

Chapter 1: The evolution of the Web

1. Introduction

The Web has transformed a lot since it began in the 1990s. It's a fascinating story of change and growth. This change has had a big impact on how we talk to each other, share information, and use the Internet.

In the beginning, the Web was simple, like a basic tool. But as time passed, it turned into something amazing. This happened because people had new ideas, created impressive websites, and because we wanted more from the Internet.

As the Web continued to evolve, it had some important moments. One significant moment was when Web 2.0 emerged. It made the Internet more exciting and interactive. We started sharing our own content online, like photos and posts, and social media became a big part of it.

Mobile phones made it even better. Now, we can access the Internet on our phones from anywhere. Plus, advancements in technology, like artificial intelligence, made the Internet more interesting and more personalized for each of us.

In this course, we will explore how all of these changes happened. We will take a closer look at the key moments in the Web's history and see how technology, culture, and challenges shaped it into what it is today. From the early Web with static information to today's dynamic and interactive Internet, it's a story of global connection and how it keeps reshaping our world.

2. Definition of the Web

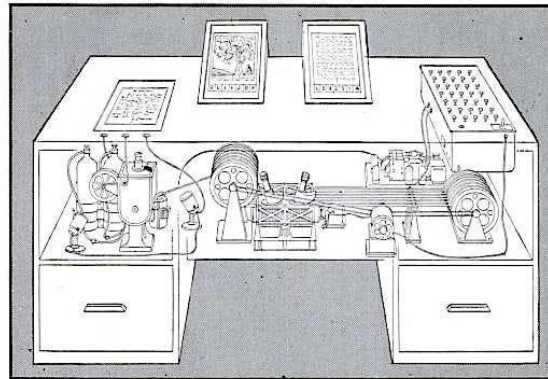
The World Wide Web (WWW), commonly referred to as the Web, is a global system of interconnected hypertext documents and multimedia content that is accessed via the Internet using web browsers. It allows users to view, navigate, and interact with a wide range of information, including text, images, videos, and links, all of which are hosted on websites and web servers distributed across the world. The Web is a fundamental component of the Internet and provides a user-friendly interface for accessing and sharing information and resources online.

The Web is one of the applications of the Internet, distinct from other applications like email, video conferencing, and file sharing. In fact, the Internet is the global infrastructure that connects millions of computers worldwide, while the Web is a specific application that uses this infrastructure to provide publishing, sharing, and access to information in the form of web pages. The Web is one of the many applications that run on the Internet.

3. The beginning of the Web

Before discussing the Web, it is necessary to first talk about a certain sequence of events that gave birth to the Web. Starting with **Vannevar Bush**, who is one of the pioneers of the Web, through his article "As We May Think," published in 1945 in The Atlantic Monthly magazine, in

which he predicted the invention of **hypertext**. The article describes a system called Memex (Figure 1 Memex system , which can store books and link them together for easy retrieval. It already discusses the concepts of links and navigation, drawing inspiration from the associative functioning of the human brain (Gandon).



MEMEX in the form of a desk would instantly bring files and material on any subject to the operator's fingertips. Slanting translucent viewing screens magnify supermicro-film filed by code numbers. At left is a mechanism which automatically photographs longhand notes, pictures and letters, then files them in the desk for future reference.

AS WE MAY THINK CONTINUED

Figure 1 Memex system (Memex: hypothetical proto-hypertext system)

Twenty years later, **Ted Nelson** drew inspiration from the Memex and proposed a structure that allowed documents to be divided into parts and linked together. He called this structure **hypertext**¹, with the idea of navigating between text excerpts from different documents (Figure 2 The data structure proposed by Ted Nelson . Later, he expanded the concept of hypertext to **hypermedia**, which allows navigation between text, images, audio, video, etc.

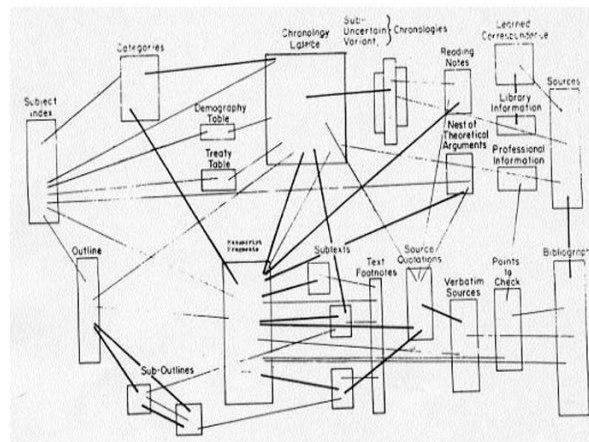


Figure 2 The data structure proposed by Ted Nelson (Nelson, 1965)

¹ text that contains links or references to other documents or parts of the same document. These links, often referred to as "**hyperlinks**," allow users to navigate between different pieces of information.

In 1989, building upon these ideas, **Tim Berners-Lee** decided to take a step further by suggesting that hypertext links should not be restricted to a single machine and that documents should be distributed across the network. The concept was to navigate from one document to another that might not be stored on the same machine. He proposed a system of linking documents that wasn't limited to a single machine and could span the entire network. This led to the invention of the World Wide Web, as detailed in an article titled 'Information Management: A Proposal' (Figure 3). The linking system proposed by Tim Berners-Lee.

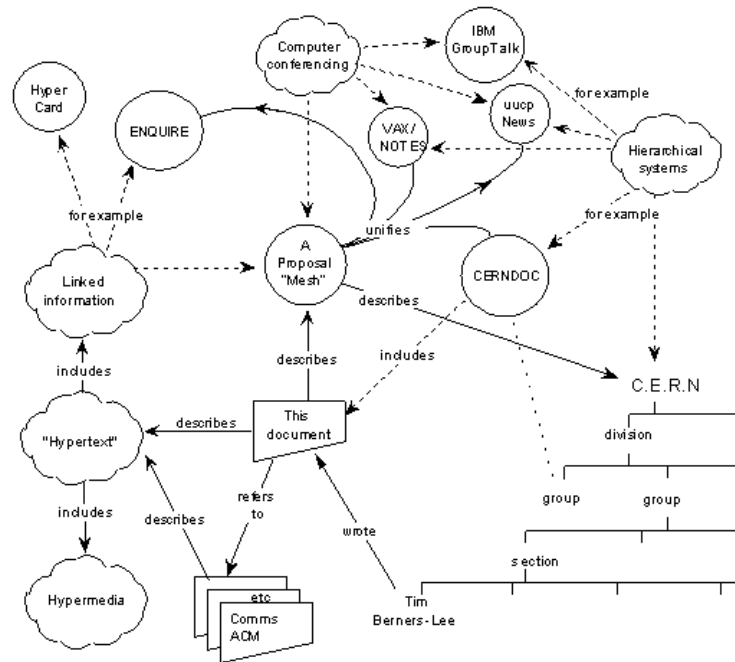


Figure 3 The linking system proposed by Tim Berners-Lee (Berners-Lee, March 1989, May 1990)

4. Web 1.0

- Period: 1990s
- Objective: Distribute information

Referred to as the "classic," "traditional," or "documentary" web, the first-generation Web is characterized by static pages containing minimal hypertext or hypermedia content. These pages are rarely updated and operate primarily in "read-only" mode. Typically unidirectional, websites of this era often take the form of catalogs designed for businesses or organizations, sometimes transactional, targeting individuals. They allow publishing information accessible to all users at any time and establishing an online presence. Email and discussion forums are the two communication modes associated with this Web.

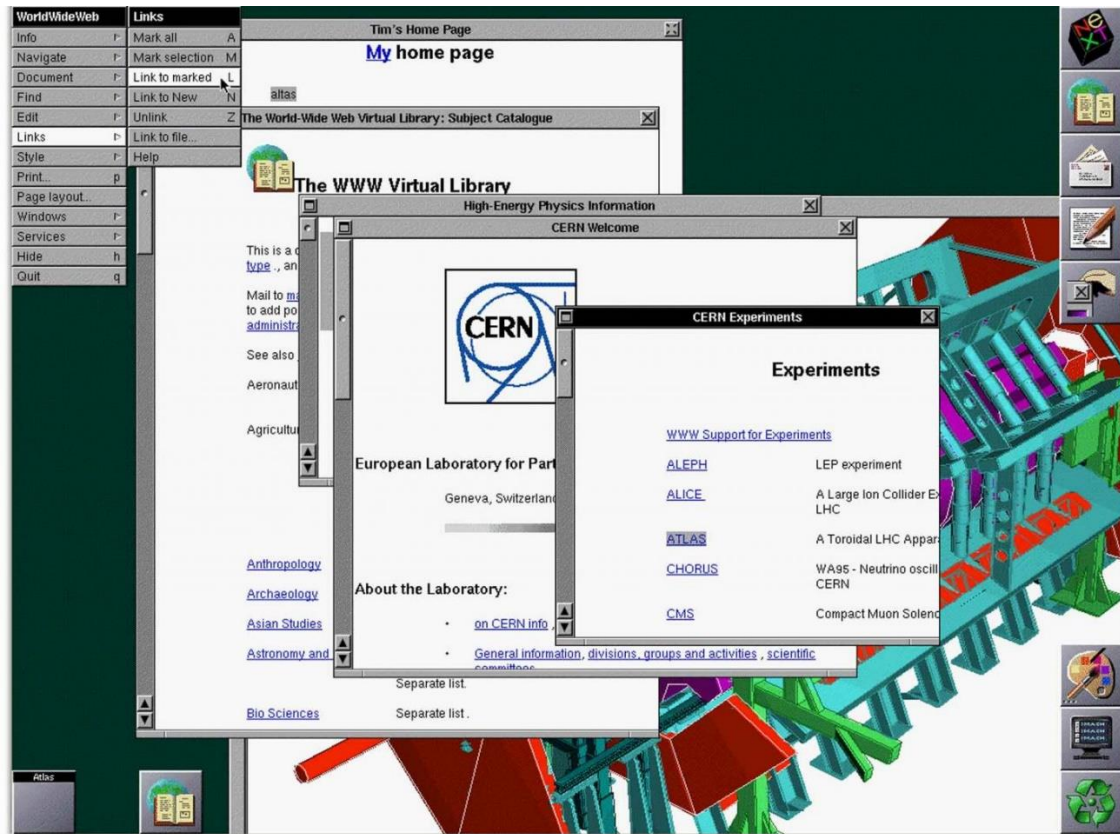


Figure 4 Screenshot of the original NeXT Web browser in 1993. (Les 20 ans d'un web libre et gratuit, s.d.)

5. Web 2.0

- Period: 2000s
- Objective: Connect users and share content

The lack of interaction possibilities among users led to the development of Web 2.0. This era marked a certain democratization of the web, as users became both participants and creators. In addition to offering more possibilities than the previous version while being easier to use, this Web allows users not only to extract information but also to share, modify, and create content. Platforms are interactive, applications are participatory, websites are dynamic, and content is unlimited. This Web, accessible to both professional and amateur users, connects them worldwide and becomes a space for socialization as users form communities and social networks emerge.

The advent of the version 2.0 of the Web notably brought about blogs and microblogs, encyclopedias and collaborative portals, content syndication² and citizen journalism. It also marked the explosion of online commerce and the rise of centralized applications — the web

² Content syndication is a process in which content, typically articles, news, blog posts, or other types of information, is distributed and shared across multiple websites or online platforms. This allows content creators to disseminate their articles or updates to a broader audience without users having to visit the original website.

giants that have become Facebook, Google, Amazon, and Microsoft, commonly referred to as the acronym GAFAM. These companies gather data from their vast user base, which is subsequently monetized through targeted advertising (Meilleur, 2021).

6. Web 3.0

- Period: 2010s
- Objective: Make sense of data, connect knowledge, and guide users based on their context

The Web has provided widespread access to an overwhelming amount of information while allowing the collection of vast user data, leading to the creation of what we call "big data" due to its colossal volume. In this era of information overload, it became imperative—and possible, thanks to artificial intelligence (AI)—to integrate technologies into the Web that facilitate the sharing and reuse of content among users (humans), between "machines" (programs or software agents), and between users and machines.

The use of metadata (data about data) has become essential for describing and categorizing big data, hence the use of the term "semantic" (related to meaning and significance) to describe this Web 3.0. This semantic approach organizes web resources so that machines can better respond to user requests. Think, for example, of using keywords for searching and promoting content, individual filters for online shopping, or visual search that provides information on a topic from an image. In summary, the 3.0 generation of the Web aims to give meaning to data, connect knowledge, and offer a more personalized experience to the user, who is now more mobile than ever (Meilleur, 2021).

7. Web 4.0

- Period: Since 2020
- Objective: Connect the real and virtual, facilitate interactions between users and objects, promote collective intelligence, and innovation.

Web 4.0 exists in a context where artificial intelligence (AI), machine learning, the Internet of Things, and augmented and virtual reality, among other technologies, are becoming increasingly sophisticated and integrated into our daily lives. It's just the beginning of a world where the digital and physical realms merge. This fourth-generation "intelligent" web will tend to become more direct, invisible, omnipresent because it will be in symbiosis with the connected objects in the user's environment. These objects and this web will better understand natural language (our language), analyze user behavior to meet their needs, sometimes even without their intervention and without the need for them to go through a digital screen. This web is therefore a key element in what is called the virtualization of the world or the "phygital" revolution (for the fusion of the physical and the digital), in which humans and computers will interact with increasing fluidity. Here are some examples of how the concepts associated with Web 4.0 are starting to manifest in

various technologies and applications: Smart Assistants (Siri), IoT and Smart Cities, Autonomous Vehicles, Natural Language Processing (ChatGPT), etc.

8. Architectural principles

A client-server architecture represents the environment in which client machine applications communicate with server-type machine applications. In this context, a web browser from a client requests the content of a web page from a web server, which then sends back the result (response).

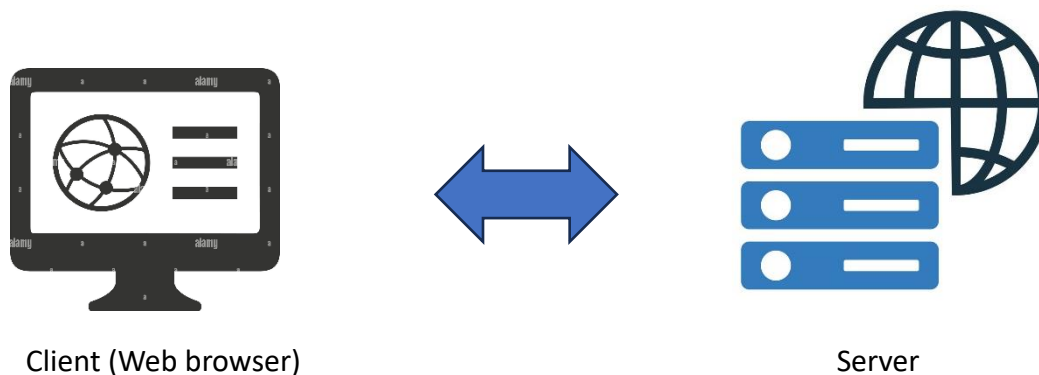


Figure 5 Client-server architecture

This architecture is based on three main concepts:

1. The first concept is that of addressing, identification, or location through web addresses or **URLs (Uniform Resource Locators)**. An example of a web address that corresponds to the university's website <https://www.univ-constantine2.dz/>.
2. The second concept is the protocol, which means how your browser will query a remote server and receive a representation of the requested page. This protocol is called **HTTP (Hypertext Transfer Protocol)**.
3. The third concept is the representation language, specifically the language used to represent web pages, generally **HTML (Hypertext Markup Language)**.

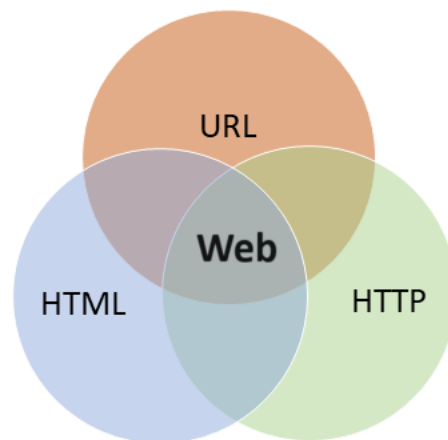


Figure 6 Components of the Web architecture

In this course, we will focus on understanding how the representations that flow through the Web can extend beyond documents and web pages. Specifically, we will explore how these representations can represent everything in the world.

9. The standardization of the Web

Since the creation of the web, the number of servers and web pages has grown very rapidly. On the other hand, in the 1980s, many web browsers were developed, and each browser extended the standards in an incompatible way and offered its own extensions. In other words, you could visit a website from one browser but not from another. This put the web at risk because there was no longer interoperability.

To address this issue, the World Wide Web Consortium (W3C) was created in 1994 and led by Tim Berners-Lee. The goal of the W3C is to harmonize the developments of the Web and standardize them in a way that ensures interoperability. The standardization process at W3C follows well-defined steps.

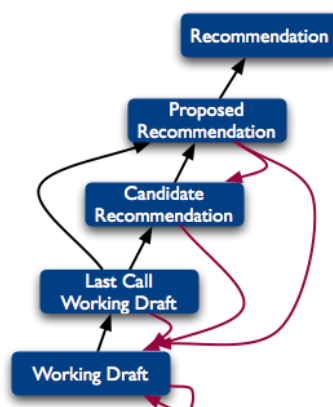


Figure 7 W3C Process Flow (W3C Process Flow, 2013)

The first step is the creation of a workshop to determine if there is enough interest in the topic. Following this workshop, if there is sufficient interest, a working group is formed, which begins working on a draft of the standard: the working draft.

When this draft is stable enough, it proceeds to the second step, the "last call". For several months, the public can provide feedback, identify errors, and suggest corrections. If there are major issues, the document returns to the draft state; otherwise, it moves to the next step called "candidate recommendation".

In this new step, the focus is on verifying that the standard can be implemented. For example, in the case of a new version of HTML, it is checked whether browsers can transition to this new version. Of course, if it's not feasible, it returns to the draft state; otherwise, it proceeds to the next step, which is the "proposed recommendation".

In this step, the standard is reviewed by the W3C Director and a specific group within W3C called the Technical Architecture Group (TAG). They ensure, among other things, that the standard is compatible with other standards because the goal is to maintain a unified architecture of the Web. If that's the case, the standard moves to a final step called "recommendation". A recommendation is the term used at W3C to refer to a standard (Gandon).

10.Conclusion

In conclusion, this chapter has provided a comprehensive overview of the evolution of the web, highlighting key milestones and technological advancements that have shaped its development. As we move forward, the next chapter will delve into the exciting realm of the 'Web of Data,' exploring how linked data and semantic web technologies have transformed the way information is interconnected and shared on the internet.

References

- Berners-Lee, T. (March 1989, May 1990). Information Management: A Proposal. *CERN*.
- Gandon, F. (n.d.). *Web sémantique et Web de données*. Retrieved from France Université Numérique Mooc: <https://lms.fun-mooc.fr/>
- Les 20 ans d'un web libre et gratuit*. (n.d.). Retrieved from <https://home.cern/fr/news/news/birth-web/twenty-years-free-open-web>.
- Meilleur, C. (2021). <https://knowledgeone.ca/du-web-1-0-au-4-0/?lang=fr>. Retrieved from <https://knowledgeone.ca>.
- Memex: hypothetical proto-hypertext system*. (n.d.). Retrieved September 2023, from Medium: <https://medium.com/@quartopiano/memex-hypothetical-proto-hypertext-system-bc0c60564eb8>
- Nelson, T. H. (1965). A File Structure for the Complex, the Changing and the Indeterminate. *Association for Computing Machinery: Proceedings of the 20th National Conference*, 84–100.
- W3C Process Flow*. (2013). Retrieved from <https://www.w3.org/>: [https://www.w3.org/2013/dd-epasorg.htm#\(8\)](https://www.w3.org/2013/dd-epasorg.htm#(8))