For everything: Tim Berners-Lee, winner of the 2016 Turing award for having invented... the Web

Fabien Gandon

For everything: Tim Berners-Lee, winner of the 2016 Turing award for having invented… the Web

Fabien GANDON, Inria, Université Côte d’Azur, CNRS, I3S

Introduction

“I am myself and my circumstances.” – Jose Ortega

What could be a common point among getting information about a concert, transferring money from one’s bank account, publishing a genomics database, communicating with one’s children on the other side of the planet, and accessing information about one’s car ? [1] The fact that they can all be done through the Web. It is indeed difficult to think of a human activity that hasn’t been impacted by the Web, and as I write this article in April 2017, the current estimate is that the Web has more than 3 billion direct users throughout the world. In the same month, the British computer scientist Sir Timothy John Berners-Lee was awarded the 2016 Turing award for having invented the World Wide Web, the first Web browser, and the protocols and algorithms that enabled the passage to the scale of the Web [2]. Sir Tim, as he is called, is a professor at MIT and at Oxford University. This award is just the latest of a long list of distinctions that he has received. But the Turing award is considered the Nobel prize of computer science, and this award has been widely expected for his invention of the Web, an invention which has transformed our society since its creation in 1989. This is thus a good occasion to revisit, in this article, the history of his invention, and at the same time try to highlight many of the influences and currents that interacted to bring it about. This will also be an occasion, for us, to deconstruct certain notions and to reintegrate others, in an effort to tie together the numerous influences that weaved the web.

From Turing to Berners-Lee: a brief prehistory of the Web

“*[...] perched on the shoulders of giants,*
so that we can see more than they and at a greater distance.”

– Bernard de Chartres, XII^{e} century

Turing’s work is obviously omnipresent in computer science, but when it comes to the Web there is one particular example, that of the CAPTCHA: the awful tests that the Web makes us take all too often, with the pretext of verifying that we are human. The acronym (“*Completely Automated Public Turing test to tell Computers and Humans Apart”*) literally indicates that it’s a specific example of Turing’s imitation game, in a completely automated way and with the goal of differentiating humans from machines [26]. Incidentally, it’s also, particularly for certain giants of the digital world, a way of obtaining free brain time [27] and to Web-source the massive labeling of databases for learning algorithms (*machine learning*), for example. Beyond this anecdotal example of a direct link between Turing and the Web, we can also identify several major influences that had a more fundamental influence on the invention of the Web, including certain from Turing's time.

The search for techniques to organize and the means to efficiently access masses of collected information was an omnipresent motivation in the prehistory of the Web. In July 1945, Vannevar Bush (MIT) wrote the article "*As we may think,"* [28] in which he expressed concern that scientists would be unable to deal with the growing volume of publications that were relevant for their work. Bush thus posed, as a scientific challenge in itself, the problem of how to improve ways of accessing knowledge. He proposed, as a first element of a response, the MEMEX system, which would use mnemonic index codes to point to and rapidly access any part of imported documents. These attachment points would also allow links to be created between any two elements, thereby externalizing the points of association. We even find the literal notion of a "web" of ideas in this historic article, when Bush talks of associations of ideas in the following terms: "*the association of thoughts, in accordance with some intricate web of trails carried by the cells of the brain"* [28].

Technically, to bring about such a system, it was necessary to wait 20 years, for the appearance of computers and for their initial use in text editing. In an article prepared for the 1965 ACM conference, Ted Nelson proposed "a file structure for complex, changing and indeterminate information," [29] and he coined a neologism to refer to this collection of interconnected texts: hypertext, and even hypermedia. The subsequent years witnessed the development of the first hypertext editors using, in particular, another invention from the same decade: the mouse, with the interfaces and interactions that it allows. We owe these inventions in particular to Douglas Engelbart of the Stanford Research Institute (SRI), who himself received the Turing award in 1997 for his pioneering and inspiring vision of man-machine interactions and the key technologies that allowed their development. In particular, his NLS ("oN-Line System") system combined, in the 1960s, hypertext, a text-editing interface, and the mouse. This system would later be renamed *Augment*, in reference to the search program of Engelbart who, with an eye toward the pioneering work of the period in artificial intelligence, proposed working on augmented intelligence [30].

Although the hypertext concept had then been created, it remained, for years, essentially limited to applications that were locally executed on a computer. Communication involving packet switching (inter-networking), which was developed between 1972 and 1975 with the work of Louis Pouzin (IRIA and then Inria [31]), Vinton Cerf (SRI), and Robert Kahn (DARPA, [32]), culminated in 1978 with the (TCP/IP) standards and the beginnings of the Internet. Network communications applications then multiplied: email (SMTP), mailing lists, file transfers (FTP), remote connections (Telnet), discussion groups, etc. Here as well, we find two Turing award winners from 2004: Vinton Cerf and Robert E. Kahn.

This all-too-short intellectual genealogy already places Timothy Berners-Lee on the shoulders of giants when he arrived in 1980 as a consultant for CERN (Centre européen de recherche nucléaire), as a young physics graduate (Queen's College, Oxford) and self-taught computer scientist. Faced with the quantity of information and documentation that he had to deal with in his work and with his group at the CERN, Tim wrote a program (*Enquire*) to store information and to provide an unlimited number of links within it, in particular to document the programs, resources, and personnel with which he was working. An important difference between Enquire and the other hypertext systems of the era is that it worked on a multi-user operating system and thus allowed multiple users to access and contribute to the same data [6]. In comparison to the CERNDOC documentation system, which was based on a constrained hierarchical structure, Tim recalls that the arbitrary and bidirectional links within Enquire were more flexible and scalable. Tim then returned to the private sector between 1981 and 1983, where he worked on real-time remote procedure calls, and thus in the field of networks and network programming, before returning to CERN in 1984.

Crystallization point: breaking the mesh and the birth of a web

"The World Wide Web was precisely what we were trying to PREVENT – ever-breaking links, links going outward only, quotes you can't follow to their origins, no version management, no rights management."

– Ted Nelson

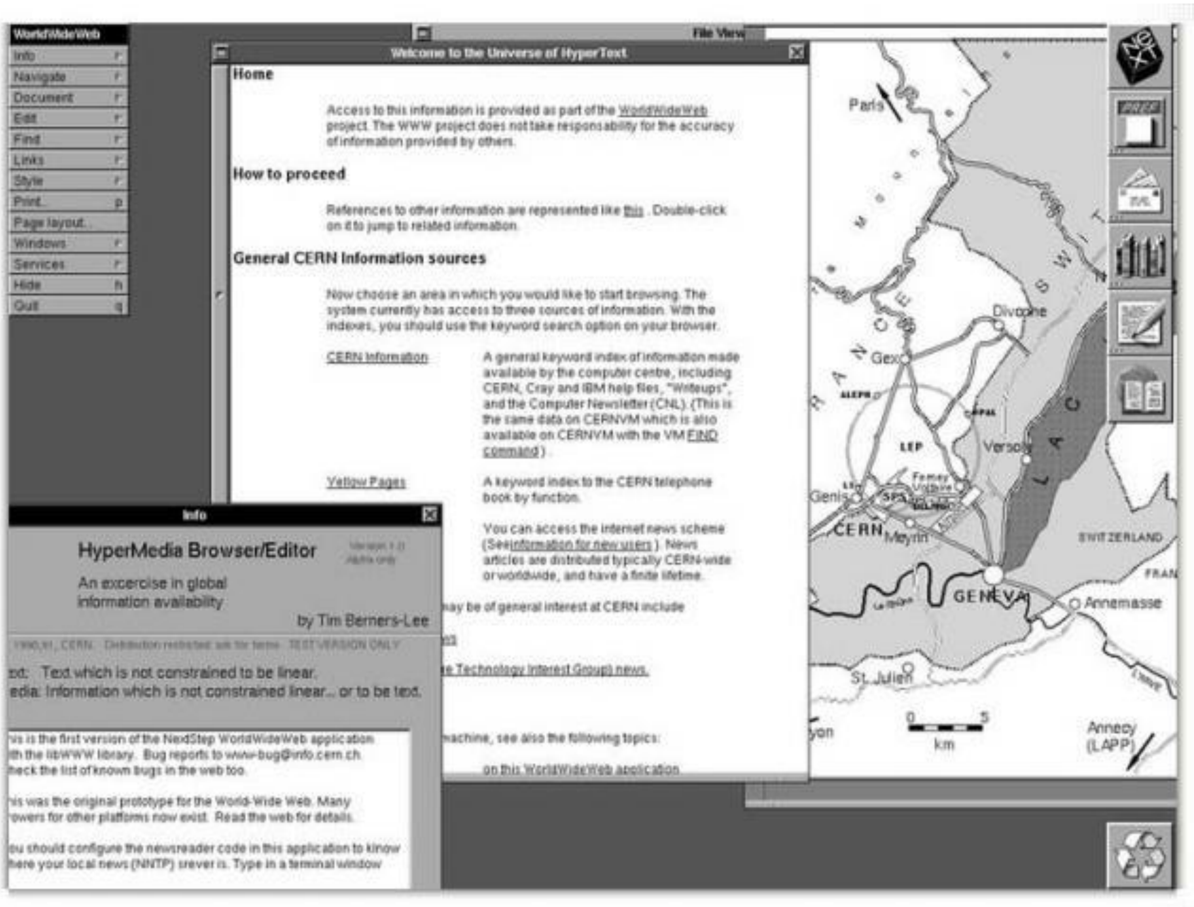
In reality, Tim's primary motivation for creating the Web was because he needed it himself for his work at CERN, a campus where thousands of people crossed paths from different areas of specialization and who were working on different instruments [13]. Tim was convinced that there was a need for a global hypertext system at the CERN research center. In March of 1989, in an effort to persuade the CERN management, Tim wrote a project proposal entitled "*Information Management: A Proposal"* [6]. In this proposal, Tim gave himself the objective of building a space where all of the information or references that were deemed important could be stored as well as the means for subsequently retrieving them [6]. In the absence of a better name, he described a system that he called "Mesh", for which he suggested the use of distributed hypertext to manage the information at CERN, including notably the notion of hotspots (equivalent to the entries of Ted Nelson) that would allow a fragment of text or an icon to be declared as the starting point for an actionable link with the mouse. This proposal was a remarkable exercise in striking a balance between the integration of and the departure from the existing and emerging paradigms of the day.

In his book [4], Robert Cailliau revisits certain important points regarding the general context at CERN that can explain several of the choices made by Tim in his proposal. We can start with several results that existed at the time and that Tim integrated at the heart of the Web architecture.

In 1989, CERN was the largest Internet hub in Europe. In this context, Tim had the idea of extending the link references to the network addresses of documents, in order to weave a "mesh" between documents made available by different machines. Conceptually, he incorporated the principles of hypertext and integrated TCP and DNS into them. More precisely, Tim and his team worked diligently to have RPC (Remote Procedure Call) accepted and deployed at CERN, allowing a program running on a given machine to invoke procedures from programs running on other machines. We can view RPC as the missing link in the distribution of a hypertext where clicking a link becomes, conceptually, a remote call for procedures, making the Web less a network of documents and more a network of procedures. Years later, the dissertation of Roy Fielding [33] would incidentally introduce the REST architecture style (Representational State Transfer) to characterize the distributed hypertext system that makes up the Web. The time that Tim formulated his proposal was also a period in which the Unix operating system was very popular, which natively integrated functions such as support for Internet protocols and multi-user environments.

Another relevant point, concerning the problem of shared documentation between the different disciplines present at CERN, is that there was already within the center a culture in which a programming approach was taken with respect to documents (e.g., LaTeX, SGML), and this had a direct influence on the idea of turning to a simplified markup language. Another, relatively poorly known characteristic of the Web at the time of its conception, is that the Web, as the inheritor of hypertext editing systems, was conceptually open to writing: everything in its architecture allowed not only the consultation (HTTP GET [34]) but also the modification of the Web (HTTP PUT, POST, DELETE [34]). The first client prototype imagined by Tim was a WYSIWYG browser called W3 [20]. This editing function was long maintained at Inria through the work of Vincent Quint, Irène Vatton, and their team working on the Thot, Grif and Amaya libraries and editors [17]. Vincent had, incidentally, worked with Louis Pouzin on the Cyclades network before becoming interested in structured and networked documents, and Vincent would later become the Inria representative within the consortium for the standardization of Web architecture (W3C) before handing the torch to me in 2012. But for diverse reasons (security, ease of use, etc.) and because of a lack of editors [20], the most popular browsers and servers did not include this possibility of modifying Web pages. It would only be reintroduced, at a certain level, with the invention of Wikis (WikiWikiWeb) in 1994 by Ward Cunningham [44].

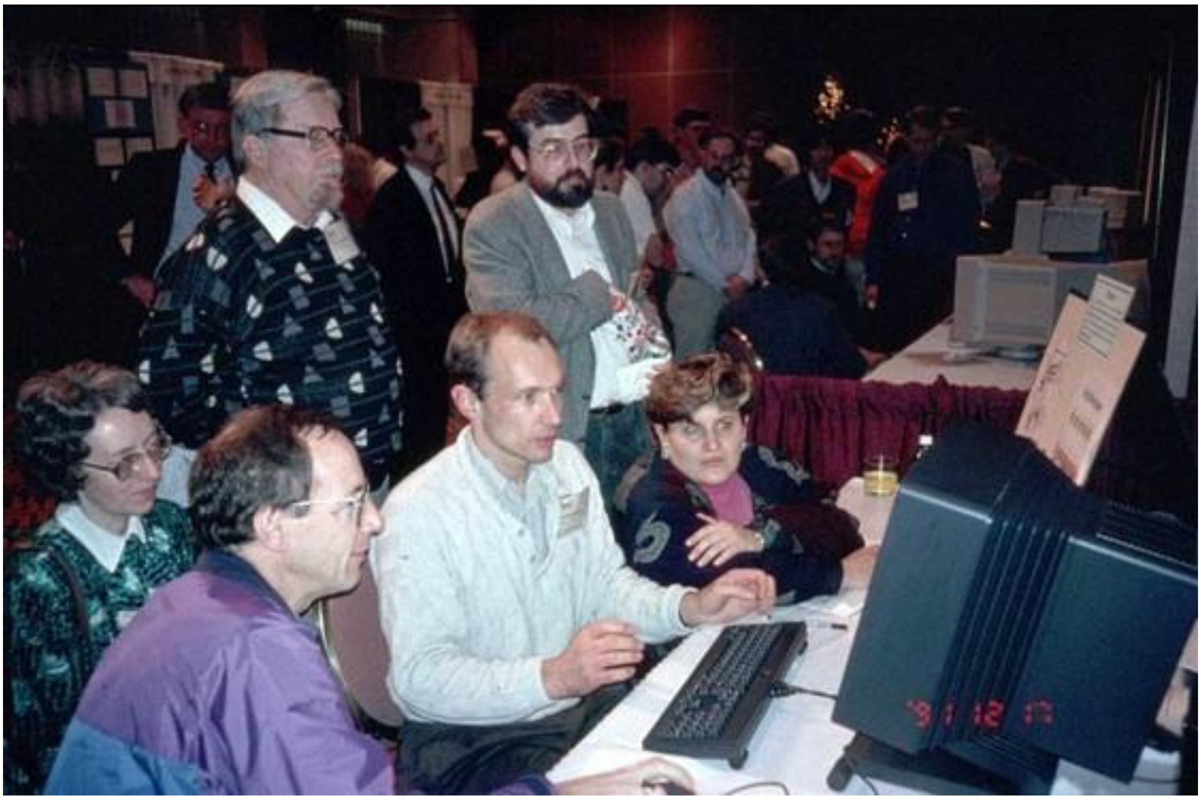
Finally, as Tim's proposal was connected to R&D in computer science, it could not be a CERN project itself as it was considered to be outside the core research scope of the institute. The proposal was thus presented as a project for testing a new computer received at the CERN [13], designed and delivered by the NeXT company, which was led at the time by Steve Jobs, the co-founder of Apple Computer and following his forced resignation. The NeXT Cube was a state-of-the-art workstation and was used by Tim as the first Web server, as well as for the design of the first browser. Its NeXTSTEP operating system became OPENSTEP, and then, following the acquisition of NeXT by Apple in 1996, Apple Rhapsody and finally Mac OS and iOS. It was also a powerful object-based and graphical programming environment that would be used for the development of legendary games such as Wolfenstein 3D, Doom, and Quake. For its first developments, then, on Tim's desk sat a NeXT Cube which brought about the convergence of the different influences mentioned above and which allowed him to build on the experience of hypertexts, the advent of the Internet, the large-scale networking of computers (TCP/IP and Internet), high level programming languages, progress on human-machine interactions, and notably graphical and multi-window interfaces. This context would be an accelerator for the development of the first server and the first Web browser, built on top of a platform providing a large number of high level construction elements. In this way, at the end of 1990, the first server and the first browser were tested through an Internet connection. The browser was called World Wide Web, which would become the name of the hypertext that it would give rise to and which would send the name "Mesh" to the dustbin of history.



The world's first Web site went on line at CERN on December 20, 1990.

But, as noted above, Tim Berners-Lee also chose to break with a certain number of the characteristics of the existing solutions at the time. One reason is that he supported the idea that it was necessary to work with a universal information system, in which generality and portability were more important than other extensions.

The proposal was designed to cover both public and private resources and their links. The solution needed to allow random associations between arbitrary objects as well as an incremental and trivial contribution for objects and for links by different contributors, justifying once again a decentralized as opposed to a hierarchical approach [5]. From the beginning, it was also a question of moving beyond the silos imposed by the existing applications for note-taking, publishing, editing, documentation, help, etc., and searching for independence with respect to the field, task, hardware, and operating system [13]. We can still see this particularity of the Web in its choice of having a weak pairing between the server and the client and in their being independent. Indeed, regardless of the browser used and of the server interrogated, the communication works as long as the standards are respected. The browser even masks the use of different protocols (HTTP, FTP, etc.) [20]. We note that we can see the same preoccupation with breaking out of silos in the current efforts of Tim Berners-Lee for the decentralization of the Web. The diversity of the platforms used at CERN led not only to a search for independence with respect to operating systems, but also for isolating from file paths and systems [5].



Tim Berners-Lee demonstrating the World Wide Web at the 1991 Hypertext Conference, San Antonio, Texas.

In Tim’s proposal, hypertext has been liberated from the central server: the data and links are decentralized on the Internet. In addition, the links are unidirectional and are no longer necessarily maintained: the 404 error was born. This rupture with respect to the fundamental aspects of hypertext is what Ted Nelson was referring to in the citation at the beginning of this section, and it also explains why Tim’s submission to the third ACM Hypertext conference in 1991 was only accepted as a demo. However, these breaks from what had previously existed were the necessary conditions for allowing passage to the scale and virality of the Web. In addition, a certain number of the changes that came with the Web concerning hypertexts led to the creation of major Web applications and services. For example, the absence of a central index and of bidirectional links would lead to the creation of directories (e.g., Yahoo! in 1994) and search engines (e.g., Altavista, Google) to locate pages and links. And, looking at it closely, pages such as those generated by search engine results now make up a substantial part of the Web we browse.

Overall, we can point to three fundamental notions that were, from the beginning, at the heart of the Web architecture:

— The first, and the most important, is that of identifiers. The major prerequisite for weaving a web is to have anchors. The notion of identification goes through UDIs, URLs, URIs, and IRIs [13]. URLs (*Uniform Resource Locator*) and URIs (*Universal Resource Identifier*) are formats for addresses and identifiers allowing the localization, or simply the naming and referencing on the Web, of any type of resource. When an identifier (URI) provides a path for accessing a representation of the resource on the Web, then it becomes a URL. Today everyone knows about these addresses, even if at the beginning they were not intended to be directly manipulated by Web users. For example, “https://www.inria.fr/” is the URL for the Inria home page. URIs are considered universal in the sense that they must allow the identification of all members of a universal collection of network addresses, as long as that the protocol has a notion of resources. It is thus a resource-oriented vision of the network [20].

— The second fundamental notion of Web architecture is that of communication and data transfer. For this, the HTTP protocol allows a request to be made, starting with a URL address, for a representation of the resource identified and localized by the URL, and then the data corresponding to the representation is either obtained or error codes are produced indicating an encountered problem, such as the famous 404 error indicating that the requested resource has not been found.

— The third fundamental notion is that of the representation of exchanged data. As the Web was initially motivated by the representation and exchange of documents, HTML was the first language proposed to represent, store, and communicate Web pages. Starting in 1990, the Web documentation was itself in HTML and the Web began self-documenting, opening up the possibility for everyone to add to this documentation, and to learn to "weave through weaving." This type of reflexivity made the Web a space designed for humans and machines to co-evolve and collaborate. It supports co-understanding, co-documentation, and the co-conception of all the objects that are part of it, starting with the Web itself.

Finally, we point out that while URIs, HTTP, and HTML are inventions, the way in which they are combined in a simple and elegant architecture is even more of an invention [13], and this has greatly contributed to the virality of the Web.

Vague, but exciting: the beginning of virality

"*Build small, but viral.*" – Tim Berners-Lee

It is with this formulation, "*vague, but exciting,*" that Tim Berners-Lee's hypertext proposal Lee was received by his supervisor, Mike Sendall [35]. At a time when many scientists are questioning the effectiveness of research management through calls for projects, this reply, succinct but accepting a degree of risk-taking, can make one dream.

As noted above, at the end of 1990, the first server and the first browser were tested through an Internet connection. On August 6, 1991, Tim posted a summary of the World Wide Web project on several newsgroups, including the forum dedicated to hypertexts, alt.hypertext. This is the moment when the Web project became an application that was publicly available on the Internet [18]. The first Web site went on-line on the same day, at the address http://info.cern.ch. This also served as the source of documentation about the Web itself, the starting point, the self-generating kernel in a sense that can still be found archived on-line [16]. Also in 1991, the first Web server outside of Europe was installed at the Stanford Linear Accelerator Center. From this starting point, the number of Web servers would increase by a factor of 10 each year for the first two years. At the start of 1992, we can identify approximately 10 Web servers, with new browsers appearing during the year (Erwise, ViolaWWW, MidasWWW, Samba for Macintosh, etc.). In 1993, the CERN management officially announced that the Web technology would be made royalty-free and with an open license [18]; at the start of the same year, there were around 50 servers. Two new browsers appeared (Lynx, Cello, Arena), with the most important of them being Mosaic, which was available under Unix, Windows, and Mac OS. It allowed, in particular, images to be visualized directly in the page text. With the Mosaic browser, the Web would truly expand throughout the world, leaving its ancestors Gopher, WAIS, and FTP behind. After Mosaic there would be, in the family tree of browsers, Netscape, then Mozilla, and then, finally, FireFox. In 1994, more than 600 servers were on-line. The next year, more than 10,000 Web servers were available, and Microsoft launched its Internet Explorer, which would be the browser under Windows, with which it was distributed. 1995 would also witness the birth of JavaScript, supported by Netscape. In 1996, the threshold of 100,000 servers was crossed, and in 1998 that of one million. At the start of the 2000s there were more than 26 million, and in 2004 well over 46 million.

Another indicator of the Web's virality was provided by Tim in a 1994 article, in which he noted that the rate of access to Web documentation on the info.cern.ch server doubled every four months between April 1991 and April 1994 [20]. This virality could be explained by several choices made with the design and creation of the Web. In addition to the architectural choices described above, several elements were decisive in allowing and maintaining the virality of the Web. Some of these were technical, starting with Tim's decision to stress the importance of generality, portability, and extensibility, and his position that they were more important than the satisfaction derived from taking advantage of the latest capabilities of computers (such as graphics). Tim could have integrated more complex technologies or aimed to incorporate multiple supplementary functions, but if the Web today has made a major, sustainable technical contribution to the computer science community, it is in particular due to its simplicity, elegance, and extensibility [2].

In his proposal, Tim also stressed that the result must be sufficiently attractive in use to allow the information contained to surpass a critical threshold, and that this mass of information would in turn encourage ever more use and contributions. To achieve this, he proposed that, from the beginning, there should be a link between existing and new databases [6], and that the Web be made systematically compatible with what was already in existence. From its creation, the Web integrated gateway servers to import legacy resources and to allow access to other applications using techniques such as CGI (Common Gateway Interface). In this way, the Web would integrate, inter-operate, and ultimately absorb existing systems, most notably WAIS and Gopher. This approach facilitated the complete migration of communities of users of older systems to the Web. Tim also designed retro-compatibility or descending compatibility with earlier protocols (FTP, Gopher, WAIS, etc.) as an inter-operability constraint, a show of flexibility, and above all as an assurance of its evolutive nature for the future [5, 20]. Further, with approaches like the CGIs, the automatic generation of dynamic pages and the possibility of referencing and reading them immediately played a key role in the weaving of a Web that was sufficiently resourced (numerous hubs), connected, and dense (numerous links) [20].

Tim thus recognized that simplicity was a necessary condition for the generalized adoption of the Web, particularly among the scientific community that he was targeting. His simplification of protocols, including his insistence on the absence of states in the HTTP protocol, made the design easy to implement. In addition, the use of human-readable scripts made the system easy to debug, as well as virally reusable by copy-paste-adapt [2]. Tim systematically took inspiration from what already existed to encourage adoption of the Web. For example, HTTP was inspired by SMTP and NNTP, and the headers used in HTTP exchanges were an extension of MIMEs (Multipurpose Internet Mail Extensions) in order to open the door to the integration of hypermedia, email, news, etc. [20]. With respect to URIs, we can point to the integration of DNSs and slashes (/) for their structure [13]. Similarly, the deliberate proximity of HTML to SGML encouraged the adoption of HTML by the documentation community [5]. Finally, with respect to interaction, the coupling of hypermedia with textual forms proved itself to be sufficiently simple and powerful to satisfy numerous needs, while at the same time remaining easy to use and to implement [20].

Beyond these technical choices, other choices made for economic, legal or social reasons also made a difference in the acceptance of the Web. First of all, the architecture and the technological foundations of the Web were open source, free of rights, and without any fees. There as well, the historical context was important: the Free Software Foundation (FSF), which was created by Richard Stallman in 1985 to promote open software, was at the origin of the GNU project and of GPL licenses which, as of 1989, would set the legal conditions for the distribution of open software. The beginning of the 1990s would also witness the negotiations around and the passage of Unix (BSD) to open source. By making the Web free and with an open code, the CERN literally donated the Web to the world. In particular, in 1992, Tim Berners-Lee and Jean-François Groff worked on a version in C that would lead to the open source creation of the famous software library, libwww, which would subsequently be used in numerous implementations [20]. Tim himself never sought to monetize his work, and advocated from the beginning a Web that was open source and without rights. This important turn of events would allow the viral penetration of Web technologies in all organizations and in their applications.

Another decisive initiative was the establishment, in 1994, of the World Wide Web Consortium (W3C), which would play a primordial role in the normalization of the evolution of the Web architecture, allowing it to grow without losing the standard inter-operability that gave it its universality. Because CERN was required to dedicate all of its resources towards the construction of the LHC (*Large Hadron Collider*), at the end of 1994 it announced that it could no longer continue to invest in the Web project (referred to as the WebCore project). With the support of the European Commission, CERN's Web activity was transferred to the W3C, with the founding members: MIT in the United States, Inria in France, and the University of Keio in Japan. Jean-François Abramatic, who was the director of industrial relations at Inria, played a crucial role in this transfer and went on to become, soon afterwards, the president of the W3C. Subsequently, Inria handed over the role of European host of the W3C to ERCIM (European Research Consortium for Informatics and Mathematics), which the institute had helped to create with its European partners in 1989. Before the W3C, Web standards had been published in the form of RFCs (Request For Comments). They would thereafter be published as recommendations from the W3C. The consortium has never stopped seeking to improve its activities since its creation, refining its mission [21], its organization [23], its way of functioning [24], and its ethics policy [25]. The work of the W3C has been, and still is, decisive in allowing the Web to overcome major crises, whether technical (for example, the war between browsers and incompatibilities), commercial (for example, proprietary extensions), political (for example, the PICS standard to respond to concerns over inappropriate content), etc. The W3C rapidly opened working groups to address different evolutions and aspects of the Web, such as accessibility or internationalization. Today, numerous standardization or simply discussion groups exist covering a broad range of topics.

Evolutions of the Web: weaving a global web of resources

"*When I took office, only high energy physicists*
had ever heard of what is called the World Wide Web...
Now even my cat has its own page."

– Bill Clinton, 1996

Tim revisited the Web architecture in 1996 [5] and slightly changed the three pillars we mentioned in the previous section, putting an accent on addressing, protocols, and content negotiation. This content negotiation is a native mechanism of the HTTP protocol, offering the possibility of proposing, for a given URI, different versions of the same resource, and was directly inspired by the format negotiation mechanism of "System 33" from Steve Putz at Xerox PARC. The HTML language was no longer considered to be one of the available formats. Therefore, starting in 1994, Tim himself underlined that HTTP was perhaps misnamed because it was not so much limited to HTML transfer as destined to efficiently exchange arbitrary data in the context of links and jumps in distributed hypermedia [20]. In a certain way, the content negotiation mechanism downgrades HTML by allowing the negotiation of all format types from the beginning. URIs permit the identification of all types of resources. HTML therefore returned to just being a prerequisite for a browser [20]. It would first permit a uniform format to be provided for hypertextual documents and the documentarization of the network of resources that became the Web, before evolving towards a programming language for Web applications. Inversely, the Web showed its independence from any particular models or structures of data.

Already in 1994, Tim noted the need to have the Web evolve towards more real time functionality (e.g., videoconferencing, virtual reality) and envisaged supporting commerce on the Web, in particular for online payments [20]. Tim also identified, in this article from 1996 [5], various directions in which the Web could evolve that he considered important and which would open up numerous directions for research and development which are still relevant today. He cited as a first area of work the infrastructure and performance of the Web, as well as the Web as not only a social space but also one for machines. He even suggested that this aspect, which we can now refer to as hybrid communities [36], was vital and he explained that machines on the Web would be necessary for to scale human interactions. In the article he also recalled one of the initial, and still unachieved objectives: the possibility of easily and securely linking the different scales and spheres of private and public documents, going from personal information systems (PIMs) to global discussion systems, and offering group-based tools at all levels (personal, organizational, public) while maintaining the capacity to link through these levels. Tim also noted with regret that in 1996 the Web still lacked a security infrastructure, and it is true that little progress had been made since the arrival of SSL and HTTPS with Netscape in 1994. It was even necessary to wait for the RFC2818 in 2000 for the first formal, official specification of HTTPS. Finally, Tim stressed again the importance of systematically seeking greater levels of decentralization: seeking openness in the writing and contribution of content, in externalizing links, externalizing annotations, and externalizing verification.

The PICS standardization, which was one of the first recommendations of the W3C in 1996, allowed the filtering of inappropriate content, in particular for children. It was also an example of the type of decentralization that was desired: the filters capturing user preferences were created and stored in the browsers; the descriptors were stored in the consulted sites, but were generated by third-party authorities. Incidentally, in doing this PICS opened up the idea of the Web architecture moving systematically towards a generic approach to specific problems, which could then be re-used in other usage scenarios. Thus, the idea of labeling by, and in reference to, a third-party site leads more generically to breaking the binary (browser/page)-(server/site) link to move towards more complex links, and also to avoid limiting the labeling to problems regarding content acceptability. This therefore represented an opening towards the general notion of metadata on the Web, and was in line with another evolution towards a Web of documents and structured data.

As part of this evolution, CSS was an important step that marked the beginning of the separation of content from the way in which it is presented on the Web (CSS vs. HTML). Style sheets allowed the formatting of a document to be made independent of and separated from its structure, and also to use the same formatting for multiple documents, or inversely to vary the formatting of a single document. This step allowed the multiplication and diversification of content and its presentation in independent and creative ways. CSS also enabled the Web to present content in professional, highly advanced ways that could ultimately rival the types of production allowed by the best editing tools for electronic documents. Here as well, the notion of style sheets was present in Tim's first prototype, although he did not publish this feature nor the syntax that he used. This capacity was therefore basically lost in the first browsers to be developed, such as Mosaic. Håkon Wium Lie and Bert Bos were the first to work on the standardization of style sheets, leading to the CSS level 1 recommendations in 1996. This success was also achieved thanks to Thomas Reardon and Chris Wilson of Microsoft, who already in 1995 provided assurances that they would support CSS. And indeed, Microsoft Internet Explorer 3 included the first implementation of CSS, even when its specification was still only in draft form. The history of CSS [45] is also one of the first examples of an activity that would require the development of suites of dedicated tests and demonstrators (for example, CSS Zen Garden) to accelerate the implementation and adoption of a recommendation and evolution in the Web architecture.

With Web content now separated from its format, the content was able to evolve through the creation and management of its own data and document structures. This was achieved with the standardization of XML in 1998 and, in its wake, of several languages allowing its validation (DTD, XML Schemas), interrogation (XPATH, XQuery), transformation (XSLT), management (XProc), etc. On a personal note, I particularly recall from this time the roadmap for the Semantic Web that Tim published in 1998 [37]. This roadmap followed from his presentation of 1994 [19] and his article from the same year, in which he expressed a preference for an evolution in Web resources, which were at the time essentially documents destined for humans, towards resources with semantics that were more oriented towards machines to allow for more automated processing [20]. The 1998 roadmap [37] would open the door to all of the work that was carried out on the Web of Data and the Semantic Web (RDF, RDFS, SPARQL, OWL, etc.). In France, Rose Dieng-Kuntz, Olivier Corby, and their team at Inria immediately identified this evolution and initiated research on the subject [10, 11]. This roadmap was also decisive for me, as I was starting my Ph.D. work in their lab in 1999 on the coupling of software architecture for distributed artificial intelligence (multi-agent models) with formal models of distributed knowledge (Semantic Web models) [12]. Today, Tim is still a passionate advocate of evolutions in the Web such as the Semantic Web [2].

In parallel with this progress with respect to content, this period saw the birth of Web programming and the Web applications that would become an important pillar of the Web and its usage. At the beginning, at the birth of the Web, the principle was very simple: a page would be generated on request following a click on a simple link, or in response to a form being submitted. With each interaction, the page would be entirely reloaded. On the server, code written in C, C++, Shell, or other languages would use the CGI method (Common Gateway Interface) to process requests and generate pages in return. With the introduction of frames in around 1996, it became possible to divide the page into frames and, as a result, only reload part of the page that is being displayed. The same year, JavaScript appeared, and in Web applications it became possible to use, even conjointly, programming on the side of the server as well as on the side of the browser. This marked the beginning of a genealogy of programming languages and techniques for the Web, some of them more server-oriented (ASP, PHP, C#, Python, Ruby, Perl, JSP, etc.), others more client-oriented (e.g., Plugins, ActiveX, CSS+HTML), and some that could be used by both (Java Servlet and Applet, JavaScript). Finally, with the *XMLHttpRequest* component, which was proposed by Microsoft in 1998 and then added to JavaScript and rapidly adopted by the majority of browsers, it became possible to exchange data between a page and its server, and also, thanks to its DOM, to modify a displayed page without necessarily reloading it. This technique would be named AJAX in 2005 and was massively adopted in Web applications to conjugate code being executed on the client with code being executed on the server, allowing fluid interactions with the user. In parallel, the architecture of the Web was being increasingly studied and formalized, for example with the Ph.D. work of Roy Fielding, who introduced the REST architecture to characterize it [34]. In the early 2000s, the proposition of evolving towards a Web of services (SOAP and WSDL standards) opened up a new direction for work, to use the Web as a programming platform, which is currently being realized through APIs and associated languages such as JSON. Some go so far as to talk of the Web as an operating system that goes beyond the global collection of services that are offered on the Internet, and that is independent of the computers and individual objects that are connected to it. They see the Web as the programming and execution environment *par excellence* for Internet applications. Within this movement, one upcoming evolution that is currently under study is to make the Web a universal application platform for connected objects, called the *Web of Things* [38]. Today, we literally envisage making a web of everything. But we can also remember the initial influence that RPCs had on the conception of distributed hypertext and see this evolution as a form of returning to the source. In addition, in representing the Web as a mesh of potential calls for procedures that we invoke with each link that we follow, we understand better why ambitions such as that of crawling, indexing, or archiving the Web are complicated or even paradoxical.

Another parallel trend was that, as of 1996, companies such as Nokia in Finland were interested in providing Web access to mobile phones. In 1999, it was NTT DoCoCo in Japan, with its i-mode. The same year, the QR code, created by the Japanese company Denso-Wave in 1994, was made open source, which would contribute to its distribution and its being made available in recognition applications on mobile phones, in particular for gathering URLs displayed around us. In the same period, WAP (*Wireless Application Protocol*) and WML (*Wireless Markup Language*) were proposed for adapting Web access and content to the constraints of mobile phones and their connectivity issues. They would later be abandoned, when phones and networks attained levels of performance that allowed them to directly access the classic Web and Internet. These first attempts, as well as those made with PDAs with wireless network cards, would have to wait until the mid-2000s and the arrival of smartphones to have platforms upon which they could reach their full potential. In 2004, the W3C would launch its Mobile Web Initiative (MWI). In 2005, the first ".mobi" domain name was proposed but was criticized, in particular by Tim, as a solution that detracted from the independence of the Web vis-à-vis terminals. In 2007, in an invited presentation at the GSM conference, Tim took a position against private domains created by proprietary platforms and defended the open platform offered by the Web [46]. The problem was first to overcome the limitations of mobile connections (screen, bandwidth, limited interactions, limited computational power, connection costs, etc.). Then, either by resolving these problems (for example, through compression) or through their disappearance (for example, with increases in the power of the terminals and networks), they gradually passed from relating to simple mobile access to the question of the deeper adaption of usage under mobile conditions (for example, geo-localization, adaptation of interactions and interfaces, access to personal data, contextualization, audio and vocal interaction, augmented reality, coupling of several terminals, etc.). Today, mobile Web applications are common, and many native mobile applications have in fact largely been developed using the languages and standards of the Web. In addition, lower entry-level prices and the democratization of smartphones mean that in certain areas of the world, mobile phones are now the primary means that people have of accessing the Web and are the most widespread way of having an initial contact with the Web [39].

With the Web of Data, Web of Applications, Web of services, Mobile Web, and a Web that is accessible, international, etc., the Web therefore initiated, very early, its transformation towards a generic hypermedia architecture for programming and interaction, and especially the generalization of a Web potentially linking all types of resources, whether computational or not. The Web can reach anything, as everything can be identified with an URI. As the principles of the Web are extensible and generic, they have allowed us to pass from a documentary vision of a global library to a network of protean resources. One of the major strengths of the Web is in its universality, but we will also see that it requires constant vigilance to be preserved.

Categories and variations: a multi-faceted web

"As soon as you externalize an idea,
you see facets of it that weren't clear
when it was just floating around in your head."
-Brian Eno

When speaking of the high points in the evolution of the Web, we see references now to what is called Web 1.0, Web 2.0, Web 3.0, etc., terms that I'm not particularly fond of because they give the impression of corresponding to major software evolutions in the Web whereas they are more changes in practice or even in the understanding that we have of the Web, and they do not take into account the multiple directions in which the Web has evolved in parallel. Web 1.0 corresponds essentially to the initial document-based and distributed hypermedia vision of the Web. Web 2.0, which is also called the Social Web, reflects both the opening up of the Web to writing, including the AJAX approach for interactions, and the enormous social exchanges that it permitted, with Wikis, blogs, forums, social media, etc., and the corresponding impact that we are all familiar with. Web 3.0 generally covers the integration of Semantic Web practices and the Web of Data into Web 2.0, such as with RDFa in the Facebook OGP protocol or the use of Schema.org in numerous sites integrating social functionalities (for example, votes, ratings, etc.). However, as we saw in the preceding section, more than representing evolutionary leaps from one version to another, the Web is constantly undergoing a number of concurrent evolutions that require constant and considerable work by the W3C to keep them compatible, but that at the same time, in an evolutionary approach, allow it to send out multiple probes in search of upcoming changes and the intersections among them.

Another set of aspects of the Web are: the surface Web, the deep Web, and the dark Web. The surface Web is the Web that is indexed and covered by search engines (through services called crawlers). This Web is public and makes up the most emerged part of the Web. It also includes numerous gateways and applications that open onto the deep Web, with modes of access, pathways, and dedicated search engines. The deep Web is often unfairly called hidden or invisible, as in fact we see it every day. It involves the part of the Web that is essentially dynamically generated like search engine result pages, and that is indeed visible and accessible but that is not indexed by engines for different reasons: because the links or content is dynamically generated, because the content is only available following authentication or another type of complex interaction going beyond the simple following of a link, because the content is not connected to the rest of the Web, because the content is covered by a policy of being excluded from indexing robots (robots.txt), etc. This part, called the deep Web, in fact represents the largest proportion of the content of the Web.

The term *deep Web* is sometimes wrongly confused with the terms *darknet* and *dark-web,* which in fact have different meanings. The dark Web, with its association with dark matter, is a shadow Web that is used for activities for which users seek anonymity for whatever reason, for example a political opposition figure seeking to avoid oppression in his or her own country, or a criminal organization seeking to use the tools available on the Web for their illicit activities. The threads of this Web are deliberately disconnected from the classic Web so as to not be indexed and found. These threads are woven on the darknet, which uses techniques for anonymization, cryptography, networking (e.g. peer-to-peer), and security to mask the identities of and exchanges between users. By combining these techniques with the Web architecture, it is possible to weave and browse (e.g., with the TOR browser) hidden webs. The terms *dark Web* and *darknet* are often used interchangeably, even though the notion of *darknets*, referring to deliberately isolated networks, already existed in the Arpanet era in the 1970s, and therefore well before the Web, and their applications extend to other services as well (email, IRCs, forums, etc.).

Finally, we can also mention the IntraWebs (the use of the Web as an intranet) which, behind VPNs and company firewalls, use Web solutions within organizations to create webs reserved for their members. There as well, security techniques are used to control access to the components of the specific Web. It is not necessarily sought to avoid links between these webs and the public Web, but access is made to be secure.

The developments described in the preceding section and the facets that were just mentioned nevertheless all come from the same, single architecture: the standardized architecture of the Web. As insisted on in the W3C's mission statement (*One Web*), it is not a question of having different Webs but of having different facets of one and the same Web, with one and the same architecture [21]. Within the W3C itself, there is a special group called the TAG (Technical Architecture Group) [22], which is led notably by Tim in his role as the Director of W3C and which ensures, in particular, that all of these changes in the Web remain compatible and continue to respect and be combined within a coherent and universal architecture. These different examples of the evolutions of and the facets of the Web nevertheless lead us to a new need: that of developing the means for studying the Web and its changes.

The unknown aspects of the Web

> *"Web science – what makes the Web what it is, how it evolves and will evolve, what are the scenarios that could kill it or change it in ways that would be detrimental to its use."*
>
> *-Dame Wendy Hall*

Even today, it is striking to see to what point the Web is both well-known and at the same time poorly understood, as we can see for example with the tenacious and all-too-often seen confusion between the terms Web and Internet. Despite the fact that their respective inventors have received distinct Turing awards, in 2004 and 2016, respectively, for two very different inventions, the terms Internet and Web are still frequently used interchangeably. Let us state it again: the Internet allows for the inter-connection between computer networks and connected objects in general. It provides a communication infrastructure that supports numerous applications such as: email, telephoning, and videoconferencing; and the Web is the distributed hypermedia that has become the primary software architecture for applications on the Internet.

Beyond this confusion, the term Web is also often used in a way that fails to differentiate between the founding principles of the software architecture and the object that emerged from it, i.e. the web weaved by its billions of users. Even in 1994, Tim noted that the term World Wide Web had rapidly evolved to cover multiple things, including, on the one hand, the architecture, and on the other hand the set of data made available on the Internet using that architecture [20]. The architecture and the object that emerged from it have two histories that are linked but that cover different aspects. Each of these two aspects exhibits, in different ways, complexity that calls for both research and development.

The architecture of the Web is based on protocols, models, languages and algorithms that need to be specified, designed, characterized, and validated, as well as on, systematically, constraints such as those related to upscaling, efficiency in time and memory, interoperability, and internationalization. To achieve this, the architecture of the Web and its extensions are the subjects of research, in particular in the digital sciences. Within computer science and the digital sciences -- and we will see in the following section that this is true in other fields as well -- the Web has expanded both as a new tool for work but also as a new subject bringing with it both solutions and new problems and needs.

With respect to the object that has emerged from the use of that architecture, the complexity of its usages, the heterogeneity and volume of its content, services, and data, the dynamic nature of some of its traffic, and the life cycles of its resources and communities, are equally sources of complexity that also call for a scientific approach and for theoretical, applied, experimental, and multi-disciplinary research. One proposal, among others, illustrating this idea of studying the Web as a complex object is, for example, to have Web observatories [40], complete with the observational instruments and methods of the experimental sciences, to study the Web; to develop, in a sense, microscopes and telescopes for studying the Web and the scientific method to turn these instruments towards the web.

In addition, the more the Web grows in its architectural complexity and in the resources linked to it, the more it calls for transdisciplinary research and development [41]. Tim would even go so far as to say that, ultimately, the Web is more of a social creation than a technical creation [3].

Historically, the Web quickly became a subject of research. Again under the impetus of Robert Cailliau and Tim Berners-Lee, in 1994 a WWW conference series called “The Web Conference,” was started at CERN and became the annual meeting place for the research, development, companies, and major actors of the Web. Having accepted on two occasions, in 2012 and 2018, to be general co-chair of this conference, I can attest to the high level of research that the Web elicits. And this research community has grown and diversified, for example with resolutely multidisciplinary initiatives such as “Web Science” or more specialized conferences such as ISWC, WI, WebIST, etc. To conclude this section, we can note that in 2006 Tim became the Director of the Web Science Research Initiative (WSRI), a scientific research program on the social and technical aspects that underlie the evolution of the Web [14], and that he is also a founder of the Web Science Trust [15], a British non-profit organization dedicated to the interdisciplinary study of the Web and its effects on society [2].

Central portion of the final panel at the 1994 WWW conference. From left to right: Dr. Joseph Hardin, Robert Cailliau, Tim Berners-Lee, and Dan Connolly.

Shifting to a Web-Wide World

"We can only see a short distance ahead,
but we can see plenty there that needs to be done."
-Alan Turing

"*This is for everyone*," it is with this message that Sir Tim Berners-Lee presented the Web during the opening ceremony of the 2012 Summer Olympic Games in London. This now-famous tweet inspired the deliberately ambiguous title of this article ("For everything"), as now the Web is for and affects everyone, is becoming part of everything around us, and is being deployed in all corners of the world. The Web, for everyone, everywhere, and for everything. This Web, with its worldwide reach, a name which at the beginning could have been seen as immodest, revealed itself in just a few years to be a self-fulfilling prophecy, where its being designed to be universal allowed it in fact to evolve towards the universal.

The Web has modified our relationship with time and space by giving us the possibility of interacting with distant people and things, by giving us access to information that is not locally available, by recording traces of our actions, by documenting a variety of activities, and by allowing us to revisit them and re-experience them *a posteriori* and remotely. We no longer note down the address of the dentist, we can find it on the Web. We no longer program our VCR or DVR, we look for a recording on-line. We no longer go to the train station to buy tickets, but rather download an electronic ticket. The omnipresence and hypermnesia of the Web are so established now that we can't stand it when it's not there to respond to us, and the "no results found" page of search engines is virtually never seen anymore.

The Web is also an outstanding tool for integration and interoperability, a space for exchange between applications. The text that you are now reading was partly written through the Web interface of a collaborative text editing application, and partially dictated through a mobile phone connected to the same application through the Web. The spelling and grammar were largely corrected in real time using tools that have been statistically trained on the Web. To connect to such resources, to integrate them, synchronize them and ensure their accessibility regardless of the application or terminal used, all of this is ensured by the standardization of the Web.

At the historical scale of computer science, the Web has not only uniquely proven itself as an architecture that has passed the test of time, it has defined a new era: there is an era before and an era after the Web. An era where the problem was to have information about or access to a service, and an era where the challenge was to find one's self within the overwhelming quantity of information. An era of fragmentation, and then an era of hyper-integration, even over-integration, with the risks that that entails.

The Web has brought about a certain number of new, now everyday things, such as: Web servers, Web browsers, Web pages, personal pages, Web sites, Web addresses, and Web applications. Some of them, like "Web sites," remain only informally defined. The Web has also created hybrids of previously existing things and activities, such as: web documentaries, web series, webisodes, web tv, web radio, webinars, web sourcing, web publishing, web ID, web mail, web commerce, web advertisement, webcams, webcasts, weblogs, web conferences, web journals, web OS, etc. What's most frightening about these terms is that my spellchecker already recognizes them. From the moment a Web resource is programmed as a type of object, as soon as these objects form hybrids and come into contact with the principles and practices of the Web, they become new objects calling for new practices. And the phenomenon is fractal: when a Web technology is developed, while remaining as close as possible to its architecture and universal principles (for example, wiki), it acquires virality and the potential to develop in different forms (for example, wikipedia, wiktionary, wiki travel, wikileaks, etc.). Once the Web left behind its initial, pupal form with the metaphor of a universal library, the Web as a sharing tool is now so commonplace that it provides metaphors for other fields, such as for example the mycorrhizal networks called the *Wood-Wide Web* by some biologists [9].

As can be seen through the references and allusions made in this article, Tim continues to champion the Web and to invest in it without keeping score, to help bring the Web to its full potential[2]. This Turing award rewards not only the fact that he invented the Web, but also the fact that he has worked his entire life defending it. Because the defense of the Web remains necessary. It is universally useful and used, but it remains fragile, and its initial ideals could prove to be just a passing phase if we do not constantly keep watch on their preservation.

Tim Berners-Lee is still fighting today against all forms of re-centralization, for example the application-based centralization supported by monopolies of certain companies for certain Web services, and for the neutrality of networks in general and the Web in particular. The stakes (neutrality, decentralization, democratization, etc.), the dangers (re-centralization, levels of access, a Web at different speeds, etc.), limitations (infrastructure needs, energy, costs, etc.) make the Web a never-ending project more than a final achievement. Tim's greatest fear is still that a political or commercial entity takes control of the Web, which would mean, for him, its death sentence. This is what motivates his engagement, for example, in support of Net Neutrality, of the re-decentralization of the Web, of the shattering of filter bubbles [42], and the re-appropriation of data by users [13]. The architecture of the Web is and must remain robust, even in a hostile environment, neutral, even if its underlying layers are compromised, and resilient, even if the infrastructure is lacking or limited, etc. To protect it, we will have to design the spiders that weave the Web architecture to be extremophile animals.

Inversely, the Web can itself be perceived as a danger. In 1996, Tim wrote about how the diversifying force provided by geography could be weakened by the Web [5]. He did this at a time when, as I was giving my first classes on the Web to my classmates in Mathematical Engineering at the INSA in Rouen, one could hear concerns being expressed in France about the Web becoming a tool for the hegemony of the English language and of English-speaking culture. In the same year, Tim drew attention in a general way to the ethical and societal stakes of the Web: he recalled the impact that the choices made concerning Web architecture had on the forms of society in which we were living; the necessity of revisiting the notion of copyright in a space where copying could take on so many forms (for example, placing copies of content in the cache, and the legal status of those copies); the privacy-related problems introduced by the multiple opportunities to capture data; the impact that Web-based information has on a voting population; the necessity of working hand-in-hand with legislative systems; etc. [5].

In my opinion, these subjects raised in 1996 foreshadowed needs that would be even more urgent two decades later, and the need to actively search for interdisciplinarity in the study of the Web and of its evolutions. We all recognize that the World Wide Web has had an enormous impact on society [2]. The Web is a subject of research, development, economic and commercial activity (digital giants), a vector for social action and structuring, a political tool and subject, a new legal object and space, and even an area of philosophical questioning [8]. It is thus important to understand the Web holistically, in all its complexity, and to encourage the

“Web Science” movement towards transdisciplinarity. For me, the three W’s of the World Wide Web call for the three M’s of a Massive Multidisciplinary Method [41].

Further, the Web must also absolutely become a subject for education and training in itself. Its use (basic principles of browsing, searching, etc.), best practices (critical reading, cross-referenced validation, active contributions, etc.), prevention (protection of privacy, protection of children, etc.), are all subjects that every generation should be taught in school as an important element for equal opportunity. In 2008, Tim became one of the founders of the World Wide Web Foundation, a non-profit organization that promotes access to the Web for all [13] and an open Web as a public good and a fundamental right [2].

The universality of Tim’s approach, and the power to link everything imaginable and to do this everywhere in the world, were difficult to comprehend at the beginning, and few people saw the difference with respect to already existing hypertextual systems. Whereas at the beginning, before the Web arrived, the problem was to imagine a world with the Web in it, we are now in the inverse situation where people have forgotten or can no longer imagine that there was a world without the Web [13]. We are at a tipping point where we talk of augmented reality on the Web [43] and where the Web has developed to the point that our perception has inverted and it’s the world that seems to be inscribed within the Web. This idea is, for me, summarized in the inversion of the terms *Web-Wide World*, where the Web has surpassed the 3D physical world and the latter is entering the Web to be extended and augmented in an open number of dimensions. And in this perspective, the echoes of Tim’s phrase that our choices in Web architecture have an impact on our societies only increases with every evolution of the Web.

The Turing award has thus rewarded Sir Tim Berners-Lee not only for inventing the Web in 1989, but also for the fact that, ever since, Tim has never stopped extending the Web to cover an ever increasing number of domains.

“*The Web as I envisaged it, we have not seen it yet.*
The future is still so much bigger than the past.”

– Sir Tim Berners-Lee, WWW Conference 2009

Fabien Gandon, Sir Tim Berners-Lee, Louis Pouzin, Jean-François Abramatic, at the 25 years of the Web at the *Futur en Seine*, Paris, 2014.

Acknowledgements. I would like to thank my colleagues and friends from the W3C, the SIF, Inria, Université Côte d'Azur, CNRS, and the I3S for their feedback on this article.

[1] W3C Automotive Working Group Charter, accessed 10 may 2017, https://www.w3.org/2014/automotive/charter.html
[2] ACM Turing Award Sir Tim Berners-Lee, http://amturing.acm.org/award_winners/berners-lee_8087960.cfm
[3] Weaving the Web : The Original Design and Ultimate Destiny of theWorldWideWeb by Its Inventor, Tim Berners-Lee, HarperCollins, 1999, ISBN 0062515861, 9780062515865
[4] How the Web was Born : The Story of the World Wide Web, James Gillies, Robert Cailliau, January 15, 2000, Oxford University Press, ISBN-10 : 0192862073 ISBN-13 : 978-0192862075
[5] The World Wide Web : Past, Present and Future, Tim Berners-Lee, August 1996, https://www.w3.org/People/Berners-Lee/1996/ppf.html
[6] "Information Management : A Proposal" (March 1989), the original proposal for the software project at CERN that became the World Wide Web. https://www.w3.org/History/1989/proposal.html
[7] Sergey Brin, Lawrence Page, The Anatomy of a Large-Scale HypertextualWeb Search Engine. Computer Networks 30(1-7) : 107-117 (1998)
[8] Philosophical Engineering and Ownerhip of URIs, Tim Berners-Lee, an interview in "Philosophical Engineering", collected set of writings from the PhiloWeb workshops, https://www.w3.org/DesignIssues/PhilosophicalEngineering.html
[9] La Vie secrète des arbres, Peter Wohlleben, Editions Les Arènes, 2017, 2352045932
[10] Cédric Hébert, Modèle de traitement de RDF basé sur les graphes conceptuels, Rapport de stage de DEA, I3S, université de Nice Sophia-Antipolis, 1999.
[11] Olivier Corby, Rose Dieng, Cédric Hébert. A Conceptual Graph Model for W3C Resource Description Framework. International Conference on Conceptual Structures, Aug 2000, Darmstadt, Germany. pp.468 - 482, 2000
[12] Fabien Gandon, Distributed Artificial Intelligence And Knowledge Management : Ontologies And Multi-Agent Systems For A Corporate Semantic Web, Thèse, Université Nice Sophia Antipolis, 2002.
[13] Neil Savage, Weaving the Web, Communications of the ACM, Vol. 60 No. 6, Pages 20-22, 10.1145/3077334, https://cacm.acm.org/magazines/2017/6/217732-weaving-the-web/
[14] MIT and University of Southampton launch World Wide Web research collaboration Initiative will analyze and shape web's evolution, MIT News, November 2, 2006, http://news.mit.edu/2006/wsri
[15] Web Science Trust (WST), accessed 7th june 2017, http://www.webscience.org/
[16] "WorldWideWeb—Archive of world's first website".WorldWideWeb Consortium. https://www.w3.org/History/19921103-hypertext/hypertext/WWW/TheProject.html
[17] Welcome to Amaya, https://www.w3.org/Amaya/
[18] https://timeline.web.cern.ch/timelines/The-birth-of-the-World-Wide-Web
[19] Tim Berners-Lee, Plenary Talk extracted slides, First WWW Conference, Geneva 94, https://www.w3.org/Talks/WWW94Tim/
[20] Berners-Lee, Tim and Cailliau, Robert and Luotonen, Ari and Nielsen, Henrik Frystyk and Secret,Arthur, The World-Wide Web, Commun. ACM, Aug. 1994, Vol. 37, no 8, 0001-0782, p. 76–82, 10.1145/179606.179671, ACM, New York, NY, USA
[21] W3C mission, accédé le 16 juin 2017, https://www.w3.org/Consortium/mission
[22] The W3C Technical Architecture Group (TAG), accédé le 16 juin 2017, https://www.w3.org/2001/tag/
[23] W3C organisation, accédé le 16 juin 2017, http://www.w3.org/Consortium/facts
[24] W3C Process, accessed 16 june 2017, http://www.w3.org/Consortium/Process/
[25] W3C Ethics, accédé le 16 juin 2017, http://www.w3.org/Consortium/cepc/
[26] Luis Von Ahn, Manuel Blum, Nicholas J. Hopper, and John Langford. 2003. CAPTCHA : using hard AI problems for security. In Proceedings of the 22nd international conference on Theory and applications of cryptographic techniques (EUROCRYPT'03), Eli Biham (Ed.). Springer-Verlag, Berlin, Heidelberg, 294-311.
[27] Von Ahn, Luis, et al. "recaptcha : Human-based character recognition via web security measures." Science 321.5895 (2008) : 1465-1468.
[28] Bush, Vannevar. "As we may think." The atlantic monthly 176.1 (1945) : 101-108.
29] T. H. Nelson. 1965. Complex information processing : a file structure for the complex, the changing and the indeterminate. In Proceedings of the 1965 20th national conference (ACM '65), Lewis Winner (Ed.). ACM, New York, NY, USA, 84-100. http://dx.doi.org/10.1145/800197.806036

[30] C. Engelbart, and William K. English, AFIPS Conference Proceedings of the 1968 Fall Joint Computer Conference, San Francisco, CA, December 1968, Vol. 33, pp. 395-410 (AUGMENT,3954,).
[31] Pouzin, L., Presentation and major design aspects of the Cyclades Computer Network. IN : Proc. 3rd Data Communications Symposium, Tampa, Fla., Nov. 1973. pp 80-85.
[32] Cerf, V. G. and Kahn, R. E., A protocol for packet network intercommunication. IEEE Trans. Commun., Vol. COM-22, No. 5, 637-48 (May 1974).
[33] Roy T. Fielding and Richard N. Taylor. 2000. Principled design of the modern Web architecture. In Proceedings of the 22nd international conference on Software engineering (ICSE ’00). ACM, New York, NY, USA, 407-416. http://dx.doi.org/10.1145/337180.337228
[34] R. Fielding, J. Reschke, Hypertext Transfer Protocol (HTTP/1.1) : Semantics and Content, RFC7231, Internet Engineering Task Force (IETF) , 2014
[35] Tim Berners-Lee’s proposal, CERN, March 1989, http://info.cern.ch/Proposal.html
[36] Fabien Gandon, Michel Buffa, Elena Cabrio, Olivier Corby, Catherine Faron-Zucker, et al.. Challenges in Bridging Social Semantics and Formal Semantics on the Web. Hammoudi, S. and Cordeiro, J. and Maciaszek, L.A. and Filipe, J. 5h International Conference, ICEIS 2013, Jul 2013, Angers, France. Springer, 190, pp.3-15, 2014, Lecture Notes in Business Information Processing. <hal-01059273>
[37] Berners-Lee, Tim, "Semantic web road map”, https://www.w3.org/DesignIssues/Semantic.html (1998)
[38] W3C Begins Standards Work on Web of Things to Reduce IoT Fragmentation, https://www.w3.org/WoT/
[39] The Mobile Web, Worl Wide Web Foundation, http://webfoundation.org/about/vision/the-mobile-web/
[40] Tiropanis, Thanassis, et al. "The web science observatory." IEEE Intelligent Systems 28.2 (2013) : 100-104.
[41] Fabien Gandon. The three ’W’ of the World Wide Web callfor the three ’M’of a Massively Multidisciplinary Methodology. Valérie Monfort ; Karl-Heinz Krempels. 10th International Conference, WEBIST 2014, Apr 2014, Barcelona, Spain. Springer International Publishing, 226, Web Information Systems and Technologies. <http ://www.springer.com/fr/book/9783319270296>. <10.1007/978-3-319-27030-2>. <hal-01223236>
[42] Pariser, Eli. The filter bubble : What the Internet is hiding from you. Penguin UK, 2011.
[43] Fabien Gandon, Alain Giboin. Paving the WAI : Defining Web-Augmented Interactions. Web Science 2017 (WebSci17), Jun 2017, Troy, NY, United States. pp.381 - 382, 2017, <http ://websci17.org/>. <10.1038/nbt0609-508>. <hal-01560180>
[44] Michel Buffa, Du Web aux wikis : une histoire des outils collaboratifs, Interstice 23/05/2008, https://interstices.info/jcms/c_37151/du-web-aux-wikis-une-histoire-des-outils-collaboratifs
[45] A brief history of CSS until 2016, Bert Bos, Style Activity Lead, W3C,17 December 2016, https://www.w3.org/Style/CSS20/history.html
[46] The Mobile Web, Tim Berners-Lee, keynote, GSM congress, Barcelone, 2007