Journal of Cyber Policy

ISSN: 2373-8871 (Print) 2373-8898 (Online) Journal homepage: http://www.tandfonline.com/loi/rcyb20

The evolution of the Internet: from military experiment to General Purpose Technology

John Naughton

To cite this article: John Naughton (2016) The evolution of the Internet: from military experiment to General Purpose Technology, Journal of Cyber Policy, 1:1, 5-28, DOI: 10.1080/23738871.2016.1157619

To link to this article: http://dx.doi.org/10.1080/23738871.2016.1157619

Published online: 08 May 2016.

JOURNAL OF CYBER POLICY, 2016
VOL. 1, NO. 1, 5–28
http://dx.doi.org/10.1080/23738871.2016.1157619

Routledge
Taylor & Francis Group

The evolution of the Internet: from military experiment to General Purpose Technology

John Naughton[a,b]

[a]Centre for Research in the Arts, Social Sciences and Humanities (CRASSH), University of Cambridge, Cambridge, UK; [b]Faculty of Mathematics, Computing and Technology, The Open University, Milton Keynes, UK

ABSTRACT
The Internet is now over four decades old. A survey of its evolution from a military experiment conducted in the context of the Cold War to a General Purpose Technology illustrates the extent to which the network was shaped, not just by the intrinsic affordances of its underpinning technologies, but also by political, ideological, social, and economic factors.

ARTICLE HISTORY
Received 12 February 2016
Accepted 12 February 2016

KEYWORDS
Internet; history; General Purpose Technology

Introduction

The Internet that we use today – i.e. the network of computer networks based on the Transmission Control Protocol (TCP)/Internet Protocol (IP) suite of protocols (Postel 1981) – is now relatively old technology. Research on its design commenced in 1973 and the network became operational in January 1983. For the first two decades of its existence, it was the preserve of a technological, academic, and research elite. From the early 1990s, it began to percolate into mainstream society and is now (2016) widely regarded as a General Purpose Technology (GPT) without which modern society could not function. So in a relatively short period the technology went from being something regarded as exotic, to an apparently mundane utility, like mains electricity.

Since utilities tend to be taken for granted (until they break down) and are generally poorly understood (because people are uninterested in how they work) industrial society now finds itself in the strange position of being utterly dependent on a technological system that is both very disruptive and yet is poorly, if at all, understood.[1] From this, various consequences flow: industries, economies, communities – and indeed whole societies – experiencing a wave of 'creative destruction' (Schumpeter 1942, 82–85) unleashed by the resulting technological change, and struggling to adapt to a rapid, and possibly accelerating, pace of development; exposure to a range of new, and potentially dangerous, vulnerabilities; the rise of new enterprises, and indeed whole industries, which would have been unthinkable without digital technology; new kinds of crime, warfare, and espionage; and the challenges of devising regulatory institutions which are fit for purpose in the digital age.

CONTACT John Naughton ✉ jjn1@cam.ac.uk

Several factors make it difficult for citizens to appreciate the nature and significance of the Internet. One is the distortion imposed by the 'Whig interpretation'[2] of Internet history – the tendency to view its development with the 20/20 vision provided by hindsight. This provides a misleading impression of a linear progression from one great idea to the next, and obscures the paths of development that could have been, but were not, taken.

Yet there was nothing inevitable about the evolutionary path that the Internet has taken. Like all technologies, it has been shaped not just by critical technical decisions made at various stages in its history, but also by accident and by economic, social, and cultural forces.

Another factor occluding a clear understanding of the network is the short-termism of much public discourse about it, a trait that might be satirised as 'the sociology of the last five minutes'. This is largely a product of two factors: the capacity of the network to enable unanticipated innovations and launch them into society, and the obsession of the mass media with the 'New New Thing'.[3]

Other factors militating against a rounded appreciation of the Internet are definitional or conceptual. An example of the former is the widespread misapprehension that it is a unitary network rather than *a network of computer networks*. To most non-technical people, who experience the Internet as if it were a seamless whole, this will seem to be a technical detail, but as we shall see, the distinction is significant in terms of understanding the evolution of the system and in appreciating its capacity for disruption.

A final conceptual error is the widespread tendency to confuse the network with one or more of the applications that people use. Thus, for example, many users think that the World Wide Web *is* 'the Internet'. Others make the same mistake in respect of Facebook (see e.g. Moo 2016). But the Web and Facebook are just particular examples of data-enabled services that run on the infrastructure that constitutes the Internet, and mistaking them for the network is analogous to thinking that intercity trains, say, define the railway system.

Although, as observed earlier, the Internet is a relatively old technology, in another sense we may only be at the beginning of the 'Internet era', in the sense that we are only now reaching the point where the network is mature, widespread, and developed enough to be eligible for the status of a GPT. Bresnehan defines a GPT in terms of three characteristics:

(i) it is widely used;
(ii) it is capable of ongoing technical improvement; and
(iii) it enables innovation in application sectors (Bresnehan 2010, 764).

At the time of writing, over 3 billion people (40% of the global population) have access to the Internet[4] and predictions are that, largely thanks to smartphone adoption, half of the global population will be online by 2017 (Broadband Commission 2014, 12). The network is therefore not yet ubiquitous in global terms, although in most industrial countries it effectively is. As far as criterion (ii) is concerned, it has been steadily and continuously improved and extended. As far as criterion (iii) goes, the network definitely qualifies, partly because a capacity for disruptive innovation is 'baked into' its architecture, and while the network has to date clearly enabled innovation, until recently this has tended to be in predictable application sectors (e.g. information goods and services). An important question for the future, therefore, is the extent to which it will have significant impacts in application areas that are far removed from information technology.

History

As with most technologies (Arthur 2011), the roots of the Internet go back a long way, mostly to the post-World War II era, but in some respects to the late 1930s.[5] The evolution of the network to date can be summarised in terms of two main phases: its development from a military experiment to a civilian utility, and the commercialisation of the network.

Phase one: from military experiment to civilian utility (1967–1995)

Pre-history: 1956–1966

That the Internet owes its existence to the Cold War is well known. But, as ever, retrospective generalisation glosses over a more complicated story.

The first strand of this concerns the doctrine of 'mutual assured destruction' (MAD) which governed the nuclear stand-off between the United States and the Soviet Union. MAD supposedly ensured national security by guaranteeing that if one side launched a nuclear attack, the other would retaliate in (devastating) kind. There was, however, one apparent flaw in the logic, in that the doctrine could give an advantage to the aggressor if his pre-emptive strike was so devastating that it rendered the enemy's command-and-control system inoperative, thereby making it impossible to retaliate.

There was therefore an urgent need to design a communications system capable of surviving a devastating thermonuclear attack. This challenge was taken up by a researcher in the RAND Corporation, Paul Baran, who came up with a design for a mesh network based on high levels of link redundancy and a digital communications technology called packet switching (Hafner and Lyon 1996, 51).[6] At a time when communications networks were almost exclusively analogue and based on circuit switching, these were radical ideas. Baran worked out a detailed design for a network based on these principles, but for various bureaucratic and other reasons, a prototype network was never built, and the trail apparently went cold.

It transpired, however, that Baran was not the only person to come up with the packet switching idea. In the UK's National Physical Laboratory, Donald Davies was seeking to create a new kind of communications network for purely civilian applications. He sought to exploit the advantages of digital switches to enable responsive, interactive time-shared computing over large distances and realised that the circuit switching technology of the analogue telephone system was not adequate for this task (Hafner and Lyon 1996, 66). So Davies independently conceived the idea of packet switching as a means of achieving his goal.

The second strand of the story opens with the Soviet Union's successful launch of the Sputnik satellite in October 1957, an event that profoundly shocked the US defence establishment and led to the setting up of the Advanced Research Projects Agency (ARPA) within the Department of Defense. Early in its organisational life, ARPA morphed into the agency within the Pentagon that funded advanced, 'blue-skies' research which could have military applications. In due course, ARPA found itself funding the purchase, operation, and maintenance of at least a dozen (Abbate 1999, 45) expensive mainframe computers for the various university departments and institutes which held research contracts from the agency. The problem was that these machines were incompatible with one another, and therefore could not function as shared resources for the community of

ARPA-funded researchers across the US (Hafner and Lyon 1996, 41). From this came the idea, and the funding, for a network that would enable these valuable resources to be shared. ARPANET (Advanced Research Projects Agency Network) was the result.

The ARPANET: 1967–1972

The idea of a 'resource-sharing' network first emerged in ARPA in 1966 (Naughton 1999, 84). Design work, conducted in a collegial style (Abbate 1999, 56) unusual in the defence industry, then proceeded over the next two years. The contract to build the network was awarded in early 1969 to Bolt, Beranek and Newman, a Boston-based consultancy firm with strong links to MIT.

The technological and conceptual challenges that faced the network's designers have long been obliterated by the omniscience of hindsight, but they were formidable.[7] Given that the network was supposed to facilitate the sharing of expensive and scarce resources, namely the mainframe computers that ARPA had funded in various research centres across the country, a key obstacle to overcome was the fact that these machines were incompatible with one another. For each of them to participate in a network would require the creation of complex, customised networking software which would enable each machine to communicate with every other machine on the system. In the end, this problem was not so much solved as side-stepped: it was decided to build a 'subnet' of identical minicomputers (which came to be called 'interface message processors' or IMPs) each linked to a single mainframe 'host'. In that way the task of writing networking software for a host was greatly reduced: it would simply have to communicate with a single machine – the IMP assigned to it.

Given the technical challenges implicit in the task, the ARPANET was built with astonishing speed. By 1972, the network was essentially complete; the 15 original sites were all connected and operational and a major public demonstration of the system was held in Washington, DC in the Autumn of that year (Hafner and Lyon 1996, 176–186).

From the perspective of the present, three aspects of the ARPANET project stand out.

The first is that while it was a triumph of project management in the conventional sense, success was achieved in an unusually collegial way. This was sensible for several reasons: the network was intended to link high-profile researchers working in elite institutions, and such individuals are not easily herded, plus it made sense to harness the collective IQ of that community at every level, including that of graduate students. For that reason, for example, design of the central protocols of the network was entrusted to a Network Working Group that largely consisted of students. In this way was established the collaborative ethos that has been an important feature of Internet technical development up to the present day.[8]

Secondly, the ARPANET provides an interesting case study in the extent to which technologies are socially shaped. In this case, the shaping was done by the network's users, first of all because many of them were actively involved in the design of the network and therefore they were both designers and 'customers', that is, users. And in the latter capacity, they sprang some major surprises on ARPA managers. The network was intended to be a facility for resource-sharing, but it transpired that it was not much used for this original purpose. Instead, its users employed it mostly for communicating with one another, sharing files and software, and for sending and receiving email (Abbate 1999, 108). In that sense, the community of users came up with a new conception of what 'networking' meant – not so much the sharing of *machines* as the linking of *people*. As Abbate puts it, 'Increasingly people within and outside the ARPA community would come to see the ARPANET not as a computing system but rather as a communications system' (Abbate 1999, 111). Given the technical sophistication of the network's users, it is also not surprising that they were vocal in their demands for system modification and innovation. But while many user tweaks were 'encouraged or at least tolerated' by ARPA, the agency did not always welcome users' attempts to steer the evolution of the system (Abbate 1999, 93). Its reluctance in this respect may not have been due to hierarchical reflexes so much as the need to reassure Congress that the network was not a publicly funded experiment in computer science, but an administrative tool useful for military and defence purposes.

Thirdly, although the ARPANET was based on the packet switching technology that characterises the modern Internet, it was a *unitary* network: the subnet of identical, centrally managed IMPs constituted the core of the system. And the network was owned and administered by a single entity – ARPA. So although the ARPANET was the precursor of what came later, it differed in significant ways from its successors.

Development of the TCP/IP-based 'internetwork': 1973–1983

During and after the construction of the ARPANET, other significant developments in networking technology were under way. At the University of Hawaii, researchers had built ALOHA – a packet-switched network that operated, not over leased telephone lines, as ARPANET did, but via radio. Within ARPA, it was decided to build on this work by creating a packet-switched radio network (named PRNET) in the San Francisco area. The motivation for this was obvious: ARPA was part of the US Department of Defense and its planners were interested in the potential of packet switching for command-and-control in battlefield conditions. The agency had also begun to experiment with using the technology in satellite communications, for example, for linking seismic monitoring stations in Scandinavia (established to monitor Soviet nuclear testing) with the US, via a network which was christened SATNET.

By the early to mid-1970s, therefore, ARPA found itself running three separate 'experimental' networks – ARPANET, PRNET, and SATNET – all of which used packet switching technology, but in different ways. An obvious next step was to see whether a method for 'internetworking' them, so that they functioned as an apparently seamless whole, could be developed.

The 'internetworking' project began in late 1973. A key challenge for the designers of the new system was to find a way of transitioning from a unitary network like ARPANET to something that could incorporate a variety of different networks that were owned and operated by independent organisations and entities. From a technical point of view, there were various ways of achieving this goal. One was to allow networks wishing to join the new 'internetwork' to retain their existing protocols and simply construct 'gateway' computers that would translate those into a common set of conventions. The other was to require that all candidate networks adopted a new set of protocols, which would become the *lingua franca* of the new overarching network (Abbate 1999, 128).

In the end, the second option was adopted, and a suite of interlocking protocols centred on two new ones – TCP and IP – evolved. In this way TCP/IP became the cornerstone of the new 'network of networks'. The great advantage of this approach was that implicit in it was the possibility of organic growth: as long as a given network 'spoke' TCP/IP (as it were) it was free to join the Internet. And because the system was not owned or controlled by anybody (unlike the ARPANET), there were no gatekeepers to control admission to it.

But the potential for organic growth was not the only affordance implicit in the TCP/IP architecture. The designers also faced the puzzle of how to create a network that would be as future-proof as possible, that is, one that could cope with applications that had not been anticipated by the designers. Their solution was to design a system that was not optimised for any particular application (in contrast to, say, the analogue telephone network, which had been optimised for voice calls but proved inadequate for computer-to-computer communication). The Internet, concluded its designers, should do only one thing: it should take in data packets at one end and do its best to deliver them to their destination. It would be entirely agnostic about the contents or purpose of the packets. In this way, all of the ingenuity would be left to users of the network. If someone had an idea that could be realised using data packets, then the Internet would do it for them with no questions asked.[9]

This philosophy – of leaving innovation to the edges of the network – had profound implications. As Van Schewick (2012) describes it, the TCP/IP design created an architecture for 'permissionless innovation' which enabled the explosion of disruptive creativity that is perhaps the most distinctive feature of the Internet.

As the new network took shape, the ARPANET continued to operate alongside it. But because it was no longer an experimental facility, ARPA began to look for a new owner for it. It was first offered to AT&T, the national regulated telephone monopoly, but the company turned it down (Abbate 1999, 135). In 1975, the agency transferred operational responsibility for the network to the Defense Communications Agency (DCA) which provided communications services for the US armed forces. This had predictable effects: it began the process of reorienting the network away from its original research focus and towards military applications; it inevitably led to tensions between the military mindset and the more freewheeling ethos of the research community; and it also led to heightened concerns about computer security and access controls.

While this was happening, the Internet project was gathering pace and the TCP/IP family of protocols was being finalised. The imperative was to extend adoption of the new protocols to the point where network effects came into operation. This turned out to be more difficult than anticipated: many nodes were reluctant to devote the necessary resources to configure their operations around the new protocols. It was at this point that DCA management of the ARPANET was to prove decisive. In March 1981, the Pentagon announced that all ARPANET hosts would be required to adopt TCP/IP by January 1983. Not all sites were able to meet the deadline, but by the middle of 1983, every ARPANET host was running TCP/IP (Abbate 1999, 142), which is why we can say that 1983 marks the beginning of the Internet that we use today.

A few months before that, however, DCA concern about the security of the network had led to a decision to split it into civilian and military domains. From October 1982, one domain – the ARPANET – would continue as a research enterprise; the other – labelled MILNET – would henceforth be entirely devoted to military communications. The switchover was implemented in April 1983.

Transition from a military/research network to a 'civilian' one: 1983–1995

The creation of the MILNET domain meant that ARPANET returned to being a research-focused network dominated by universities and research institutions, so the breach was an essential first step towards achieving ARPA's goal of transferring the network to civilian control. The second step was to take measures to foster the dissemination of TCP/IP technology within the computer industry. To that end, ARPA funded various operators to create TCP implementations for various operating systems (notably Unix[10]) and launched a $20m fund to help computer manufacturers implement TCP/IP software on their machines (Abbate 1999, 143). So by 1990, TCP/IP was available for most computers, at least in the US market.

Until the end of the 1970s, access to the developing Internet was restricted to those working in a relatively small number of institutions which held research contracts from ARPA. As computer science became an accepted academic discipline in universities, the exclusiveness of the ARPANET/Internet club was increasingly perceived as irritating and dysfunctional. This led the US National Science Foundation (NSF) to fund the creation of the Computer Science Network (CSNET) in the early 1980s. After an initial hiccup, it was decided CSNET would use the TCP/IP protocols, which meant that a connection between CSNET and ARPANET was feasible and so, at a stroke, the community of networked computer scientists was significantly expanded. While access to ARPANET was only granted to researchers funded by the agency, membership of CSNET was open to computer scientists in any institution willing to pay the annual subscription (although *commercial use* of the network was prohibited under the NSF's 'acceptable use' policy). The result was that the network began to grow at a faster rate – from 2000 host computers in 1985, to 185,000 in October 1989, and 1,776,000 in July 1993.[11]

CSNET turned out to be just the first step in the NSF's involvement in networking. In mid-1984, the foundation began funding the establishment of several new supercomputing centres around the US. To make these available to the widest possible community of researchers, a national network was required. The original idea was for a network – NFSNET – linking the centres that would form the 'backbone' of a wider academic network, but in 1998, an agreement was reached to use the ARPANET as the (temporary) backbone of the new network while it was being built. With this decision the Internet became 'a civilian network in all but name' (Abbate 1999, 194). It transpired, however, that the ageing ARPANET proved inadequate as a backbone for a rapidly expanding national network, and so in the end a swap was done, with NFSNET becoming the backbone for the ARPANET.

On 28 February 1990, the ARPANET was officially decommissioned; the era of formal military involvement in the operation of the Internet had ended. That still left open the question of how the network should be funded in the coming decades. Should it remain a publicly funded operation, with the NSF using taxpayers' dollars to pay the costs? Among the arguments against this was that it implied a continuation of the rule that the network could not be used for commercial purposes – which would preclude the exploitation of the economic potential of the technology. In the end, the NSF decided that the only way to allow commercial use of the Net would be to privatise it – to take it off the government's books.

And this is in fact what happened. In 1994, the NSF implemented a plan to allow Internet service to be taken over by commercial companies known as 'Internet Service Providers' (ISPs), each of which would operate its own backbone, enabling the old NSF backbone to be decommissioned. Customers would connect to one of the companies' backbones, and the ISPs would operate a set of gateways at which a number of ISPs could interconnect their systems, allowing traffic to pass smoothly from one network to another, giving end users the illusion of interacting with a seamless, unitary system. What this also implied, though, was that the network was 'open for business'.

The process by which the network was privatised was a critical determinant of how the Internet evolved, and it has been largely obscured by the Whig interpretation of the network's history. But there was nothing preordained about the transition. The eventual outcome of the NSF's handover of the network to private interests was a product of judgement, foresight, consultation, political astuteness, luck, and timing.

A full account of the process lies beyond the scope of this paper,[12] so a single illustration must suffice. One of the key questions facing the NSF administrators was: what kinds of commercial entities should run the network's backbone? A decade or so earlier, the answer would almost certainly have been AT&T, the regulated telephone monopoly. In 1991–1992, when the actual decision had to be made, an equally plausible candidate might have been IBM (and indeed IBM saw itself precisely in that frame). The attractions for replacing a unitary public service provider (NSF) with a single commercial provider, were obvious in terms of continuity, stability, and order. But giving a single organisation such a degree of control over the network would violate its fundamental design axioms. It was therefore important to ensure that a number of competing ISPs should run the backbone. But that raised the question of how these organisations would co-operate in such a way that the illusion of a seamless network would be maintained. The ISPs proposed a solution: the Commercial Internet Exchange – essentially a router that took traffic from a number of competing ISPs who had agreed to pass data traffic between their networks (Greenstein 2015, 80). Since each exchange was funded by a flat-rate subscription from its members, data exchange between them was effectively free, rather than being metered by volume. This 'peering' system[13] became the predominant model for interconnection and evolved to become a central component of the network's infrastructure. But in order to make it feasible in 1992, a small yet significant amendment to the NSF's Charter had to be made – and this was achieved by a Bill, sponsored by a Virginia Congressman, Rick Boucher. The point of this story is simple: there is nothing preordained about technological development.

Phase two: the commercial Internet (1995–present)

The first Internet boom: 1995–2000

In the 1980s, 'cyberspace' – the term coined by the novelist William Gibson to describe the virtual world behind the computer screen (Gibson 1984) – was an unusual space. It was essentially a geek preserve, with a social ethos that was communal, libertarian, collaborative, occasionally raucous, anti-establishment and rich in debate and discussion (see Hauben and Hauben 1997). It had no commerce, no hierarchies, no crime and no spam, and in general it was populated by people who either knew one another, or at least knew others' institutional affiliations. In that sense, cyberspace and the so-called real world existed as parallel universes. Most people outside of the magic circle had no knowledge of the network – and even if they did, they would have found it difficult to gain admission to it.

Two developments eroded the distinction between the two universes and caused them gradually to merge. The first was the commercialisation of the network achieved by the NSF's decision to hand the backbone over to ISPs. This meant that lay people could now access the network – generally via slow dial-up connections. The second development was the arrival of the network's second 'killer application'[14] – the World Wide Web.

The Web was the creation of a single individual – the physicist and computer scientist Tim Berners-Lee, who was employed in the late 1980s and early 1990s at CERN, the multinational particle-research laboratory located just outside Geneva in Switzerland. The underlying idea was to develop a way of publishing, locating, and retrieving documents stored on Internet servers across the world, something that would be useful for a large international laboratory like CERN, which had large numbers of visiting physicists and a perennial problem with document control. Berners-Lee's idea was to take an established technology called 'hypertext' – software which created documents with extensive cross-referencing between related sections of text and associated graphics (Naughton 1999, 220) – and make it work across the Internet. In a remarkable burst of creativity at the end of 1990, Berners-Lee created a working prototype of what he dubbed the 'WorldWideWeb', in three months (Berners-Lee 2000, 30–32).

The prototype, however, did not generate much excitement at CERN, save among a small group of enthusiasts led by Berners-Lee's colleague, Robert Cailliau (Berners-Lee 2000, 48; Gillies and Cailliau 2000, 201). Most people did not see the potential of the Web at the beginning. With hindsight, this seems surprising, but in fact it is perfectly plausible when one remembers the context in which the technology first appeared. As Gillies and Cailliau put it:

> CERN's management was wary of committing resources to projects outside the laboratory's core area of physics research. The new LHC accelerator was not yet approved by the laboratory's member states, the Americans were planning to build a similar machine, and management's main priority was convincing member state governments that Europe really did need another million-dollar atom smasher. They were keen to be seen to be spending the taxpayers' money wisely, and pumping resources into what they saw as no more than an interesting curiosity did not look like the way to do it. (2000, 201)

From virtually the outset of the project, Berners-Lee had been posting documentation on the project on CERN's Internet server (info.cern.ch). In March 1991, he released the code to a selected number of CERN colleagues who were using NeXT workstations (Berners-Lee 2000, 50). In August, having obtained agreement from CERN management that the laboratory had no interest in retaining the IP rights to the software, he uploaded to info.cern.ch the program he had written to run the Web on NeXT machines, plus the code for a simple line-mode (i.e. text) browser and for a generic web server, and posted a note drawing attention to this in alt.hypertext, the Internet discussion group for researchers working on hypertext. Shortly afterwards, he opened a public Telnet server[15] on info.cern.ch which would enable anyone to dial-in and use the browser installed on the CERN machine. This had the advantage of allowing people who were not in a position to install the Web software on their own machines to experiment with it.

Despite all this, dissemination of the Web in 1991–1992 was slow and remained so until the spring of 1993, when Marc Andreessen and Eric Bina, then working at the National Center for Supercomputer Applications (NCSA) at the University of Illinois at Urbana-Champaign, released *Mosaic*, a browser they had written for the Web. Although *Mosaic* was not the first graphical browser, it was the first one to display graphics inline (i.e. as an integral component of a page, rather than in a separate window).

The launch of *Mosaic* was a landmark moment in the evolution of both the Web and the Internet. It provided a dramatic illustration of the Web's potential for both publication and commerce. It opened up the possibility that the Web could be an entertainment medium. And because one needed access to the Internet in order to use the Web, *Mosaic* triggered a sudden surge in the demand for Internet connections among the general public. One source claims that in 1993, traffic on the nascent World Wide Web increased by over 300,000% (Cassidy 2002, 57).

Andreessen left Illinois in 1994 to join with the technology entrepreneur Jim Clark in order to set up Netscape, the first company founded expressly to exploit the Web commercially. The company's first product was a browser, *Netscape Navigator*, which although resembling *Mosaic* in many respects, had been coded from scratch to avoid intellectual property disputes with the NCSA. *Navigator* rapidly became the dominant browser in the nascent market, partly because versions of it were available for a range of computer systems, including the IBM PC.

On 9 August 1995, the Netscape corporation filed for an Initial Public Offering (IPO). The stock was priced at $28 a share. In the first day of trading, the price peaked at $75 before eventually closing at just over $58, a price that valued the year-old company at $2.9 billion.

Netscape's extraordinary IPO triggered intense speculative interest in the Internet and paved the way for the first Internet boom, an extraordinary outbreak of what the economist Robert Schiller dubbed 'irrational exuberance' (Schiller 2000), and which was later christened the 'dot-com bubble'.[16] What fuelled the mania was speculative interest in the stock market valuation of a multitude of Web-based companies ('dot-coms') which materialised following Netscape's extraordinary debut, and which was amplified by the activities of fund managers, stock analysts, journalists, and pundits. As one sceptical observer put it, what really happened was that 'Wall Street moved West' (Keen 2015).

The core business model of these fledgling companies was the idea of harnessing the network effects implicit in the rapid growth of consumer interest in the Internet, in order to obtain a dominant market share in a range of sectors.[17] At the height of the frenzy, dot-com companies with few customers, little (sometimes no) revenues and few employees, briefly enjoyed stock market valuations greater than those of long-established companies with significant revenues and thousands of employees.

The boom followed the traditional pattern of speculative manias through the centuries and, eventually, in March 2000, the bubble burst. In just over a month, the total market capitalisation of companies on the NASDAQ exchange fell from $6.71 trillion to $5.78 trillion (Geier 2015). In other words, nearly a trillion dollars in value had been obliterated. And fewer than half of the dot-coms founded in the boom survived the crash (Goldfarb, Kirsch, and Miller 2007).

In the disillusionment that followed, many people drew the conclusion that the Internet phenomenon was overblown. This was understandable, but misconceived. If the first Internet bubble showed anything (other than the perennial gullibility of humans) it was that expectations can sometimes run far ahead of technological, and economic, reality. The truth is that in the last decade of the twentieth century, the Internet was still not a technology mature enough to bear the weight of the (feverish) expectations of dot-com investors.

At that time, for example, only an estimated 413m people – 6.7% of the global population – had access to the Net.[18] Most of them connected to the network via slow, noisy, and low-bandwidth dial-up lines. There was virtually no wireless networking outside of research labs, and very few people had persistent IP addresses. Most mobile phones did not connect to the network, and the few that did, connected via very limited WAP connections. There was very little 'cloud' computing. Given these realities, the Internet of the late 1990s was very much an immature technology.

The most paradoxical aspect of the first Internet boom however, is that the bubble created much of the technological infrastructure necessary to hasten the maturation of the network. When the bubble began to inflate, some canny observers quoted the old maxim of the Californian Gold Rush of the 1850s: that the people who made most money in California were not the miners or those who panned for gold, but the merchants who sold them pickaxes and shovels. The modern embodiments of those merchants were the telecommunications companies which, in the 1990s, invested heavily in building large fibre-optic cable networks and server farms to service the 'new' economy that was apparently coming into being. When the bubble burst, these companies were left with apparently unwanted assets, and some went bankrupt. But the infrastructure that they had built remained, and indeed turned out to be critical in enabling what came next. As the economist J. Bradford DeLong put it, 'Investors lost their money. We now get to use all their stuff. What got built wasn't profitable, but a large chunk of it will be very useful' (DeLong 2003).

And this is an old story. DeLong points out, for example, that the 'railway mania' of the nineteenth century lost investors a lot of money, but the extent of the railway network that was the product of the frenzy enabled completely new industries to be built.

> Americans and the American economy benefited enormously from the resulting network of railroad tracks that stretched from sea to shining sea. For a curious thing happened as railroad bankruptcies and price wars put steady downward pressure on shipping prices and slashed rail freight and passenger rates across the country: New industries sprang up.
>
> Consider, for example, the old Montgomery Ward and Sears Roebuck catalogs. Sears and Montgomery Ward discovered at the end of the 19th century that the cost of shipping consumer goods to rural America was no longer a competitive burden.
>
> Mail a catalog to every household in the country. Offer them big-city goods at near big-city discounts. Rake in the money from satisfied customers. For two generations this business model – call it the 'railroad services' business model – was a license to print money, made possible only by the gross overbuilding of railroads, the resulting collapse of freight rates, and the fact that railroad investors had had to kiss nearly all their money good-bye. Their pain was outweighed by the gain to American consumers and manufacturers, who could now order and ship goods essentially free. The irrational exuberance of the late 1800s made the railroads a money-losing industry – and a wealth-creating industry. The more money investors lost through overbuilding, the lower freight rates became, and the more railroads belched out wealth for everybody else.[19]

Profits, in other words, are not the same thing as social value. And so it proved with the Internet.

'Web 2.0': 2000–2003

The Web was originally conceived as a means of sharing information among particle physicists who were scattered across the world. Since most of that information was in the form of documents, the design was therefore for a system that would make it possible to format these documents in a standardised way, publish them online, and make them easy to access. So the first 'release' of the Web (to use a software term) created a worldwide repository of linked, static documents held on servers distributed across the Internet.

Given that it was intended as a system for academic researchers, the original Web design was probably fit for purpose in its first two years. But once the *Mosaic* browser appeared in 1993 and the commercial possibilities of the technology became obvious to the corporate world, the limitations of the original concept began to grate. The early Web did not make provisions for images, for example. And it was a one-way, read-only medium with no mechanism for enabling people to interact with web pages, which meant that it was unsuitable for e-commerce. There was no way for users to talk back to authors or publishers; no way to change or personalise web pages; no way to find other readers of the same page; and no way to share or collaborate over the Web.

From 1993 onwards therefore, there was a steady accretion of innovative technologies designed to extend Berners-Lee's creation and to overcome some of its perceived limitations. The main driver behind this was e-commerce, which desperately needed to transform the Web into a medium that facilitated transactions.

In order to make transactions possible, a whole range of problems had to be solved. For example, ways had to be found to allow interactivity between browsers and servers; to facilitate personalisation of web content; and to overcome the problem that the http protocol was both insecure (in that communications between browser and server could be intercepted and monitored by third parties) and stateless (i.e. unable to support multi-step transactions).

In time, solutions to these problems emerged in the forms of: 'cookies'; HTTPS (an encrypted version of the basic http protocol); the evolution of browsers with capabilities added by specialised 'plug-ins' which enabled them to handle audio and video and other kinds of file; and, eventually, JavaScript, which effectively turned web pages into small virtual machines. Many of these technologies had an ad hoc feel to them, which was hardly surprising, given that they had been grafted onto a system rather than being designed into it. But they nevertheless proved extraordinarily powerful in supporting the dramatic expansion of the Web from 1995 onwards.

In pondering the Web 1.0 enterprises that had survived the crash, and the new ones that had arisen afterwards, it became clear that they had several important features in common, an observation which eventually led to them being dubbed 'Web 2.0' by one prominent observer of the technology (O'Reilly 2005). One of these features was that they harnessed the collective intelligence available on the Web, either via software such as Google's PageRank algorithm (which ranks web pages using a kind of automated peer-review) or by exploiting the willingness of users to engage with the enterprise (as, for example, in Amazon's utilisation of product reviews by customers). Another example of collective intelligence at work was Wikipedia – an enterprise made possible by Ward Cunningham's invention of the 'wiki' – a web page that could be edited by anyone who read it (see Naughton 2012, 95). Cunningham's software transformed the Web from a one-way, read-only medium, into what Tim Berners-Lee later called the 'read–write Web'.

A second distinguishing feature of the 'new' Web was 'user-generated content' or 'peer production' – that is, material created and published freely by people who do it for no apparent economic motive (Benkler 2007).

Another distinctive feature of the 'new' Web was that many of the emerging services on it were dynamically interconnected by means of software tools like the syndication tool RSS and Application Programming Interfaces (APIs). The latter provide the 'hooks' on which other pieces of software can hang. What was distinctive about some of the web services that evolved after 1999 was that they used APIs to specify how entire web services could work together. A typical example is the API published by Google for its Maps service. This made it possible for people to create other services – called 'mashups' – which linked Google Maps with other Internet-accessible data sources.

Fourthly, many of the new Web services were distinctive by never being 'finished' – by being in what programmers would call a 'perpetual Beta' stage. This intrinsic, experimental ethos of the emerging Web was exemplified by the Google search engine which, when it launched, and for a considerable time afterwards, carried the subscript 'BETA'. What was significant about this was that it signalled its designers' philosophy of regarding their web-based service as a work in progress – subject to continual and sometimes rapid change – rather than as something fixed and immutable. What made this possible of course, was the fact that it was a cloud-based service, so every user's version of the software could be upgraded at a stroke, and without any effort on their part, beyond occasionally upgrading their browser software or installing some (free) plug-ins designed to take advantage of whatever new features Google had decided to add.

A final distinguishing characteristic of the post-1999 Web was that the enterprises and services that were becoming dominant were effectively using the Web as a programming *platform*. So while the Internet was the platform on which Web 1.0 was built, Web 1.0 in turn became the platform on which the iconic services of Web 2.0 were constructed. This was made possible firstly by the fact that the Web provided a common standard, and secondly by the fact that if a service was provided via the http protocol, it could bypass the firewalls used by organisations to prevent unauthorised intrusions (since most firewalls were programmed to allow 'web pages' to pass through).

Mobile connectivity, surveillance, cybercrime, corporate power, changing patterns of use and their implications: (2004–present)

The most recent phase in the evolution of the Internet has been characterised by significant changes in the ways that people access and use the network and by the ways in which the infrastructure of the network has evolved to cope with these changes.

A comprehensive survey of these developments is beyond the scope of this paper, so an outline of some of the more significant will have to suffice. Of these, the most prominent are: the rise of mobile connectivity; the rapid expansion of so-called social media; pervasive surveillance by both state and commercial entities; increase of the power and influence wielded by a small number of large technology companies and consolidation of their grip on the network; increases in cybercrime; the possibility of 'Balkanisation' of the network; changes in patterns of media consumption; and the emergence of new intermediaries like Uber, Airbnb, and Coursera which use the network as a platform on which to run businesses that are potentially very disruptive to incumbents.

Mobile connectivity

In many respects, the most significant moment in the recent history of the Internet was the arrival of the 'smartphone' – i.e. a mobile phone that can access the Internet – in 2007.[20] Adoption of smartphones (and related mobile devices, like tablet computers) has increased rapidly, to the point where it is clear that most of the next few billion Internet users, mostly from developing countries, will access the network via a smartphone (Schmidt and Cohen 2008). The implications of this development are profound.[21] On the one hand, access to the network – and all the good things that could flow from that – will come within the reach of communities that have hitherto found themselves on the wrong side of the 'digital divide'. On the other hand, ubiquitous mobile connectivity will increase further the power and influence of corporations over Internet users because of (i) the latter's dependence on companies for both connectivity and content, and (ii) mobile devices' dependence on cloud computing resources for much of their functionality.

Social media

Online social networking services[22] have quite a venerable pedigree in Internet terms (see Naughton 2012, 98–101), but in the last few years the market has been dominated by Facebook (founded in 2004), LinkedIn (2003) and Twitter (2006). Of these, Facebook is by far the most dominant. As of Autumn 2015, it had 1.55 billion 'monthly active users', 90% of whom access the service from mobile devices (Abutaleb and Maan 2015). Given that Facebook was the brainchild of a single individual, a Harvard sophomore, its current prominence is an impressive demonstration of the capacity of the Internet to enable 'permissionless innovation'.

Pervasive surveillance

'Surveillance is the business model of the Internet. We build systems that spy on people in exchange for services. Corporations call it marketing.'[23] This statement from a noted computer security expert is a hyperbolic way of encapsulating the symbiotic relationship between Internet users and companies. On the one hand, users clearly value online services like search and social networking, but they have traditionally been reluctant to pay for them; on the other hand, Internet companies wanted to 'get big fast' in order to harness network effects, and the quickest way to do that was to offer services for free. The business model that emerged from this symbiotic relationship is advertising-based: users agree that the service providers may gather data about them based on their online behaviour and use the resulting knowledge to target advertising at them, hence the trope that 'if the service is free, then you are the product'.

Up to now, this surveillance-based model has worked well for the Googles and Facebooks of the online world. But its long-term sustainability is not assured; there are signs, for example, that users are becoming resistant to targeted advertising, and use of ad-blocking software is on the rise (Naughton 2015).

The last 15 years have also seen massive expansion in state surveillance of Internet and mobile communications, stimulated in large part by the 'state of exception' (Agamben 2005) necessitated by the so-called war on terror. There was probably a vague awareness among the general public that security and intelligence services were monitoring people's communications, but it took the revelations by the former National Security Agency (NSA) contractor, Edward Snowden, in 2013, to demonstrate the scale and intrusiveness of this surveillance.

Snowden's revelations have provoked much controversy, prompted a number of official inquiries (notably in the US and the UK) and the publication, in the UK, of a draft new Investigatory Powers Bill which is scheduled to become law before the end of 2016. At the time of writing, the Bill is on its passage through Parliament, but it seems unlikely that current surveillance practices will be abandoned, though oversight arrangements may change. And although public attitudes to covert surveillance seem to be culturally dependent, at least as measured by opinion polling, all the indications are that extensive surveillance of communications has become a fixture in liberal democracies, with unpredictable long-term consequences for privacy, human rights, and civil liberties.

Corporate power

Two aspects of 'power' are important in a networked world. One is the coercive, surveillance, and other power exercised by states. The other is that wielded by the handful of large digital corporations that has come to dominate the Internet over the last two decades. This raises a number of interrelated questions. What exactly is the nature of digital corporations' power? How does it differ from the kinds of power wielded by large, non-digital companies? In what ways is it – or might it be – problematic? And are the legislative tools possessed by states for the regulation of corporate power, fit for purpose in a digital era?

The five companies – Apple, Google, Facebook, Yahoo, Amazon, and Microsoft – have acquired significant power and influence and play important roles in the everyday lives of billions of people. In three of these cases – Apple, Amazon, and Microsoft – the power they wield mostly takes a familiar form: market dominance in relatively conventional environments, those of retail commerce and computer software and/or hardware respectively. In that sense, their market dominance seems relatively unproblematic, at least in conceptual terms: all operate in well-understood market environments and in one case (Microsoft) antitrust legislation has been brought to bear on the company by both US and European regulators. So although the market power of the trio raises interesting legal and other questions, it does not appear to be conceptually challenging.

The same cannot be said, however, of the power wielded by 'pure' Internet companies like Google, Facebook (and to a lesser extent, Yahoo). Their power seems just as significant but is harder to conceptualise.

Take Google, for example. Between January 2012 and January 2015, its global market share never dropped below 87.72% (the lowest point, reached in October 2013). In Europe, its share is even higher: around 93%. The global market share of its nearest rival, Microsoft's Bing, in January 2015, was 4.53%.[24]

This raises several questions. The first is whether such dominance results in – or might lead to – abuses of corporate power in ways that have become familiar since the 1890s in the United States, and for which legal remedies exist, at least in theory.

But there is another aspect of Google's power that raises a more puzzling question. It is posed by a ruling of the European Court of Justice in May 2014 in the so-called right to be forgotten case. The essence of the matter is that individuals within the European Union now have a legal right to petition Google to remove from its search results, links to online references to them that are in some way damaging or inaccurate. Such online references are not published by Google itself, and even if Google accedes to the requests, the offending references continue to be available online, so in that sense the phrase 'right to be forgotten' is misleading. All that happens is that they disappear from Google searches for the complained-of information. It would perhaps be more accurate, therefore, to describe this as *the right not to be found by Google searches*.

One could say, therefore, that Google has the power to render people or organisations invisible – to 'disappear' them, as it were. This effect may not be intentional, but it is nevertheless real. And the capacity to make that happen through what one might call 'algorithmic airbrushing' could be seen as analogous to a power which was hitherto the prerogative of dictatorships: to airbrush opponents from the public record. So this capacity to render people 'invisible' is clearly a kind of power. But what *kind* of power is it? Are there analytical or theoretical tools that would enable us to assess and measure it?

This is just one example of the uncharted territory that societies are now trying to navigate. Similar questions can be asked about Facebook's documented power to affect its users' moods (Kramer, Guillory, and Hancock 2014) and to influence their voting behaviour (Bond et al. 2012).

Cybercrime

The term 'cybercrime' covers a multitude of online misdeeds, from sophisticated attacks on government and corporate websites, to spam emails offering fake prizes. Its rise seems correlated – at least in countries like the UK – with a fall in reported offline crime. This might be a coincidence, but a more plausible hypothesis is that it reflects the reality that the chances of being apprehended and convicted for online crime are alarmingly low. There is general agreement that cybercrime is widespread and growing but few authoritative estimates of its real scale. (One estimate puts the annual global cost at €750 billion) (Global Economic Symposium 2015). A study carried out in October 2014 reported that fully one half of Britons had experienced crime online, with offences ranging from identity theft and hacking to online abuse ('Thieves in the Night', *Economist*, 2014).

As far as companies are concerned, cybercrime is a real and growing threat and one that is chronically under-reported. According to a 2014 study by PricewaterhouseCoopers, 69% of UK companies had experienced a cybersecurity 'incident' in the previous year, but an earlier government inquiry found that businesses reported only 2% of such incidents to police ('Thieves in the Night', *Economist*, 2014).

The widespread public perception that cybercrime is carried out by opportunistic hackers is misguided. In fact, it is now a sophisticated global industry with its own underground economy, in which stolen personal and financial data are freely traded in covert online marketplaces. Stolen credit card details, for example, are available at prices in the

£1 range in such marketplaces (Hern 2015) and personally identifiable information, like social security numbers, fetch 10 times that (Greene 2015). Stolen data fuel a range of criminal activities (phishing, hacking of corporate systems, extortion via denial-of-service attacks) which are supported by 'a fully fledged infrastructure of malicious code writers, specialist web hosts, and individuals able to lease networks of many thousands of compromised computers to carry out automated attacks' (Global Economic Symposium 2015).

Since cybercrime is now a global industry, effective measures to deal with it require a co-ordinated international response. Although in some areas (e.g. child pornography) law enforcement agencies have shown that such co-operation can work, arrangements for dealing more generally with cybercrime are slow and patchy. A solution to the problem lies some distance ahead in the future.

'Balkanisation'[25]

Although the Internet has its origins in the US, from the outset it was conceived as a global network that transcended territorial boundaries. As the network expanded however, tensions began to emerge between this 'global' network and local customs, culture, and laws. The kind of free expression protected by the First Amendment to the US Constitution, for example, was deemed unacceptable in other cultures. Over time, these tensions extended to clashes between local laws and the operating and commercial assumptions of US-based Internet companies. A classic example was the fact that Nazi memorabilia which were freely available on American websites could not legally be sold in France (BBC News 2000).

Tensions between local jurisdictions determined to enforce their laws and cultural norms on the network have steadily increased over the last decade, leading to fears that the Internet would eventually be 'Balkanised', that is, split into locally controlled subnets. To a certain extent, this has already happened – for example, in China, which is now the largest Internet market in the world but where the government retains very tight and sophisticated control over the network (see e.g. King, Pan, and Roberts 2014). Other governments – notably those of Iran, the Russian Federation and some Middle-Eastern regimes – have also instituted increasingly tight control over use of the Internet by their citizens.

Until mid-2013, the momentum towards Balkanisation of the network was relatively modest. In 2013, however, revelations by the former NSA contractor, Edward Snowden, about the surveillance capability of US and allied governments, radically altered the picture. Snowden's revelations of the extent of US covert penetration and surveillance of network communications have provided authoritarian and quasi-democratic regimes with a convenient rationale for extending their control. And even in liberal democracies, the reassertion of local territorial rights has become more noticeable – as, for example, in the 'Schrems' judgment by the European Court of Justice (Court of Justice of the European Union 2015), which ruled invalid the 'Safe Harbour' agreement between the EU and the US, under which personal data of European citizens held by American Internet companies could be 'exported' to server farms in the continental US. As a result, US Internet companies like Microsoft have decided to hold European citizens' data on server farms located in EU countries (Ahmed and Waters 2015). So it is conceivable that a kind of *de-facto* Balkanisation is under way.

Changing patterns of media consumption

Since the Internet, in principle, enables anyone with a network connection to become a global publisher, it was seen at the outset as a radically different kind of medium from the *mass* media which had dominated the print and broadcast world. Whereas those earlier media were *few-to-many* systems, the Internet could be a *many-to-many* medium; its users could be *active* creators of content, rather than *passive* consumers of content created by others (see Benkler 2007). But as the network has evolved to connect billions of users, this early vision of its potential as a communications medium has been tempered by experience. Analysis of data traffic on the network suggests that the kind of passive consumption that characterised the broadcast era is returning. On North American landline connections, for example, Netflix – a movie-streaming service – accounts for 36.5% of downstream traffic in the peak evening hours (Deeth 2015). The network still offers great creative opportunities for its users, but the odds of it turning into 'billion-channel TV' may be shortening.

Changes in network infrastructure

In 1995, an estimated 16m people worldwide had access to the Internet. The current number of users is estimated to be in the region of 3.5 billion.[26] In 1995, all users accessed the network via fixed-line connections; currently over half of all users access it via mobile devices.[27] And most of the applications that are popular with contemporary users did not exist – and indeed would not have been feasible – on the network as it was in 1995. The remarkable thing about the Internet is that, while its infrastructure has had to evolve radically over those two decades to meet the ever-changing demands of its users, it still remains largely true to its fundamental architectural principles.[28]

This is an extraordinary achievement, made possible by what Greenstein (2015) describes as a 'combination of inventive specialisation and technical meritocracy'. Inventive specialisation evolved naturally from the early days because groups of engineers routinely coalesced around specific technical challenges in order to improve particular functions in the network. Technical meritocracy was likewise an inheritance from the early days of the ARPANET design – a prevailing ethos that ideas should succeed (or fail, as many did) on their technical merits, rather than on the organisational status of whoever proposed them (Greenstein 2015, 42–49).

One illustration of this process in action is the evolution of 'content delivery networks' (CDNs) – distributed networks of proxy servers located in multiple data centres across the world. Their function is to improve the speed and reliability with which digital content can reach users who demand it. Essentially, CDNs cache (temporarily store) digital content closer to where communities of users are located so that, for example, on-demand streaming video arrives promptly, and at higher quality than would be the case if it were being streamed directly from a central server farm on the other side of the globe. CDNs now serve a large proportion of contemporary Internet content, and without them, the kind of services that mobile users in particular take for granted would be impossible. In that sense, they are a rational technical solution to a problem that, if left unsolved, would have reduced the overall utility of a network that has increasingly been called upon to serve passive consumers of multimedia content.

Conclusion

This survey of the evolution of the Internet over four decades highlights a number of themes.

The first is the extent to which its development was socially shaped. The intrinsic affordances of digital technology did, of course, play an important role in the network's evolution. But it was also shaped by non-technological forces: its military provenance, for example; the surveillance-based business models that evolved to support the 'free' services provided by companies like Google and Facebook; the collaborative, non-commercial social ethos of the engineering community that developed the protocol layer of the network.

This last factor deserves special attention, because the ethos of the developer community was effectively 'baked into' some of the protocols. Just to take one example, the original (1982) version of the SMTP protocol (Simple Mail Transfer Protocol),[29] which governs how mail servers handle messages, had no provision for authentication – i.e. for checking that the 'sender' of a message was actually the real sender – because at that stage email traffic took place between researchers and institutions that were known to one another. But once the network was commercialised, absence of authentication was what enabled the spoofing of email addresses and enabled the rise of spam: the social context had changed.

Secondly, like many other technologies, the evolution of the Internet, from its earliest beginnings to a mature technology, involved *both* public investment and private capital. The ARPANET, and the development of the TCP/IP network that followed it, were exclusively funded by taxpayers' money, initially via the Department of Defense and later via the NSF. Privatisation of the network in the mid-1990s then brought in a torrent of investment capital in a five-year speculative bubble. But while this bubble caused a financial crash, it also resulted in a massive expansion in the communications infrastructure needed to turn the network into a ubiquitous public utility. In that sense, as the economist and venture capitalist William Janeway observes, *bubbles matter*.

> They matter because they not only transfer wealth from greater to less-great fools, and to the knaves that prey on the former. Occasionally – critically – they transfer wealth to fortunate opportunists and insightful entrepreneurs in the market economy who are granted access to cash on favourable terms and put it to work with astounding consequences. (2012, 181)

Thirdly, we can see an intriguing contradiction emerging between the affordances of digital technology as time progresses. In the period from 1995 to (roughly) 2005, the architecture of the network definitely facilitated 'permissionless innovation' (Van Schewick 2012). Software is pure 'thought-stuff', the barriers to entry were very low and so entrepreneurs and inventors like the founders of eBay, Google, Skype, Facebook, and others were able to launch, with very little capital, services and enterprises that eventually became global corporations.

But the technology has other distinctive affordances too, which became progressively more pronounced from 2005 onwards. The most significant of these affordances are: zero marginal costs; powerful network effects; the dominance of Power Law distributions in cyberspace; and technological lock-in (Anderson 2014). The resultant of these affordances points towards winner-takes-all outcomes – as seen, for example, in Google's domination of the search market, Facebook's in social networking and Amazon's in online retailing and cloud computing. So while the Internet in principle still facilitates permissionless innovation, the chances of insurgents displacing incumbents – at least in pure Internet businesses – seem less likely. Whereas Google was able to displace other search engines in 1996 simply by having a better page-ranking algorithm, nowadays a newcomer with an innovative idea in search will face an incumbent with huge troves of user data and large server farms distributed across the globe. The barrier to entry has thus been raised to a formidable extent.

What this suggests is that the Internet has evolved into a GPT, much as railway networks did in the late nineteenth century. At the same time, it is now mature and extensive enough to serve as a foundation on which new kinds of innovation, much of it in areas apparently unrelated to information goods, can build. In that sense, it is conceivable that enterprises like the cab-hailing application Uber, or the room-hiring service Airbnb, may turn out to be the contemporary equivalent of the mail-order services of the nineteenth century: unthinkable before the technology, and unremarkable afterwards.

Finally, there is the overriding issue of control. In his magisterial history of the dominant communications technologies of the twentieth century (Wu 2011), the legal scholar Tim Wu discerns a pattern – a cycle. Each new technology: the telephone, radio, movies, and television, initially engendered waves of creativity, excitement, and utopian hopes. But each, in the end, was 'captured' by corporate interests, sometimes with the connivance of government (as with telephony and AT&T). Wu notes that the Internet, at *its* inception four decades ago, also engendered the same kind of excitement and utopian hopes, and asks whether, in the end, it will suffer the same fate. It is a good question. The historian will say that it is too early to tell. Others may disagree.

Notes

1. 'The internet is the first thing that humanity has built that humanity doesn't understand, the largest experiment in anarchy we've ever had' (Dr Eric Schmidt, executive Chairman of Google, quoted in Taylor (2011)).
2. To use Herbert Butterfield's celebrated phrase, see Butterfield (1931).
3. To borrow the title of Lewis (1999).
4. http://www.internetlivestats.com/internet-users/ (accessed 11 November 2015).
5. For example, the first articulation of the ideas that eventually found their expression in hypertext and later in the World Wide Web appeared in a paper by Vannevar Bush which he wrote in 1939 but did not publish until after the war in 1945. See Naughton (1999, 214).
6. 'Circuit switching' refers to the technology of analogue telephony, in which communication was enabled by creating a continuous electrical circuit between sender and receiver which was maintained for the duration of a call; 'packet switching' involves disassembling a digitised message into discrete data 'packets' of uniform size, and then dispatching them through a mesh network of switches (or routers) which pass them on to other routers that are closer to their destination, where the packets are eventually reassembled to re-create the original message. For an explanatory animation, see https://commons.wikimedia.org/wiki/File:Packet_Switching.gif.
7. Abbate (1999) provides an excellent account.
8. This was expressed in an early declaration often attributed to Jon Postel, one of the students involved: 'we believe in rough consensus and running code'. 'Rough consensus' was

subsequently defined in the 'Working Group Guidelines and Procedures' of the Internet Engineering Task Force (IETF):

> Working groups make decisions through a 'rough consensus' process. IETF consensus does not require that all participants agree, although this is, of course, preferred. In general, the dominant view of the working group shall prevail (However, 'dominance' is not to be determined on the basis of volume or persistence, but rather a more general sense of agreement). Consensus can be determined by a show of hands, humming, or any other means on which the WG agrees (by rough consensus, of course). Note that 51% of the working group does not qualify as 'rough consensus' and 99% is better than rough. It is up to the Chair to determine if rough consensus has been reached. (1998, 13)

9. Subsequently, the spirit of this design philosophy was encapsulated in a variety of memorable phrases: 'dumb network, smart applications'; the 'end-to-end' principle; and 'Net Neutrality'.
10. Bill Joy added an implementation of TCP/IP to the Berkeley variant of Unix in 1982. See McKusick (1999).
11. Abbate (1999, 186) and http://www.netvalley.com/intvalstat.html.
12. For a comprehensive study, see Greenstein (2015).
13. See https://en.wikipedia.org/wiki/Peering (accessed 29 January 2016).
14. A computer application that is so compellingly useful that people are motivated to purchase or adopt the software or hardware on which it runs. The first such application was email which, as observed earlier, took the designers of the ARPANET by surprise.
15. Telnet was an Internet protocol which provided users with a command-line interface to a remote host.
16. See Cassidy (2002) for an excellent history.
17. Which is why 'get big fast' was the dominant mantra of the time in pitches to venture capital firms.
18. Internet Live Stats, http://www.internetlivestats.com/internet-users/ (accessed 9 November 2015).
19. Internet Live Stats, http://www.internetlivestats.com/internet-users/ (accessed 9 November 2015).
20. The first iPhone was launched by Apple in the summer of 2007.
21. For a prescient summary, see Zittrain (2008)
22. Defined as

    > web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with they share a connection, and (3) view and traverse their list of connections and those made by others within the system. (Boyd and Ellison 2007, 211)

23. Internet security expert Bruce Schneier, speaking at a cybersecurity conference in Boston on 9 April 2014 https://www.schneier.com/news/archives/2014/04/surveillance_is_the.html (accessed 11 November 2015).
24. http://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/ (accessed 11 November 2015).
25. This term is understandably irritating to citizens of the Balkans, but has become shorthand in Internet governance discussions.
26. http://www.internetworldstats.com/emarketing.htm.
27. http://www.statista.com/statistics/284202/mobile-phone-internet-user-penetration-worldwide/.
28. Though there are continual tensions over principles like 'Net Neutrality'. See, for example, Madrigal and Lafrance (2014).
29. https://tools.ietf.org/html/rfc821 (accessed 29 January 2015).

Acknowledgements

The author gratefully acknowledges helpful comments from Richard J. Evans, Rolf Fredheim, Alfred Moore, Andrew McKenzie-McHarg, Rachel Hoffman, Richard Danbury, Nóra Ní Loideain, Daniel Wilson, Julia Powles and two anonymous referees.

Disclosure statement

No potential conflict of interest was reported by the author.

References

Abbate, Janet. 1999. *Inventing the Internet*. Cambridge, MA: MIT Press.

Abutaleb, Y., and L. Maan. 2015. "Facebook Revenue, Profit Beat Forecasts; Shares Hit All-Time High." *Reuters*, November 4. Accessed November 11, 2015. http://tinyurl.com/qeov4nv.

Agamben, Giorgio. 2005. *State of Exception*. Chicago, IL: University of Chicago Press.

Ahmed, Mured, and Richard Waters. 2015. "Microsoft Unveils German Data Plan to Tackle US Internet Spying." *Financial Times*, November 11.

Anderson, Ross. 2014. "Privacy Versus Government Surveillance: Where Network Effects Meet Public Choice." http://weis2014.econinfosec.org/papers/Anderson-WEIS2014.pdf.

Arthur, Brian W. 2011. *The Nature of Technology: What It Is and How It Evolves*. London: Free Press.

BBC News. 2000. "France Bans Internet Nazi Auctions." May 23. Accessed November 11, 2015. http://news.bbc.co.uk/1/hi/world/europe/760782.stm.

Benkler, Yochai. 2007. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven, CT: Yale.

Berners-Lee, Tim. 2000. *Weaving the Web: The Past, Present and Future of the World Wide Web*. London: Texere.

Bond, Robert M., Christopher J. Fariss, Jason J. Jones, Adam D. I. Kramer, Cameron Marlow, Jaime E. Settle, and James H. Fowler. 2012. "A 61-Million-Person Experiment in Social Influence and Political Mobilization." *Nature* 489: 295–298.

Boyd, D. M., and N. B. Ellison. 2007. "Social Network Sites: Definition, History and Scholarship." *Journal of Computer-Mediated Communication* 13 (1): 210–230.

Bresnehan, T. 2010. "General Purpose Technologies." In *Handbook of the Economics of Innovation*, Vol. 2, edited by B. H. Hall and N. Rosenberg, 761–791. Amsterdam: North-Holland.

Broadband Commission. 2014. *The State of Broadband 2014*. September. International Telecommunications Union and UNESCO.

Butterfield, H. 1931. *The Whig Interpretation of History*. London: G. Bell.

Cassidy, John. 2002. *Dot-con: How America Lost Its Mind and Its Money in the Internet Era*. New York: Harper Collins.

Court of Justice of the European Union. 2015. "The Court of Justice Declares that the Commission's US Safe Harbour Decision Is Invalid." October 6. http://curia.europa.eu/jcms/upload/docs/application/pdf/2015-10/cp150117en.pdf.

Deeth, Dan. 2015. "Global Internet Phenomena Report: Latin America and North America." May 28. Accessed November 18, 2015. http://www.internetphenomena.com/page/5/.

DeLong, J. Bradford. 2003. "Profits of Doom." *Wired*, 11 (4). Accessed November 11, 2015. http://www.wired.com/2003/04/profits-of-doom-2/.

Geier, Ben. 2015. "What Did We Learn from the Dotcom Stock Bubble of 2000?" *Time*, March 12. Accessed October 10. http://time.com/3741681/2000-dotcom-stock-bust/.

Gibson, William. 1984. *Neuromancer*. New York: Ace.

Gillies, James, and Robert Cailliau. 2000. *How the Web Was Born: The Story of the World Wide Web*. Oxford: Oxford University Press.

Goldfarb, Brent, David Kirsch, and David A. Miller. 2007. “Was There Too Little Entry During the Dot Com Era?” *Journal of Financial Economics* 86 (1): 100–144.

Global Economic Symposium. 2015. “Dealing with Cyber Crime – Challenges and Solutions.” Accessed November 11, 2015. http://www.global-economic-symposium.org/knowledgebase/the-global-polity/cybercrime-cybersecurity-and-the-future-of-the-internet/proposals/dealing-with-cyber-crime-2013-challenges-and-solutions.

Greene, Tim. 2015. “Anthem Hack: Personal Data Stolen Sells for 10X Price of Stolen Credit Card Numbers.” Network World. February 6. Accessed November 11, 2015. http://www.networkworld.com/article/2880366/security0/anthem-hack-personal-data-stolen-sells-for-10x-price-of-stolen-credit-card-numbers.html.

Greenstein, Shayne. 2015. *How the Internet Became Commercial: Innovation, Privatization, and the Birth of a New Network*. Princeton, NJ: Princeton University Press.

Hafner, K., and M. Lyon. 1996. *Where Wizards Stay Up Late: The Origins of the Internet*. New York: Simon & Schuster.

Hauben, Michael, and Ronda Hauben. 1997. *Netizens: On the History and Impact of Usenet and the Internet*. Los Alamos, NM: IEEE Computer Society Press.

Hern, Alex. 2015. “Stolen Credit Card Details Available for £1 Each Online.” *The Guardian*, October 30. Accessed November 11, 2015. http://www.theguardian.com/technology/2015/oct/30/stolen-credit-card-details-available-1-pound-each-online.

IETF (Internet Engineering Task Force). 1998. “IETF Working Group Guidelines and Procedures.” RFC 2418, September. Accessed November 4, 2015. https://tools.ietf.org/html/rfc2418.

Janeway, William H. 2012. *Doing Capitalism in the Innovation Economy: Markets, Speculation and the State*. Cambridge: Cambridge University Press.

Keen, Andrew. 2015. *The Internet Is Not the Answer*. New York: Atlantic.

King, Gary, Jennifer Pan, and Margaret E. Roberts. 2014. “Reverse-Engineering Censorship in China: Randomized Experimentation and Participant Observation.” *Science* 345 (6199): 1–10. Accessed January 25, 2015. http://j.mp/16Nvzge.

Kramer, Adam D. I., Jamie E. Guillory, and Jeffrey T. Hancock. 2014. “Experimental Evidence of Massive-Scale Emotional Contagion Through Social Networks.” *PNAS* 111 (24): 8788–8790. doi:10.1073/pnas.1320040111.

Lewis, Michael. 1999. *The New New Thing: A Silicon Valley Story*. New York: W.W. Norton.

Madrigal, Alexis C., and Adrienne Lafrance. 2014. “Net Neutrality: A Guide to (and History of) a Contested Idea.” *The Atlantic*, April 25. Accessed January 20, 2016. http://www.theatlantic.com/technology/archive/2014/04/the-best-writing-on-net-neutrality/361237/.

McKusick, Marshall Kirk. 1999. “Twenty Years of Berkeley Unix: From AT&T-Owned to Freely Redistributable.” In *Open Sources: Voices from the Open Source Revolution*, edited by Chris DiBona, Sam Ockman, and Mark Stone, 31–46. Sebastopol, CA: O’Reilly.

Moo, Craig. 2016. “The Facebook-Loving Farmers of Myanmar.” *The Atlantic*, January 21. http://www.theatlantic.com/technology/archive/2016/01/the-facebook-loving-farmers-of-myanmar/424812/.

Naughton, John. 1999. *A Brief History of the Future: The Origins of the Internet*. London: Weidenfeld.

Naughton, John. 2012. *From Gutenberg to Zuckerberg: What You Really Need to Know about the Internet*. London: Quercus.

Naughton, John. 2015. “The Rise of Ad-blocking Could Herald the End of the Free Internet.” *The Observer*, September 27. Accessed November 11, 2015. http://www.theguardian.com/commentisfree/2015/sep/27/ad-blocking-herald-end-of-free-internet-ios9-apple.

O’Reilly, Tim. 2005. “What Is Web 2.0? Design Options and Business Models for the Next Generation of Software.” September 30. Accessed November 11, 2015. http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html.

Postel, Jon, ed. 1981. “Transmission Control Protocol: Protocol Specification.” RFC 793, September. Accessed November 21, 2015. https://tools.ietf.org/html/rfc793.

Schiller, Robert J. 2000. *Irrational Exuberance*. Princeton, NJ: Princeton University Press.

Schmidt, Eric, and Jared Cohen. 2008. *The New Digital Age: Reshaping the Future of People, Nations and Business*. London: John Murray.

Schumpeter, Joseph A. 1942. *Capitalism, Socialism and Democracy*. New York: Harper.

Taylor, Jerome. 2011. "Google Chief: My Fears for Generation Facebook." *Independent*, October 22. Accessed November 21, 2015. http://www.independent.co.uk/life-style/gadgets-and-tech/news/google-chief-my-fears-for-generation-facebook-2055390.html.

"Thieves in the Night." 2014. *Economist*. December 20.

Van Schewick, Barbara. 2012. *Internet Architecture and Innovation*. Cambridge, MA: MIT Press.

Wu, Tim. 2011. *The Master Switch: The Rise and Fall of Information Empires*. New York: Atlantic Books.

Zittrain, Jonathan. 2008. *The Future of the Internet – and How to Stop It*. New Haven, CT: Yale.