rare occasions to fail and cause death or serious injury, and the rest of it, you would have recourse, you'd have legal recourse because the manufacturer would be legally liable.

One of the great mistakes of the last thirty years is the failure of governments to understand the potential damage that can be caused by inadequate or faulty, or vulnerable software. If the governments of the world had said from the beginning sure, you can sell as much software as you like, but you'll have to stand over it, and you're legally liable for its flaws, then we'd have had a very different business, and we'd have much less trouble from malware, and we'd have much less angst and grief. But in one of those extraordinary lapses of governmental regulatory attention, we never thought about it. If you wanted to clean up the industry quickly that would do wonders. Of course the industry would scream like hell because the software people would say well, that will slow up innovation, if you require us to be legally liable for something that some crazy cracker may do in due course, then our business is shot to hell, and the rest of it as well. The pharmaceutical industry might say the same thing, you require us to spend all this money testing drugs and the rest of it when we,

you know that's a curb on our innovation. Well we accept the curb on the innovation in the case of the drug companies and one day, I think, we will have to accept it in the case of software.

Penny Boreham
You mentioned concerns about 'crazy crackers'. What about hackers? Where do they fit into the story.

John Naughton
From the very beginning there have been groups of people with the requisite know-how, the deep technical ingenuity you need to understand this stuff, and actually a good heart, and these are people that said look, we have found this vulnerability, we're not going to exploit it but we want something done about it and they have, on occasions, tried to persuade these large software companies that they really ought to attend to these security violations, to these security vulnerabilities, and they've had an indifferent response. In some cases they have then gone public and revealed it. I've always thought that that's a very ethical and laudable thing to do, and I've found it very annoying and distressing that even today large software companies regard these people as pests. They're not; they're public citizens doing a good job, in my opinion.

Penny Boreham
John, you've been talking about at some surprises the internet has sprung over the last forty years. Where do you see it going in the next forty?

John Naughton
Because I wrote a history of the internet some people think that I have some special insight into where it's headed, and they say well what does it mean, and where's it heading, and the rest of it, the honest answer is that I have no more idea than anybody else. I say to them sometimes if you want to know the future go buy a crystal ball. If I knew what was going to happen I'd be rich and famous.

The first thing to start with is humility, that nobody knows. But there's a good reason for that. Naughton's first law of technology is that we always overestimate the short-term impact of communications technologies, and we always grievously underestimate their long-term implications. The truth is that nobody knows what the long-term implications of this stuff is. We've had a complete transformation in our communications environment, absolutely total, as revolutionary as print. We have no idea sitting here what that means in the long run.

To illustrate that let's take print. You could date the invention of printing by movable type back to the Gutenberg Bible in 1455. He published it in Mainz, a small town in Germany. Tim Berners-Lee released the Web in '91 so we're less than two decades into it. Now, just imagine this for a minute. Think yourself back to Medieval Germany. You're standing on a bridge in Mainz, twenty years after the invention of print, and you've got a clipboard and you're going to do one of these public surveys, and you stop people and you say, can I ask you some questions? This invention of printing, now I want you to tell me on a scale of 1-5, one being it's not likely, and five is very likely, that the invention of printing will undermine the authority of the Catholic Church, lead to the Reformation, enable the rise of modern science, change our conceptions of childhood, lead to the creation of completely new social classes – what do you think is most likely on a scale of 1-5 in these cases? They would have no idea. Printing transformed the world entirely. For the next four hundred years it's shaped the way our societies are, it transformed the entire world. We can see that now from the the comfortable hindsight of three hundred or four hundred years. We now can see the size of the revolution of the print book, we're trying to have some kind of perspective on the Web and on the internet, and the answer is that we have to think like that. Long-term we have no idea. We can be sure it is transforming our world already, but the transformation's only just started.