Chapter 11

Making Links: Everything Really Is Deeply Intertwined

A Talk for Ted

Wendy Hall

Ted, this is for you. I've flown in for this event from Hong Kong. If I become incoherent it's because there's a 15-h time difference. I'm flying to London tonight because I have to be back for the weekend, so I can't stay for the dinner, but I wanted to be here for you. I only have 30-min and I've got a lot to say, in fact, I've got about 30 years of stuff to say. It is a great honor to be here. Thank you very much for inviting me. I really wanted to honor Ted, and support this wonderful event.

Throughout my life I have always tried to make links. When I got my chair at Southampton my inaugural lecture was entitled, "Making Links." It was about the hypermedia work we were doing at Southampton, but I'm also a social linker and I've always tried to link the different research communities, such as hypermedia, multimedia, the Web, the Semantic Web and others, to try to be a bit of glue in there that gets everybody talking together. Over my life I've found that everything really is deeply "intertwined" and I'm very proud to say that this is my first Ted Talk. Well, it depends how you parse that. I first heard "TED Talk" in 1989, but I'm calling this *my* first Ted Talk. Now I'm going to tell you about me because I reckon if I do that Ted can't say I'm wrong. He could of course challenge the references to how he has inspired my life but I'll let him do that.

This talk is based on a standard talk that I give, but intertwined with how Ted has inspired my career and my work, and my life generally. We were both inspired by Vannevar Bush's paper "As We May Think" [2]. Ted read it long before I did of course. I first read it in about 1987. I started my career as a pure mathematician—my PhD topic was algebraic topology—but in the Eighties I became increasingly interested in the application of personal computers in education. I got very, very interested in what we would now call multimedia. Do you remember the old twelve-inch analog video discs? I got really excited about how you could put a video on a computer and then teach people using this new interactive media.

W. Hall (✉)

Web Science Institute, University of Southampton, Southampton, UK

e-mail: wh@ecs.soton.ac.uk

As the new personal computers emerged onto the scene, I gradually moved more and more into computing. I went back to the University of Southampton in 1984 as a member of the computer science faculty, which surprised everybody including me because I loved pure mathematics. But the rest, as they say, is history because that move opened so many doors for me. So around 1986/1987, as I was beginning to find my feet as junior member of faculty, I started working in this new exciting area of multimedia.

In 1987 when I read Vannevar Bush's paper I also began to hear about this 'new' idea called hypertext. I began to hear about Ted and Doug Engelbart, both of whom equally inspired me: Ted talking about everything being deeply intertwined, and Doug, talking about augmenting the human intellect. Again, I don't need to tell this audience about these two men. When I give talks to a non-expert audience I always include reference to them because it was their ideas—I hadn't met them at this point—that inspired me. The year of 1987 was a key one for hypertext. It was the year of the first ACM Hypertext conference, and the year Apple released HyperCard. It was also the year that the archive of the Earl Mountbatten of Burma arrived at the University of Southampton. Mountbatten was a cousin of the Queen, and very famous in the UK for his various military leadership roles during and after the Second World War and as the last Viceroy of India. What does this have to do with my research story?

The Mountbatten family estate is just outside Southampton, and after he died in the 1970s, the University of Southampton took over custodianship of his archive, which consisted of about 250,000 papers, 50,000 photos, audio recordings of his speeches and various film and video recordings. This was multimedia as it was in 1987. It was pre-digital library. Nothing was born digital then, we had to go through the process of digitizing it before we could view it on a computer let alone over a network. I thought how marvelous it would be if we could digitize all this material and use this new hypertext/hypermedia idea to make it available to anybody who wanted to find out what was in the archive without having to come to Southampton. We could link it all up! And we could, maybe, have different links for different people, so that if school children wanted to find out about something that was in the archive they would get different links to historians who were looking for evidence of what had happened when and why.

As I was mulling these ideas over, I was lucky enough to have a 6 month sabbatical at the University of Michigan in Ann Arbor in 1989, and this is when I first heard Ted talk at a Computers in the Humanities conference in Toronto. He was the keynote speaker. I was spellbound. I bought a copy of his book *Literary Machines* [5], and he signed it for me. We didn't really get talking at that conference but the book became my hypertext bible. I taught from it. I learned about hypertext and Ted's definition of the link and everything about Xanadu, tumblers, transclusions, micropayments and much, much more. They were all such exciting ideas, so much ahead of their time. Ideas that the world still hasn't fully appreciated, solving problems that the world is only just beginning to realize it has.

I came back from that sabbatical year really fired up about what we could do with digital archives and what was then becoming called the digital library. At Southampton we built a hypermedia system called Microcosm using the Mountbatten

archive as our first demonstrator [3]. I wanted the system to be able to automatically create links on keywords in the application such as linking the names of key characters mentioned in the Mountbatten archives to their biographies or to a photograph. We called these generic links and they became a significant feature of Microcosm. Ted said we should have called them something other than links, as a generic link didn't fit his definition of a hypertext link, but by the time Ted saw Microcosm it was too late to change the naming of the links. Microcosm was an open hypermedia system in that all the links were stored in a database as first-class entities that could be reasoned about and applied to any document. Each link was a triple that consisted of a source, a destination and a description. Little did I know at the time how prescient of the Semantic Web these ideas would be. Of course, there are problems with automatically making a link on a word without knowing its precise semantic meaning. There are a lot of different people with the name Mountbatten in the Mountbatten archive for example. So working out the context in which the link was being applied and therefore the meaning of the word became a key focus of our work: problems we are still dealing with as the Semantic Web develops today.

We did also have specific links in Microcosm that were more like standard hypertext links because they were embedded in the documents and represented to the user through highlighted buttons, and you could trace them backwards though the link database or linkbase as we called it. But the really novel idea was the generic links that were stored in a separable hyper-structure and created on the fly.

We had the concept of a viewer in Microcosm, which you might call a browser today. We had a control system, which is where we went wrong because everything was centralized around that control system, even though the system itself could be distributed. And when the user said, "I want to see what links we've got on this word," then the query would be run through the control system and through a set of filters which determined the list of links that came back to the user through the link dispatcher. I had a lot of post-graduate students developing both the main system and experimenting with different types of filters. It was a great plug and play system for experimental research. There were a lot of ideas in this system that you now see working at scale in the Web and in the Semantic Web.

Remember we were doing all this using personal computer workstations and videodisc players. It all seems very primitive now when you think about the way we use the Web on our mobile phones, but the ideas ran deep. We developed tools to generate links based on the metadata description of documents, and if you do that the *docuverse*, to use Ted's term, just falls out. The pure mathematician in me just loved making these patterns. We also generated links on the *data* in the documents and using that idea we able to integrate (make links between) different types of documents: text documents, multimedia documents, spreadsheets, databases, etc.. We created links in picture, video, and audio files, and we did lots of experiments using the data to create links in multimedia archives. It was all very exciting and very prescient of the work we are doing in the Semantic Web and the world of big data today. I'm very proud to say that and not at all bitter when today's researchers don't make the links back to previous work. I just think how privileged we were to be able pioneer these ideas in the research lab and what fun it was.

Meanwhile, of course, at CERN in Switzerland, Tim Berners-Lee was thinking his own great thoughts that were going to change everything. It was 1989 when he wrote his paper “Information Management: a Proposal” for his boss at CERN in which he defined what would become the World Wide Web. At the time, it was quite complicated to download a document from the Internet. Most people used FTP. There were emerging systems such as Gopher and WAIS that were providing easier interfaces and that had elements of hypertext functionality but were not that easy to use or readily available. In Tim’s paper you see all the elements that he had distilled from the hypertext community, the generalized mark-up community, and the network community. It was those three ideas that came together in his design for a global hypertext system. His main aim was to enable anybody, but in particular physicists, to share information over this new thing called the Internet, basically at the click of a button. Tim’s boss, Mike Sendall, marked his proposal as “Vague, but exciting” and this gave Tim permission to carry on developing his ideas as part of his day job.

I first met Tim at the European Hypertext Conference in Paris in December 1990. Our first paper on Microcosm was accepted at the conference [3]. Tim was there with his colleague from CERN, Robert Cailliau, talking about the system they were building that they hadn’t yet called the World Wide Web. The list of attendees at the conference is available on the Web now so you can see who was there.

In 1991, the ACM Hypertext conference was in San Antonio in Texas. Tim had created the first website over the Christmas holidays in 1990, and he christened the system the World Wide Web. It was my first trip to Texas, one of my first trips to the States for work. This was the conference that famously rejected Tim and Robert’s paper on the World Wide Web. Our second paper on Microcosm also ended up on the reject pile for this conference. What you do when you have a paper rejected from a conference is you submit a poster or a demonstration so you can still go to the conference. I was still junior faculty at the time remember. So we were in San Antonio demoing the Microcosm system, and Tim and Robert were demoing the World Wide Web. This was the first time I saw them demo it. I remember looking over Tim’s shoulder and thinking, “There’s nothing new here. This hypertext system is hardly hypertext.” In those early days, I used to set my students an essay with the title, “Is the Web hypertext?” because according to all the definitions at the time it wasn’t. And I remember thinking “he embeds his links in the documents and they only go one way!” I mean, this was just so primitive. It wasn’t going to go anywhere. And I wasn’t the only person thinking that at the time. How wrong we all were.

I also remember thinking how pretentious to call it the World Wide Web. How was he going to persuade the whole world to use this rather primitive system? Well, he did. It has become the hypertext system that the whole world uses to exchange information on the Internet. Tim understood that the network was everything, and he recognized the need for the system to be distributed, decentralized and open, with universal standards to enable it to scale. He got the fact that scruffy works, in other words he allowed for human frailty by allowing for the links to fail. He also gave it away so there was no economic barrier to people using it. I think that’s what

really made the difference to the uptake. The problem we now have of course is that nobody owns the system that has subsequently been created, but that's another whole story.

We must also give credit of course to Marc Andreessen and his team at NCSA for their development of the Mosaic browser for the Web in 1993, which made it much easier for people to access the Web. At ACM Hypertext'93, half the demos were Web-based, and the first Web conference was in May 1994. We were still developing Microcosm as a research system, and we produced a commercial version in 1994, which did very well for a while. The company we set up raised over £13 million pounds of investment funding and is still in existence today, but by the end of the 1990s my research group was almost entirely focused on Web-based developments and much of the interesting hypermedia work encapsulated within the Microcosm project had to be shelved until the Semantic Web was mature enough to become an alternative development platforms. But of course, like others we were riding the wave of a truly World Wide Web which was changing society in ways it would have been hard to imagine just 10 years previously.

It was during this transition period at IEEE Multimedia 1996 in Hiroshima that Ted started to become interested in the Microcosm project. Two of our PhD students—Stuart Goose and Jonathan Dale—presented a paper about Microcosm at the conference and spent time talking about our ideas there with Ted [4]. They invited Ted back to Southampton to talk more and meet the rest of the team including me. I remember sitting with Ted and Marlene in the staff bar at the University just talking and talking. We talked all night I seem to remember, or as long as the bar was open anyway. To cut a long story very short this led to Ted becoming a visiting professor at Southampton and to him spending a year with us as a part of a visiting fellowship scheme. It was an amazing year. It was during this year that Ian Heath—one of the original Microcosm team—developed the first version of Cosmic that demonstrated Ted's ideas showing the connections, or links, between the same parallel documents. We could do this in Microcosm because the links were in separable hyperstructure. That bit was easy. The hard bit, as always was, the visualization. Ian did a lot of work on this for Ted to produce the Cosmic book demo.

The other thing that was happening at the time was that the web community had spun away from the hypertext community, and I couldn't get the two communities to work together. The WWW conference had taken on a life of its own and the hypertext community—despite the fact that the Web was really launched at a Hypertext conference—resolutely refused to accept that its future was the Web. We had agreed to host ACM Hypertext'97 in Southampton, and we had lined Ted up as the keynote speaker. I remember saying to Robert Cailliau, who was co-ordinating the Web conference series, "We have fixed our dates, please don't fix the WWW conference for the same week." And sure enough, when the dates for WWW'97 came out there was a complete clash. Bebo White from Stanford was the WWW conference chair. I was so determined to try to make links between the two communities that I negotiated with Bebo for Ted's keynote to be beamed live across the Atlantic and for us to have a joint panel session afterwards. I chaired the panel in Southampton, and Robert chaired the panel in Santa Clara. The technology to

achieve this in those days was really very complicated. Dave De Roure managed the technology at our end. I don't know who managed it in the U.S., but it was quite an achievement and it all worked amazingly well.

The Web community is a very different community from the hypertext community culturally, and I don't know how much they understood of what Ted said in his keynote. The most memorable thing for me were the words Ted used to close his keynote—given to a hypertext community in Southampton but addressed to the Web community in Santa Clara—"Your future is my past."

That is so very true, but even now most people in the Web community have no idea just how true it is. As I reflect on how the Web has developed over the last 25 years, I see the ideas that Ted had 30, 40, 50 years ago emerging because they have to emerge. But it's taken a long while, and whilst I love the Web for all it's given us I also understand exactly how it slowed things down as well. I felt this with the Microcosm project too.

There was another major seminal moment for me at WWW'98, which was in Brisbane. Three things happened at that conference as I recall. Tim started talking about the Semantic Web again in his keynote for the conference. He had talked about it at the first WWW conference in 1994 [1] and the idea of making links on data in the information management proposal he wrote in 1989. As far as he was concerned in 1998, the web of linked documents was beginning to emerge but his vision wasn't complete until it was also a web of linked data, and so he started to re-educate the community about this at the Brisbane conference.

Ted was also at the Brisbane conference to pick up a special award. I remember him demoing ZigZag to us in the bar one night at that conference. He was so excited, and we were all mesmerized. So I had heard Tim talk about the Semantic Web and I saw Ted demo ZigZag at the same conference, and I didn't fully appreciate either of them at the time. I understood the principles, but I didn't understand the detail. It's taken me a long time to appreciate both the Semantic Web and ZigZag, but as my understanding of both of them has increased I now firmly believe what I suspected all along: there is a one-to-one correspondence between the two ideas, and that you can implement ZigZag in the RDF graph.

Someday I'll find the time to prove that. I need to get Ted involved in making that happen. I really believe that these two amazing people—Tim and Ted—have the same idea of how you can make links on data to create an incredibly rich hyperstructure for generating knowledge. Tim will never talk about it like that. His idea with the Semantic Web is that machines can, if you describe the data using a vocabulary like an ontology, make inferences about the information contained in the data that couldn't be made in any other way. This is what Ted said in the paper he sent us about his closing keynote for this event, that actually if you take all these ideas to their extreme we will generate more knowledge.

The other thing I remember about the WWW conference in Brisbane is that this is where Sergey Brin and Larry Page published their paper about the algorithm that became the Google search engine. So for me this really was a seminal conference with so many truly ground breaking ideas emerging at the same time, apparently orthogonal to each other but actually all the same thing as time has confirmed, since

the Google Knowledge Graph is the Semantic Web or ZigZag by another name. It's all about linking data. This is a much quieter revolution than that initiated by the document Web but it will be much more far reaching. Linked data will become an integral part of the development of data-driven systems architectures that will revolutionize the way we build and maintain information management systems. Linked data architectures will supersede relational databases, make websites easier to build and unify the worlds of hypertext, document management, and databases to create rich interlinked knowledge-based systems as envisaged by the pioneers such as Ted and Doug over 50 years ago.

But the linked data revolution was very slow to take off—largely because it's hard to explain the key concepts to people and what the benefits are. In 2004, it seemed to have completely stalled. Analyzing why this was the case is a much longer story than I have time to tell here, but as a by-product of doing this analysis at the time, Tim, Nigel Shadbolt, Danny Weitzner, and I started to look back at the factors that made the web of linked documents take off in order to try and understand why the web of linked data wasn't. We realized that to understand the ecosystem that is the Web we have to take a socio-technical approach. It cannot just be thought of from the perspective of computer science. For good or bad we called this new approach Web Science, and set about launching a new research and education discipline around this idea. This has been the focus of my work ever since.

I passionately believe that this is a very important new area of study. We work with social scientists and people who understand human and organizational behavior, and it is a very interdisciplinary activity. I spend a lot of time now bringing together different disciplines to study how the Web has emerged, what it means for the future in terms of policy making and what it means for society—how society shapes it as much as how it shapes us. Through the Web Science Trust we organize conferences and workshops, and we have created an international network of web science laboratories. We have also just launched a major new Web Science Institute at Southampton. For me, one of the most exciting things about Web Science is that it attracts as many women as men because it is so interdisciplinary, and so for the first time in my career my classes have as many women as men—unlike the average computer science class! Last year we launched a MOOC (massively open on-line course) in Web Science—a case of using the medium to teach about the medium. If you are interested to learn more, it can be found at <http://www.southampton.ac.uk/moocs/webscience.shtml>.

Another major project being managed by the Web Science Trust is the development of a distributed global repository of data that people can use to do longitudinal research in this area. The idea is to enable researchers around the world to share data and data analytics about what is happening on the Web in the same way as the astronomers do with the data they collect from their telescopes. The physicists are trying to explain the mysteries of the physical universe, we are trying to do the same for the digital universe.

What's next? Well, I'm becoming increasingly involved in policy work in areas such as net neutrality, Internet governance, and security, privacy, and trust on the Internet. These are huge issues for us as a global society. I'm a member of the new

Global Commission for Internet Governance being organized by Chatham House and the Canadian Centre for International Governance Innovation. There's talk in the UK of a charter for Internet access rights. The Magna Carta is 800 years old in 2015, so it is very timely to talk about a Magna Carta for the Internet in that year. In these days when we have such tensions between the need for digital surveillance versus the protection of personal privacy on-line, Internet governance is a major topic that needs to be seriously addressed across the world.

As I come to the end of my talk, I want to turn back to the reason why I'm here. I want to pay honor to Ted for so many different reasons. Ted has inspired me throughout my entire career. But it has also been wonderful getting to know him as a person and to be able to talk about Ted and Marlene as two of our closest friends. My husband, Peter, and I have shared some marvelous times with them—one of the most memorable being when we had Doug and Ted and Marlene and Karen to dinner at our house. What an evening! We also have wonderful memories of spending time with them on their houseboat in Sausalito. But there is one memory that just Ted and I share. Ted always works late into the night as I often do. One night we had both been working late at Southampton and long after everyone else had gone home we found ourselves walking out of the office to the car park together. It was mid-winter, dark and a bit misty and Ted said to me, "Can you hear the nightingale's singing, Wendy?" Without Ted I would never have noticed them. And that about sums up my career in hypertext. So Ted, thank you for the nightingales and thank you for letting me share your world.

Open Access This chapter is distributed under the terms of the Creative Commons Attribution Noncommercial License, which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Berners-Lee T (1994) Plenary at the first World Wide Web Conference. Geneva, Switzerland. <http://www.w3.org/Talks/WWW94Tim>. Accessed 20 Jan 2015
2. Bush V (1945) As we may think. *Atlantic Monthly* 176(1):101–108
3. Fountain A, Hall W, Heath I, Davis H (1990). Microcosm: an open model for hypermedia with dynamic linking. *Hypertext: concepts, systems and applications: proceedings of the first European conference on hypertext, INRIA, France, November 1990: 298–311*. Cambridge University Press
4. Goose S, Dale J, Hill G, de Roure D, Hall W (1996). An open framework for integrating widely distributed hypermedia resources. In: *Proceedings of the third IEEE international conference on multimedia computing & systems*
5. Nelson TH (1981) *Literary machines: the report on, and of, Project Xanadu concerning word processing, electronic publishing, hypertext, thinkertoys, tomorrow's intellectual revolution, and certain other topics including knowledge, education and freedom*. Theodor H. Nelson, Swarthmore