THE DIGITAL HUMANIST

A CRITICAL INQUIRY

DOMENICO FIORMONTE
TERESA NUMERICO
FRANCESCA TOMASI

TRANSLATED FROM THE ITALIAN BY
DESMOND SCHMIDT & CHRISTOPHER FERGUSON

The Digital Humanist: A Critical Inquiry

First published in 2015 by
punctum books
Brooklyn, New York
http://punctumbooks.com
ISBN-13: 978-0692580448
ISBN-10: 0692580441

Library of Congress Cataloging-in-Publication Data are available from the Library of Congress.

Cover image: Masaccio, *Young Man* (~1420), National Gallery, Washington DC
Cover design: Chris Piuma
Typographic design by Vincent W.J. van Gerven Oei

Domenico Fiormonte,
Teresa Numerico &
Francesca Tomasi

THE DIGITAL HUMANIST

A CRITICAL INQUIRY

*Translated from the Italian by
Desmond Schmidt with
Christopher Ferguson*

MMXV
punctum books
Brooklyn

Table of Contents

PART I

THE SOCIO-HISTORICAL ROOTS

Chapter 1
Technology and the humanities: A history of interaction

The history of modern computing is not simply the history of one particular technology, but is part of a larger history of culture and knowledge, which itself implies changes of perspective, epistemic loops and the emergence, at the end of a complex process, of certain devices at the expense of others.

There are several ways to look at the development of technology, and at least two possible ways to look at the history of computing. One is to focus on the hardware, on the engineering development of the machine, and the other is to study the evolution of machine languages and their logic, and hence the history of software development. Each approach has a limited perspective, and each interprets events of the past in the light of the present. Rather than simply describe the history of these unique and extraordinary machines, this chapter will instead try to pursue a metaphoric, epistemic and perhaps even socio-political path: a way of looking at machines that were meant to simplify, represent, and control the production of information in ways never before imagined. It will try to look at the development of the electronic programmable machine (what we normally call a *computer*) as a project for the production of meaning destined to deeply influence other disciplines and our culture as a whole.

The objective of all this is to investigate if and how the humanistic paradigm — in its broadest sense seen as a relation between philosophical anthropology and the development of human–machine interaction — may have influenced the emergence of information science. A strong case can be made that it did, in several ways. The emergence of a certain humanistic perspective can be traced from the end of the Second World War: the idea that machines should not only be able to solve equations, but could also provide a simple and unequivocal answer to any problem of representation, organization, or enhancement of knowledge.

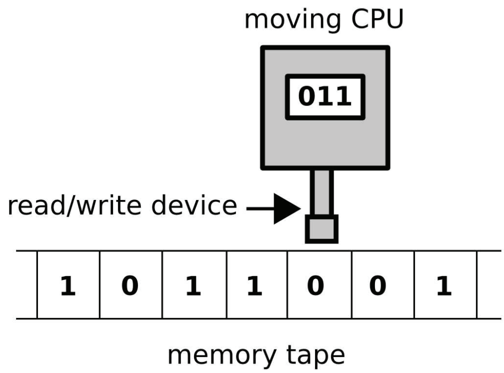

Figure 1.1. Turing Machine

1.1 From Alan Turing to the modern computer

This section discusses the origins of the theoretical model behind the computer, and how it influenced the design of the first programmable electronic machines. The aim is to show the profound historic relevance of certain events, which turned out to have a massive impact on the future development of technology. While the influence of Turing's theoretical machine is widely accepted (especially in historical accounts that focus on software), the historical contingency underlying his model is generally overlooked. This work of reconstruction aims to highlight the happenstance of the story, and to strip away from the computer the fate-like aura of inevitability that often surrounds it.

Information science has not always existed in its current form. According to tradition, it emerged during the Second World War, when the need to carry out large numbers of calculations prompted researchers from various countries to devise machines that could do exactly that. However, these early computers relied heavily on the logical-theoretical model devised by Turing in his studies on logic back in the 1930s. Turing had invented a theoretical machine that was able to solve any task whose execution could be automated (Turing 1937). Initially, the idea of the computer was not born as a theoretical model. In fact, it stemmed from the need to demonstrate the undecidability of logic (Petzold 2008). Turing designed the most powerful theoretical device ever imagined by man, but this almost omnipotent device came with an original sin: it was unable to solve one class of problems, which included the "decision problem" or *Entscheidungsproblem*. The Turing Machine (TM) was extremely simple: it had a two-dimensional tape, divided into squares and as long as needed (although not infinite). It also had a head for reading, writing, erasing and moving along the tape, and a table of instructions representing the task at hand in a precise and unambiguous way (Figure 1.1).

Turing later described his model in the context of his work on the ACE (automatic computing engine):

> Some years ago I was researching on what might now be described as an investigation of the theoretical possibilities and limitations of digital computing machines. ... This type of machine appeared to be sufficiently general. One of my conclusions was that the idea of a "rule of thumb" process and a "machine process" were synonymous. .... Machines such as the ACE may be regarded as practical versions of this same type of machine. There is at least a very close analogy. (Turing 1947/2004, 378–379)

In this passage, written around ten years after his earlier work on the *Entscheidungsproblem*, Turing sought to define the relationship between his theoretical machine and practical computers: the Turing Machine was the study of the "theoretical possibilities and limitations of digital computing machines." John von Neumann certainly became aware of Turing's work during the latter's two-year stay at Princeton (Numerico 2005, 45–48), and he also relied on Turing's constraints and specifications in 1945, when he wrote what is known as the *First Draft,* the document marking the beginning of the history of informatics.

Taken on its own, Turing's 1937 article does not seem particularly promising as the (still unwitting) beginnings of a theoretical computer structure. But computing theory had established the limits within which the discipline could develop (Mahoney 2000, 22), and described the class of problems that cannot be solved by a TM, i.e. those equivalent to the famous *halting problem*.

The TM also produced the awareness that performing calculations was a simple manipulation of symbols, and that there was really no qualitative difference in principle between numbers and formulae.[1] In fact, computer science would develop from the notion of computers as manipulators of symbols and not as "simple" calculators, so laying the foundations for the informatics revolution that came about after the Second World War. The identity of each Turing Machine was represented by a precise table of instructions specific to that machine, and identified by a serial number. In this context, Turing also developed another interesting concept: the concept of the universal machine. This device could emulate any other TM, and later in part inspired the development of the computer. The universal machine was able to replace the others, provided it was supplied with their "programme." Turing's thesis, which equated the TM and effective computability (i.e. the informal notion of definable calculability), was extremely credible. However, it could not be genuinely demonstrated, since there was no demonstration that could connect a formal notion to a semi-formal one such as computability through a TM. One of the arguments Turing used to back his thesis was the identification of the ability of both humans and machines to calculate. In some sense, it was at least a partial realization of a project of the XVII century German philosopher Gottfried Wilhelm Leibniz (1646–1716). The aim of his *Characteristica Universalis* was to find a system that could allow researchers to sit around a table and "calculate" the solution to any proposition.[2] The method, also suggested by McCarthy (2005, 165–169), was a mix of theory and experiment, using a unique technology that could represent propositions as well as calculate their "value." Up until at least the early 1960s, a large section of computer science could relate to Leibniz's project of representation, organization and the production of knowledge. Later in this section Leibniz's influence on artificial intelligence and Semantic Web projects will be discussed in more detail (cf. § 2.4).

The development of the computer resulted from many different requirements and pure coincidences. History did not bear any categorical imperative that necessitated its development or identified it in the digital form it later acquired. In the 1930s the most powerful computing machine was the Differential Analyzer, designed and built by Vannevar Bush (cf. § 1.3), based on an analog computing model. However, due to the interdisciplinary nature of the research and the contact between its various research groups, work on the development of the digital computer certainly contributed to its realization.

Among the fields that most contributed to the development of the digital electronic device, cryptography deserves especial mention for its role in the difficult task of decoding the secret messages of the German military. The protagonists active in this field of research, which had primarily developed in Europe, were mathematicians, engineers, physicists, and logicians, but also experts in linguistics and enigmatography, chess masters, statisticians, etc. The most important research group working on the problem was based in the sleepy London suburb of Bletchley Park. Turing actively participated in the group's mission, marrying his great talents with the diverse and hybrid skills of the colleagues he met there, since this was the best way to ensure success in such a critical situation. In this context what was perhaps the first fully electronic machine was constructed: the Colossus, designed and built in record time between January and December 1943 at Bletchley Park. The realization of that early computer was undoubtedly a complex, collective, and interdisciplinary endeavor, which saw its creators changing their original vision along the way, under the influence of practice.[3]

In this context, it is interesting to note that in the 1940s, when Turing focused on building a device capable of emulating some of the skills of the human brain, he modified the model he had himself created, and had inspired the development of the computer:

> So far we have been considering only discipline. To convert a brain or machine into a universal machine is the extremest form of discipline. Without something of this kind one cannot set up proper communication. But discipline is certainly not enough in itself to produce intelligence. That which is required in addition we call initiative. … Our task is to discover the nature

> of this residue as it occurs in man, and to try and copy it in machines. (Turing 1948/2004, 429)

In other words, Turing realized that, in order to achieve more interesting outputs in terms of machine intelligence, a method was needed that not only fully obeyed rules, but was also capable of exercising initiative. That was because a machine, however universal, might not be enough on its own to carry out intelligent tasks.

Turing's contemporaries understood his prediction that including components other than "discipline" was only needed to obtain intelligent responses from the machine at a later stage in its development. Once they had built devices that were powerful enough, they could reflect more deeply on the possibility of endowing those same machines with abilities comparable to, or even exceeding, human skills.

The next section examines the common history of the analog and digital models, and analyzes the reasons behind the ultimate success of the latter, while keeping in mind the continuing relevance of rhetoric to the orientation of the technological project under investigation.

1.2 What computers cannot do: from analog to digital

Why did the Turing Machine seem so convincing that it became synonymous with Twentieth Century machines, when Turing himself questioned its ability to perform "intelligent" human tasks? According to some technology historians (Eduards 1997), the success of the digital paradigm was due not only to a specific technological development, but also to the power of a metaphor: a machine that could tirelessly control and execute computations and always provide precise and repeatable results. The idea behind this external approach to the history of technology arises from the compatibility of the values represented by the digital machine with the social and political model of a closed society, typical of the Cold War years (Eduards 1997, 66–73). Others have seen the rise of "cyber-science" as a response to the need to reorganize the complexity of science in ways more suited to the current power regime (Keller 1995). This would in turn allow the reformulation of a knowledge-model to guarantee the control — albeit in new ways — of the emerging dynamics of information. And yet, even if it is accepted that the cultural dimension and its "discourse" do not play a central role in the choice of technology, the question remains: how could a machine with a very precise set of constraints and limitations, even before it was built, become the solution to all problems of calculus and "artificial" intelligence? One of the possible sources of innovation in technological development is the exploitation of already available tools. Following this pattern, the development of the computer was based on theoretical studies in computability theory carried out at least a decade earlier, in the mid 1930s. As described above, the theoretical machine developed in that context had a set of specific limitations relating to its ability in solving decision problems. Some problems cannot be solved by a Turing Machine, and they are rather relevant problems too: such as deciding whether or not a formula belongs to the set of theorems of a certain formal system, or again (and this really amounts to the same thing) determining if a certain program will come to an end or get stuck in an infinite loop.

Nevertheless, when moving from the functionality of theoretical to that of real machines, one must also consider other kinds of problems, as von Neumann began to realize by the late 1940s. In his contribution to the famous *Hixon Symposium* (an interdisciplinary meeting held in 1948 to discuss "Cerebral Mechanisms in Behavior") von Neumann introduced a new series of constraints specific to practical machines, as opposed to those valid for purely abstract ones. As a matter of fact, in logic "any finite sequence of correct steps is, as a matter of principle, as good as any other." In the case of automata "this statement must be significantly modified" because "automata are constructed in order to reach certain results in certain pre-assigned durations, or at least in pre-assigned orders of magnitude of duration" (von Neumann 1948/1963, 303). In other words, the recognized father of the computer was introducing the issue of tractability as distinct from the issue of decidability. But what does that mean? For a computer to be unable to solve a certain problem it is not necessary for that problem to be undecidable; it is sufficient for it to be intractable, i.e. the number of steps needed for its solution can be finite, but superior to the physical capabilities of the machine in actually performing the task within a reasonable time. Such intractable problems are in fact rather common: the so-called "traveling salesman problem" (finding the shortest route between a number of cities), school timetables or filling spaces according to certain characteristics. All these problems are known as NP-complete and it is unclear if they can be solved with polynomial (i.e. tractable) algorithms (see Harel 2000 for more details). It is also worth noting that not only are computers unable to solve all these problems for practical purposes, they are also unable to simulate the behaviors of deterministic problems described by precise equations, when the problems are particularly sensitive to the initial conditions. In such cases even a minute variation in those conditions leads to unpredictable results, and hence their behavior cannot be simulated by devices designed to provide unequivocal answers to a certain input (Longo 2009a).

But if things are so complicated, why was the computer so successful as a problem-solving tool in calculation, control, and communication? The answer may lie in the ideology behind the logic that dominated the first half of the twentieth century. In order to provide a foundation to mathematics, logicians set out to model an imaginary agent. This agent would only manage symbols. It operated in complete isolation, and was only able to preserve the truth of its premises in the truth of its result by manipulating other symbols, and by following a set of precise inference rules. In short, formal systems like those used until the 1930s by the Hilbert school always modeled their agent as omniscient and isolated — a vision that aptly represented the perspective of logic in contrast to that of mathematics or other analogous symbolic systems. When forced to defend the validity of his hypothesis, Turing worked on the concept of a machine capable of emulating the absolute computing agent, i.e.

a human being in the isolated and pure act of calculation. Turing even imagined it would be possible to parcel out its operations as a list of basic and accurately describable steps. It would then be possible to represent these steps by means of instruction tables so precise and exhaustive that even a machine could execute them. Obviously, if mathematics is seen purely as the manipulation of symbols according to detailed, strict schemes, whose sole function is to be applied without any ingenuity or intuition (i.e. without any form of individual or collective creativity) then the digital machine becomes the perfect device for representing this kind of behavior, and electronic technology can be used to realize it. However, the reduction of mathematics to the elementary and transcribable steps of calculation takes into account neither the practice of the mathematician nor the limitations of real machines, which, contrary to abstract ones, simply cannot keep the process going for long enough.

Nevertheless, this model provides some very simple, powerful, and convincing answers to the issues of process unification and simplification: the reduction of all operations to elementary ones, a precise and unambiguous instruction list for the computing agent to execute, and the notion that everything that is calculable can be performed by a single machine simply by changing its instructions.

Analog devices were more flexible than digital ones in representing phenomena, but created greater problems in terms of interpretation and repeatability of results. Moreover, with analog devices it was not possible to create a precise and unequivocal instruction table for use with each instance of the same problem. For all these reasons, the analog model did not seem very appealing as a way of simplifying procedures, and digital devices later took over even the control tasks for which analog machines were well suited.

The pioneers of digital technology were well aware that it would be difficult to use their devices without giving them the ability to emulate human self-organization skills. Nevertheless, this was not enough to change the widespread public belief that digital computers could function as real electronic brains. David Golumbia (2009, 10) has labeled this misconception "computerism": the belief that computers can solve any problem, even those that cannot be formulated in mathematical terms (Golumbia 2009, 14–19). Turing's cautious stand on this point has already been noted. Von Neumann also declared himself unsatisfied with the use of traditional mathematical logic by machines, but his objection remained virtually unanswered.

In a posthumous document published in 1958, von Neumann agreed with those who saw language in all its forms as a historical fact. In his view:

> … it is only reasonable to assume that logics and mathematics are similarly historical, accidental forms of expression. They may have essential variants, i.e. they may exist in other forms than the ones to which we are accustomed. (von Neumann 1958, 81)

Von Neumann went on to argue that, since the brain uses logical systems at a lower depth than we are used to, it is only reasonable to assume that, in order to recreate the decision-making and problem-solving mechanisms of the brain, we must change the traditional language and rules of logic.

These critical voices, however authoritative, could not challenge the power of the digital machine: a device that could apparently solve any problem and emulate any behavior (even intelligent behavior) so long as it was furnished with adequate instructions.

1.3 Bush's visionary dream

The complex relationship between logic and language as forms of communication is not an isolated issue: it lies at the heart of information science, and of software development. For how can we communicate with machines, and how can we interact with them? And how can we formulate our tasks for the machine in a way it can understand? Far from being simple, these questions lie at the core of another inspiring force of computer science. Although this force emerged at the same time as the early computers, it began to wield its influence only some twenty years later. The next section will introduce the analog component into the history of computer development, championed by the father of analogical machines: Vannevar Bush (1890–1974). He was one of the most inspiring figures of contemporary technology. His contribution was significant not only for its scientific results in the field of electrical engineering, but also for his influence on American research policy before and during the Second World War. He was the inventor of the Differential Analyzer, an analog machine, built in 1936, that, during the years of World War II, was still the most powerful calculating device. This was a large mechanical device weighing about 100 tons, which used wheels and disk mechanisms to solve differential equations (Bush 1931).

In 1939 he was appointed Chairman of the Carnegie Institute in Washington and left his career at MIT, where he had been vice president and dean of the Faculty of Engineering. He became one of the most trusted advisers of President Roosevelt, and in 1940 was placed in charge of the institution he himself invented to support the war effort from the scientific point of view, the Office for Scientific Research and Development (OSRD). From this position he organized all the efforts of scientists in favor of the war, and found himself at the center of the largest scientific network ever built for military support, which was one of the key elements of the American victory. After the war, his role became less significant, although he continued to influence the scientific policy of the United States, being one of the founders of National Science Foundation (NSF), the body that still seeks to drive and fund part of the scientific research in the United States.

It was probably this experience, as a scientist and administrator in charge of funding, at a very delicate moment in the history of his country, that prompted him to investigate the role of technology in scientific development and beyond, and the management of information in particular. Starting in the 1930s, his thinking led him to formulate a visionary project made public in 1945, fortuitously the same year in which von Neumann and his group formulated the first project for a calculating machine. Among the articles he published to describe the project, "As We May Think," which appeared in the July 1945 issue of *The Atlantic Monthly,* remains a major source of inspiration for successive generations. Bush's point of departure was the need to redirect the efforts of scientists towards activities more appropriate to a period of peace, without losing opportunities for collaboration, particularly between disciplines.

Two problems facing the world of science were *information overload* and the increasing specialization of scientists, which meant that "the effort to bridge between disciplines is, correspondingly, superficial" (Bush 1945, 6). Bush's solution included the construction of a machine, called the Memex, which was to support the human effort by managing information dynamically and efficiently. He also understood that soon most information would not be produced in text format, but would use other types of media. The dynamic aspect of information included the need to continually update data and to store it adequately, but also to make it easily accessible. The idea of information as an uninterruptible process, which must be shared by people, would also be espoused by Wiener (cf. § 1.5), although, in common with most of his contemporaries, he did not understand completely Bush's position. Indeed, the problem of information overload is still as critical for us today (Yeo 2007).[4] Although digital technology may have made things worse, the use of analog devices, as proposed by Bush, might even today stimulate alternative approaches. Apart from the technology used, the design of the Memex was based on alternative techniques to those traditionally used in a library. Bush's proposal represented a genuine paradigm shift in relation to the access, retrieval, creation and representation of information, and to the management of knowledge itself.

This was Bush's crucial insight: it was reworked in several subsequent revisions of the project he worked on until the 1960s, a witness to how important it was in the construction of this new machine:

> Our ineptitude in getting at the record is largely caused by the artificiality of systems of indexing. When data of any sort are placed in storage, they are filed alphabetically or numerically, and information is found (when it is) by tracing it down from subclass to subclass ....
>
> The human mind does not work that way. It operates by association. With one item in its grasp, it snaps instantly to the next that is suggested by the association of thoughts, in accordance with some intricate web of trails carried by the cells of the brain. (Bush 1945, 32–33)

This idea of simulating the associative strategy adopted by the mind when selecting a set of ideas gave a new perspective to the field of information management and, in a sense, implicitly connected with an empirical tradition in the construction and connection of ideas. Although he may not have been aware of it, Bush's idea was part of a tradition that ran from Locke to Hume, but in opposition to Leibniz (whom Bush cited as a prophet of the machine). This tradition argued that the empiricist construction of associations is the primary and most efficient method for the creation of ideas. In contrast to connection, association may make sense only for the individual who created it, but it is still a powerful tool for finding information and organizing raw data. Bush did not believe that the machine could actually emulate human memory, at least not in the short term, but he was convinced that the machine could "improve" the natural power of the human brain in constructing useful and effective associations. Using analog devices, the Memex would allow the user to select associative blocks and store them in its extensive memory. Bush's machine had nothing to do with the huge, powerful electronic stored-program machines that were beginning to come into service. It was a desktop device, designed not only for scientists but for all categories of workers and professionals. Bush specifically mentioned lawyers, doctors, chemists and historians, who could input useful data for their work, and find it again quickly when needed. He did not recognize the birth of digital technology as such a huge transformation in modeling information; rather, from the very beginning his analog vision recognized that integrated interactive devices would need to retain the analog characteristics of users' cognitive habits. The interface must therefore contain analog elements in order to interact in a friendly manner with the mechanisms of communication and control used by human beings. This is perhaps why he is often regarded as one of the fathers of the development of the Web (cf. § 2.2.2). For digital humanists one of the most inspiring characteristics of the project was Bush's firm belief that "Progress … depends upon the advent of new technical instrumentalities, and still more upon greater understanding of how to use them" (Bush 1959/1991, 183). The main legacy of Bush, then, can be identified in the centrality of the interaction and integration between humans and machines. Technology is revolutionary only if it is perceived and defined by the relationship between people and their needs. So while Wiener (as will be seen in the next section) was looking for ways to build "the human use of human beings," Bush was creating the social and intellectual space for the "human use of technology" (cf. § 3.6). This aspect can be traced back to the origins of the research tradition that focuses on increasing human intelligence via technical means, instead of trying to simulate and replace it by simple use of the machine.

Bush referred to all those technologists who had contributed to the development of the computer as a tool of friendly interaction between man and machine, and developed the idea that the computer should augment human intelligence, not replace it. Among these can be mentioned important figures who did not regard the computer simply as a tool for performing calculations, but for managing and exchanging information between machines and humans. A special mention should be made of Douglas Engelbart, who repeatedly acknowledged his debt to Bush, and can rightly be considered as the father of some of the most successful projects that transformed computer interfaces, and increased their friendliness for humans. These included the invention of icons, group work and groundbreaking devices like the mouse.[5]

To conclude, it can be seen that Bush's project, even though he was an engineer who became a policy-maker, undoubtedly had a strong humanistic component, because he put the improvement of the conditions of mankind at the center of technological development. The focus of his Copernican revolution was not just on the management and organization of information — typically work that would be directed in any case to the humanities — but on humanity over technology.

1.4 A mathematician with a Ph.D. in philosophy

As explained above, the theoretical paradigm of computer science that arose in the second half of the 1930s was later associated with the economic, social and cultural life at the time. The process of elaboration was long and complex, and at times involved opposing conceptions of knowledge, intelligence, and the ways in which these could be represented. One of the most important steps in this debate, which took place at the same time as the transition from the electronic machine of Turing and von Neumann and the launch of Bush's Memex project, was taken during the *Macy Lectures* (1946–1953). These were inspired by Norbert Wiener (1894–1964), among others, and around them was born cybernetics, "the science of control and of communication in both the animal and the machine" (Wiener 1948/1961).

Participants in the Macy Foundation conferences in New York, in addition to those scientists who have made it into the history of information technology, such as John Von Neumann, Norbert Wiener, Walter Pitts, Warren McCulloch and Claude Shannon, included figures from the social sciences such as Lawrence K. Frank and Gregory Bateson, anthropologists like Margaret Mead (all founding members of the group), linguists like Roman Jakobson, the psychologists Hams Lukas Teuber, Donald G. Marquis, and Molly Harrower, but also philosophers, physicists, physicians, biologists, chemists and psychiatrists. Although the meetings mostly focused on technical and scientific issues (Heims 1991, 22), one of the common concerns was to understand how the cyber-technologies emerging at that time could illuminate the workings of the mind and human behavior (and provide application models), while also asking what the consequences of such applications on society might be. Thus a wide range of technical, theoretical and ethical issues, including social development, as well as the diffusion of information technology, were raised during these meetings. Despite the collective effort in the cybernetics revolution, if there was any one man at the center of the intense cultural exchanges that characterized the birth and growth of this new scientific discipline, that man was Norbert Wiener. As a privileged witness and visionary of what he himself called "the second industrial revolution," he had the ability to recognize in advance some of the toughest challenges of the society of information, a society in which not just culture, but also a large part of the economy, would become highly dependent on information technology and communication.

From a strong multi-disciplinary background — at age 17 he received his doctorate in philosophy at Harvard (Conway and Siegelman 2004, 34–40) — Wiener considered himself a mathematician, but his research fields ranged from control engineering, to physics and physiology. He was among the first to recognize the necessity of a theoretical and ethical dimension that went beyond the boundaries of technology. His most significant work in this regard is *The Human Use of Human Beings.* Among its many interesting pages (worth noting is his warning that the protection of technological inventions will be impractical, almost an anticipation of the movement for free software), there is a passage in which he reports a sentence from a review of his first book:

> Perhaps it would not be a bad idea for the teams at present creating cybernetics to add to their cadre of technicians, who have come from all horizons of science, some serious anthropologists and perhaps a philosopher who has some curiosity as to world matters. (Wiener 1950/1954, 180)

Wiener had, in fact, a genuine interest in further research in the "no man's land" on the border between disciplines, and believed that only in this free space could real innovation flourish. Communication between scientists from different backgrounds was therefore at the heart of cybernetics, not only as a scientific objective, but also as a key tool to promote the opening of new fields of research with implications for cybernetic projects. According to Wiener's vision, *communication* and *control* were two interrelated concepts, because control was a special case of communication: "When I control the actions of another person, I communicate a message to him, and although this message is in the mandatory mode, the technique of communication does not differ from that of a message of fact" (Wiener 1950/1954, 16). Communication was an interaction with other people or machines in an attempt to get feedback. The concept of feedback lies at the center of cybernetics research as the mechanism of reaction and re-balancing of agents in response to messages from outside. Although, by its very nature, this remains inaccessible to the producer of the message, it nevertheless ensures the effectiveness of the interaction. Wiener therefore focuses his attention on the very concept of communication beyond its human dimension. In his text *The Human Use of Human Beings* he clarifies his position on this issue:

> ... society can only be understood through a study of the messages and the communication facilities which belong to it; and that in the future development of these messages and communication facilities, messages between man and machines, between machines and man, and between machine and machine, are destined to play an ever increasing part. (ibid.)

One of the primary purposes of Cybernetics, as a transdisciplinary field, was the study of "a language and techniques that will enable us indeed to attack the problem of control and communication in general" (Wiener 1950/1954, 17). The importance of the role of communication and subsequent research on the languages and technologies that have made this interaction between machines and humans possible was at the center of the line of research that we can call cybernetics. Although the discipline did not have a long or easy life, the seeds it sowed were longer-lasting and more influential than is generally recognized in the subsequent history of science and technology. For Wiener, future efforts would be directed towards constructing the most profitable and efficient interactions between humans and machines. Although, by the time the *Macy Conferences* eventually came to a close, the momentum of cybernetics had been partially depleted, still Wiener's tireless cultural and organizational activities never stopped. He continued to promote meetings, such as the *cybernetics dinners* held in Cambridge, which were attended by many of the scientists who later became part of the technological revolution of the 1960s, and among them certainly Licklider, whose contributions will be examined in the following sections. The problem of language and the integration of communication into the most varied of applications, from prostheses to the education of machines capable of learning, always lay at the center of these interdisciplinary exchanges that interested Wiener so much. In particular, his cultural leadership at Princeton, and more generally throughout the world, from Russia to China, from France to Mexico, meant that his name became associated with cybernetics and a prophecy of the near future, which expanded his influence beyond his immediate field of influence. This can be seen in the interest in man-machine interactions, in the search for "high-level" programming languages (those more oriented to users than machines) and finally to the growth of computer networks to facilitate communication between humans and machines.

1.5 Wiener's ethics and politics of the computer

Before delving into the influence of cybernetics on the development of information and computer technology, it is crucial to recall Wiener's vital contribution to the understanding of the historical, ethical and social implications of the diffusion of information technology in society. He realized very soon that

> Those of us who have contributed to the new science of cybernetics thus stand in a moral position which is, to say the least, not very comfortable. We have contributed to the initiation of a new science which ... embraces technical developments with great possibilities for good and evil. (Wiener 1948/1961, 28)

The moral dilemma to which he refers echoes the loss of innocence among scientists, which took place as a result of the dropping of the atomic bomb during World War II. As Michel Foucault suggests (1970, 21–23), in that context scientists realized their work could not be neutral with respect to good and evil. It was necessary to take a stand and to take responsibility for activities in the gleaming, sterilized and lonely research laboratory. The position taken by Wiener was publicly clear, and he personally paid for his choice in terms of isolation. After the War he refused to be involved in any research project funded, even if only indirectly, by the Department of Defense, or in research involving private enterprise as privileged partners of an academic institution.

It is worth investigating what, according to Wiener, were the risks society would run when the "second industrial revolution" was complete, i.e. when the promises of cybernetics had come true. Once built, machines designed for control and communication would be able to replace not only manual work but also the minds of many workers, and this would cause first of all intellectual unemployment:

> … This new development has unbounded possibilities for good and for evil. …. It gives the human race a new and most effective collection of mechanical slaves to perform its labor. Such mechanical labor has most of the economic properties of slave labor, although, unlike slave labor, it does not involve the direct demoralizing effect of human cruelty. However, any labor that accepts the condition of competition with slave labor accepts the conditions of slave labor, and is essentially slave labor. (Wiener 1948/1961, 27)

The consequences of this situation would have been that "the average human being of mediocre attainment or less has nothing to sell that it is worth anyone's money to buy" (ibid.). This would mean either that the mechanisms by which the working classes were treated would have to be changed, or it would be necessary to rethink the value-structures of society so as not to leave out relevant layers of citizenship. In an effort to create a civil consciousness with respect to the progress of science and its effects on the world of work, Wiener became involved in raising awareness among trade unions, but was disappointed with their inability to understand the situation.

The second area of risk that he saw, and in which he was engaged, was the analysis of the evolutionary processes of science. Wiener opposed the idea that the findings should be shrouded in ever-greater secrecy. The argument he used against secrecy is interesting and topical. His idea was that information could not be regarded as a commodity as any "information is more a matter of process than of storage" (Wiener 1950/1954, 121). Considering the dynamic and procedural characteristics of knowledge transmitted through information, Wiener believed that imposing secrecy on research simply meant slowing it down, without hiding it from the enemy, who could simply maintain strategies that allowed him or her to find out the important details. He also likened military top-secrets to the patent systems followed by private enterprise, and increasingly also by universities, noting the negative consequences of a slowing down in the process of innovation, and the way in which knowledge produced by several people ends up benefiting only a few. His position is expressed as usual with great courage:

> The fate of information … is to become something which can be bought or sold … It is my business to show that it leads to the misunderstanding and the mistreatment of information and its associated concepts … beginning with that of patent law. (Wiener 1950/1954: 113)

His personal crusade against intellectual property, already mentioned, brings him even closer to modern research, the battle for free code and shared knowledge (Hess and Ostrom 2007, Lessig 2004, Berry 2008). The situation nowadays is, however, quite complicated. Although "knowledge is more a matter of process than of storage" it is now easier to store the process and to manage it. This is exactly what happens with social network services that host the process on their servers and then claim ownership of the information produced. Wiener could not expect that storage and communication tools would have evolved so quickly. However, he clearly had in mind one of the characteristics of the knowledge society, the centrality of information access and distribution.

This centrality becomes a strategic asset in the future imagined by Wiener, and, he had foreseen very precisely that if the means of communication did not adequately satisfy this need, we would risk the construction of anti-democratic mechanisms, supported by the misuse of machines (of which the computer would be just an example), favoring the centralization of power:

> … such machines … may be used by a human being or a block of human beings to increase their control over the rest of the human race or that political leaders may attempt to control their populations by means not of machines themselves but through political techniques as narrow and indifferent to human possibilities as if they had been conceived mechanically. (Wiener 1950/1954, 181)

This risk and others related to the automation of war (the so-called *push-button war*) could only be averted by producing a generation of scientists, informed and free from all ties with society and politics, who could make decisions by assessing the situation solely in their capacity as experts. This solution may seem a bit technocratic and very utopian, but its echo has certainly informed some of the more extreme positions adopted by the pioneers of ARPANET, the computer network that was the forerunner of the Internet. His idea of the machine as a potential concentrator of power is echoed in some of the most critical literature about the Web, which underlines the risks related to the controversial relation between control and freedom of the digital network (cf. § 5.9 for more details and a discussion of Chun 2006).

The second method identified by Wiener to oppose the worst consequences of the first machines programmed to control society could be categorized as resistance. It was to provide the maximum possible information on what was happening in science and technology to foster the growth of awareness in the general population and to allow the implementation of appropriate corrective measures. The decision to write his two key texts on cybernetics — that of 1948 and another, more accessible work that appeared two years later — was precisely directed at providing remedies to these risks by making them as public as possible. Its purpose was to help build a more just society whose values were more than purely commercial. Such a society would have taken on the contradictions born of an unconscious use of technology, and would choose to rule in the light of sharing and collaboration. Society so reorganized would allow "the human use of human beings" that the spread of cybernetic machines was putting in danger.

In conclusion, the positions taken by Wiener constitute an important lesson for the digital humanist today because they may represent a critical vision *ante litteram* that proposes solutions oriented to humans concerning the kinds of machines that were being built, or had recently become available. His insight into the barriers that patents and copyright posed to the development of knowledge, and his anticipation of current disputes over digital protection methods (whether in art or science) raises the question of whether knowledge can only renew itself through free and open transmission (cf. § 5.9.1). His transdisciplinary education meant that he was continually examining the development of science from a social and ethical point of view. And perhaps it is no coincidence that, despite having proposed perhaps the first project for an electronic machine with stored programming in 1941 (although it was ignored due to a lack of understanding on the part of Bush), he was never again interested in the construction of real machines. Perhaps he feared that without guidelines they would only become mechanisms of automation, instead of improving human communication.

1.6 Licklider and the man-machine symbiosis

Joseph Carl Robnett Licklider (1915–1990) was the right man at the right place at the beginning of the 1960s. "Lick," as friends and acquaintances called him, was a person capable of making everything that he worked on seem simple (Mitchell Waldrop 2002, 7–8). His training was in experimental psychology and his main research field was psychoacoustics — a discipline that currently falls within the domain of neuroscience. His doctorate in 1942 was the first to identify the areas of cat brains that were activated in hearing sound waves of varying intensity. In his field, sophisticated equipment was beginning to be used to study neurons and he was among the first scientists to come into contact with a mini-computer, the PDP1, which allowed direct interaction with its operator (Hauben 2007, 110). But Licklider was also part of a group of people who were influenced by Wiener's ideas of cybernetics: he participated in the cybernetics circle that met in Cambridge, MA every Tuesday evening, and attended their dinner seminars. He recalled those years after World War II in an interview in 1988, and admits that those meetings and dinners had a powerful influence on his career, because his training in psychology became enriched with information about computer science and the theory of communication, which were of great importance for his future as a scientist. Licklider took part in the Macy Seventh Conference, held in March 1950, giving a paper on the distortion of language and its ability to remain intelligible. Licklider's name can thus be added to the list of scientists who were stimulated by the transdisciplinary nature of the new field of cybernetics. Although unknown to the general public, a number of important functions in the development of information technology and telecommunications are due to Licklider, for he was at the head of one of the offices of ARPA (Advanced Research Project Agency), called IPTO (Information Processing Technologies Office), which was a major supporter of man-machine interface research, and which funded the construction of the ARPANET, a network that linked universities and research centers — the forerunner of the Internet.

He held this position from 1962 to 1964, and was also central in the launch of the ARPANET, although that project was only realized under the guidance of his colleague Robert Taylor in 1969. According to Jack Ruina, the director of ARPA who appointed Licklider to his new office, his job was to address such issues as *command and control* and the *behavioral sciences* — two typical areas of research for the military — but the flow of research supported by Licklider helped to interpret them for the realization of human–machine symbiosis. In fact, all projects funded by IPTO were in this area.

What drove him to pursue this line of inquiry? He shared Bush's conviction, from his own experience, that most of the work in research was taken up by putting oneself into a position where one could be creative. So he made every effort to ensure that the machine might not just be an instrument to help scientists save time on mechanical work, but also to share that part of the work devoted to thought, the part he formally called "creative." The device Licklider was thinking about from the early 1960s was influenced by his participation in the Whirlwind project. Whirlwind was to produce the SAGE (semi automatic ground environment) system, able to assist in the discovery of enemy forces through the use of radar, if war broke out. It was a simulation project in which the monitoring instruments and their signals were analyzed by operators with the help of machines. Even though the project, which ended in the 1950s, was already old in technological terms and never became operational, it was extremely important in enriching the scientific imagination: the idea of a machine capable of interacting directly with its operator, without the intermediary of the technician and long delays waiting for program output. In short, the paradigm of human–machine symbiosis was being born, based on theoretical studies on communication that lay the center of all processes in cybernetics, and from the model of a desktop machine like the Memex that would act as a supplement to human memory (cf. § 1.3).

The idea of symbiosis is explicitly opposed to what Licklider called — along with Wiener — "automation," identified with the idea of "mechanically extended man." This was about replacing parts of humans with mechanical devices that would perform the intended task instead of the operator. Symbiosis was not in competition with the first artificial intelligence projects active at that time; it opted for a lower profile, considering that an intelligent machine would probably be realistic in the medium to long term, while in the short term (5–15 years), it would be desirable to develop more sophisticated and real-time interaction with machines.

Licklider explained specifically what he meant by his idea of a symbiosis between man and machine:

> … the other main aim … is to bring computing machines effectively into processes of thinking that must go on in "real time," time that moves too fast to permit using computers in conventional ways …
>
> To think in interaction with a computer in the same way that you think with a colleague whose competence supplements your own will require much tighter coupling between man and machine than … is possible today. (Licklider 1960, 4)

His appointment as director of IPTO suddenly placed this project at the center of national science policy. And it translated into an action plan that granted substantial funding to those scholars who were moving in the same direction.

Symbiosis meant the humanization of the computer, and for this reason in his 1960 article the characteristics of dissimilarity between humans and machines were identified, along with the efforts of integration that would be needed. This list of open problems immediately became the agenda of IPTO and the research centers it funded. It also affected the birth of personal computing and the creation of a network of interconnected computers, two essential conditions for the development of digital humanities. Symbiosis, in fact, presented a vision of computers in which their different tasks of calculation and data management were fused with those of storage and information retrieval (Licklider 1960, 6). Once this relationship between such diverse tasks was built, and having integrated the component of communication into all levels (between operator and machine, between machines, and between people via machines, as Wiener foresaw) one is faced with a machine comparable to that designed in 1945 by von Neumann (cf. § 1.1), although still governed by a processor, a memory and a set of programming instructions. To give just a few examples: it incorporates mechanisms for the collective management of machine time, or introduces input–output devices, unimaginable to the inventors of the calculator. In addition, from the point of view of programming languages, it will be seen that Licklider sought to support different techniques to those adopted previously, more suited to human than to machine language. But the newest element integrated into the new machine, designed by cybernetics and related fields, is the role the user has in the development of mechanical processes. This idea was the opposite of that proposed by the Turing Machine, where the operator was replaced by an appropriate set of instructions. Following the model that may be called cybernetic, the machine was, in fact, a participant and complicit in all its activities, replacing the human operator only in the purely automatic components of the task, and supporting but never replacing him or her. From this perspective it may be said that, in the generation of projects funded by Licklider, the human being became integrated with the computer and, to achieve this, it was necessary to construct input–output tools that would facilitate its implementation. It was in this context that some devices were invented that are still useful for the so-called *human–machine interaction,* such as the mouse, the use of icons in the interface, workgroups, screens suitable for graphic applications, and so on.

1.7 Libraries and information processing

Before describing the obstacles to the implementation of the kind of symbiosis that Licklider hoped for, it will be instructive to see first how the agenda drawn up by the research group he financed and built had consequences for both computer science and eventually digital humanities. It was, in fact, a new way to manage and represent data, build information and thus create knowledge — all activities central to the research process.

Licklider was invited to a series of evening seminars on "Management and the computer of the future," organized in 1961 to celebrate the centenary of MIT. At one of these seminars, dedicated to the computer in the university, he openly declared his opinion that the computer has affected "the whole domain of creative intellectual processes" and, in particular, he supported the argument that "information processing" would one day become an important scientific field:

> Planning management communication, mathematics and logic and perhaps even psychology and philosophy will draw heavily from and contribute heavily to that science (the information processing). (Licklider 1962, 207)

Licklider recognized decisively that all disciplines concerned with creative processes of any kind were related to computing, and that every one of them, even the humanities, would contribute to its development. This message has not always been adequately understood in the context of computer science and perhaps not even in the humanities, but evidently it was clear from the beginning to all who had a significant influence on the organization of this discipline.

However, Licklider identified many obstacles to the implementation of symbiosis. The first was the difference in speed between the two entities: the human and the mechanical. His solution favored the development of *time-sharing,* a technology that would allow different *consoles* to make use of the same machine at the same time, taking advantage of the speed of execution of the big machines that were then available. To show the positive effects of this technology, which was the first sub-goal of its leadership to IPTO, he pointed to libraries as places where most could be made of the technology:

> It seems reasonable to envision, for a time 10 or 15 years hence, a "thinking center" that will incorporate the functions of present-day libraries together with anticipated advances in information storage and retrieval and the symbiotic functions suggested earlier in this paper. The picture readily enlarges itself into a network of such centers, connected to one another by wide-band communication lines and to individual users by leased-wire services. In such a system, the speed of the computers would be balanced, and the cost of the gigantic memories and the sophisticated programs would be divided by the number of users. (Licklider 1960, 8)

In the years before his assignment to IPTO, Licklider was appointed by the Council of Library Resources to explore the role technology would play in the libraries of the future. The work by his research group was carried out between 1961 and 1963, and reached fruition with the publication of the book *Libraries of the Future* (Licklider 1965). According to Licklider, the mass digitization of knowledge would become an essential new tool for the consultation of materials; also the availability of these large machines that were accessible by many users would allow everyone to get digital content, thus permitting fast, direct and complete access to all human knowledge. Seen from this perspective, computing would have an immediate impact on content organization, access and searching: some of the key fields of the humanities. Licklider did not pose the question of who should hold the expertise necessary to reorganize the libraries in his research, because he knew that in the cybernetic tradition certain objectives were either worked out within an inter-disciplinary framework or they were abandoned.

The model of time-sharing, necessary to overcome the difference in operating speed between humans and machines, had at least two consequences, one positive and the other negative. From a social point of view, the collective use of a single machine, perhaps at a distance, led to the birth of a sort of cooperative team spirit among the programmers, who laid the groundwork for a culture of sharing and respect for each other's work that became the basis for the development of the Internet, which itself began as one of many related community projects aimed at the distribution of information as an essential characteristic of these new technologies. From a technological point of view the Internet is still based in part on this time-sharing model through its client-server design, an architecture structured on the limited availability of processors and memory, that allocated the burden of work on the server while the "clients" (our computers) use only the results of that work, like the terminals connected to the mainframes of Licklider's era. In the modern world the widespread availability and evolution of technology should have rendered the hierarchical network obsolete some time ago, but once some infrastructures have been built, it is difficult to revise them. The story of the Internet is about the biases of technology and their roots in its complex and cumbersome evolution.

According to Licklider, another factor that hindered the realization of human–machine symbiosis was language. Machines and human beings were using very different languages because they were based on very different principles. The computer had to be precisely guided through a series of detailed steps, whereas humans could content themselves with simply knowing the desired goal and, being bound by the result, putting into place the technical means for obtaining it. Licklider noted that in order to achieve the symbiosis it would be necessary to make use of new principles of communication and control than those then in use. In addition to techniques being tested in the field of artificial intelligence, which consisted in the development of heuristic strategies to get results, and the consequent construction of declarative rather than procedural languages, he called for the strengthening of a second line of research that would move towards creating separate sub-routines which could be called upon by the operator directly as needed. The idea was to build real-time pathways to encourage interaction with the machine, depending on the needs of the moment. This would not be a ready-made program but an agreed set of constraints that would be manipulated by the skill of the operator and the requirements of the situation — a more flexible modular system that would allow integration with the machine.

The relationship with the language of the machine is another "humanistic" issue from the development of information technology.[6] The languages and their structure, although they may be strategies of digital technology, belong to the realm of language-related research. How should a language be built so that it can be used mechanically? Do natural languages have a formal structure? Is it possible to represent — in a procedural way — problems to be solved through the machine? These were the kinds of questions posed by the linguist Noam Chomsky (1957), whose work in developing formal descriptions of natural language formed the basis of all modern computer languages, including markup languages. Such questions cannot be answered directly because they deal with the complex general issue of what a language is and how it works. However, it is still possible to define the challenge of finding methods of communication to facilitate the interaction between humans and machines — an issue which lies at the core of computer science — as related to the philosophy of language and/or to the linguistic turn of the science and humanities of the last century.[7]

Licklider recognized the centrality of the problem and felt that the solution was not to be found in a human adaptation to the mechanical system, but rather a continuous communication between humans and machines, in the definition of a set of programs each of which could be useful in a specific context. The interactive nature of the symbiosis required the implementation of routines written in a language understandable by the machine, but at the same time capable of conforming to the communication patterns of human beings.

The last problem identified by Licklider as an obstacle to enabling a more fertile interaction between humans and machines was the state of current input–output devices that were at that time unsuited to the role of the computer as a communication tool. To overcome this limitation the IPTO concretely funded many projects in the field of graphic interfaces and input–output solutions that would allow a faster and more immediate interaction with the machine. Licklider's idea consisted of the construction of three types of input–output device, the visionary aspect of which is still impressive, given that he was working back in 1960, when machines were very different from today's desktop computers. He supported the creation of desktop displays, able to show results and to control the activity of the machine. Screens like those available only recently, were conceived by his creative imagination. He also foresaw wall displays that could serve to facilitate harmony in the case of group work, and small screens that, although separated from each other, could still display the same information. All these ideas have been realized only recently, such as the use of communal virtual screens to enable real time collaboration (as used when writing this book). And finally, Licklider imagined that systems for the recognition and production of language would eventually become available. In this area as well, very important steps have been taken, although his original vision is far from being realized.

1.8 Conclusion

This chapter has tried to argue that the contribution of the humanities to the early development of the computer is far greater than is generally believed. Many of the leading theorists who shaped the design of the modern digital computer: people like Turing, von Neumann, Leibnitz, Wiener, Bush and Licklider came from very varied educational backgrounds. They were often trained either directly in the humanities or social sciences (Leibniz, Wiener, Licklider) or they consulted with a wide range of academics, including those from humanistic disciplines. The contribution of linguists to the development of computable formal languages is especially significant. Likewise the development of human–computer interaction design can be traced back to the Licklider, and his background in psychology; and Bush's Memex was based on his neurological idea of associative rather than hierarchical links between items of information, which led eventually to the creation of the modern Web.

Another important lesson that can be learned from the development of the early computer is the reluctance of pioneers, like Turing and von Neumann, to concede that the digital computer could ever possess the ability to think like a human — a misconception that persists today. Others, like Bush and Wiener, foresaw how information would become a valuable commodity in a digital world, and that the changes set in motion by the development of cybernetics had the potential to do as much evil as good.

One final point worth making is that the success of the digital computer also owed something to chance: the development of alternative analog computers was cut short by the Second World War and not resumed, and criticism of the path taken in the development of the digital computer by leading theorists was ignored.

For the modern digital humanist the message that this historical analysis provides is one of caution: not to assume that a computer can solve any problem, or that the continued development of digital computing in all its aspects is necessarily in the best interests of society. It should rather be regarded as one particular path which has been followed thus far to the exclusion of all others, and which may still be retraced at some future time. And in this retracing humanists will doubtless play, as before, an important role.

The next chapter will follow the development of the Internet and the World Wide Web, starting from Licklider's involvement in the ARPANET project. It will also investigate the social inequalities and cultural biases that the new global system of information, as envisaged by the pioneers of the computer, created.

Chapter 2
Internet, or the humanistic machine

2.1 The design of the intergalactic network

In April of 1963, as manager of IPTO, Licklider wrote a famous letter addressed to a group of staff scientists, whose names read like a veritable who's who of the leaders of computer science at the time. The letter was sent to postpone a meeting, and Licklider was the point of contact for the projects and their points of intersection, and also managed joint projects relating to longer-term goals. The members of this "intergalactic network" were striving to combine forces to overcome various barriers to communication, with the ultimate goal of creating an interactive and cooperative system. The project involved the construction of a network of connected computers, which would allow different operators to use a program on different machines, and to store the data in memory, even if it came from other parts of the system. To realize this dream, which was the material realization of his project of symbiosis, based on the concept of time-sharing, a number of issues that would have hindered cooperation would have to be resolved.

As seen in the previous chapter, one of the key problems was the question of which languages were to be used. This became even more crucial in the context of the network. For what should the language of the intergalactic network be? The language of time-sharing? Or was there a need for new *koinē*? The use of programs originating on other computers implied that they would have to communicate using different languages, so it seemed a good idea to at least try to standardize best practice for storing data and information. There was also the problem of who should manage the communication protocols and traffic on the machines. Licklider's prophetic suggestion was to allow the network to autonomously manage traffic, data storage and priority of access. It would then be able to modify shared files without recourse either to the users themselves, or to a higher level. For the first time, the user was put at the center of the design:

> It seems easiest to approach this matter from the individual
> User's point of view — to see what he would like to have, what

> he might like to do, and then to try to figure out how to make a system within which his requirements can be met. (Licklider 1963, 2)

The birth of user-centered design that has played such a leading role in the design of computer systems from the 1980s onwards[1] can be traced to this letter, and other communications from those years. The centrality of the user means computing at the measure of man (or woman), and especially the idea that using a computer does not rely on making things simpler for the machine, but using the machine to support human activities.

In this letter, the relationship between machines and communication devices is expressly constructed through the implementation of standards, interpreted languages or tools to facilitate access and retrieval of previously acquired resources. The computer is no longer limited to a relationship with an individual user, but is seen as one element in a complex system of communication, just as Wiener had predicted in 1948. While Wiener saw at once the risks regarding control and the concentration of power that these new models of communication would bring about (see § 1.4 and 1.5), Licklider saw them as an opportunity to join together his various interests as a scientist and to expand opportunities for collaboration between scientists and the objectives of the military.

> … the military greatly needs solutions to many or most of the problems that will arise if we tried to make good use of the facilities that are coming into existence. I am hoping that there will be, in our individual efforts, enough evident advantage in cooperative programming and operation to lead us to solve the problems and, thus, to bring into being the technology that the military needs. (Licklider 1963, 3)

Licklider, who in any case was working for an agency funded by the Department of Defense (unlike Wiener), saw in his work an opportunity to steer the computer in a desirable direction for scientists, while at the same time supporting the national war effort. So the technology of network communication was born, between Wiener's warnings of the "human use of human beings" on the one hand, and Licklider's engaging and somewhat naive enthusiasm on the other. Licklider, though, was well aware that this developing technology would enhance the ability of the army, or other similar institutions, in their control and exercise of power. Nevertheless, he did not renege: he portrayed the marriage between computers and communication as a victory for science, rather than as a tool for power. It was this step that transformed the activities of command and control that he had to develop at ARPA, into projects of human–machine interaction, realized through collaborative programming effort.

Although he resigned as manager of IPTO in 1964, Licklider's influence continued to be felt in the Research department. The real network project was led by Robert Taylor, an experimental psychologist who called himself a "professional student," having followed — for pleasure's sake alone — courses in mathematics, philosophy, English and religion. He also obtained a master's degree in psychology, but declined an invitation to complete a PhD, refusing any kind of specialism in favor of a highly transdisciplinary path. Taylor came to be considered as one of the top technology executives of the time: after setting in motion the ARPANET project, the nucleus of the first network that later became known as the Internet, he directed the Computer Systems Laboratory (CSL) at Xerox PARC, from which came the most important innovations in personal computers. It was Taylor who first realized the importance of the screen as a medium of communication between human and machine:

> Which organ provides the greatest bandwidth in terms of its access to the human brain? Obviously the eyeball. If one then contemplated how the computer could best communicate with its human operator, the answer suggested itself: "I thought the machine should concentrate its resources on the display." (Hiltzik 1999, 9)

In 1966 Taylor became director of IPTO and realized that if he wanted to communicate with MIT or Berkeley he had to move his chair and turn on different terminals in his room. So, by making use of Licklider's insights, he envisioned connecting the different laboratories with each other through a network that spoke a common language, where everyone could interact with each other's machines. In 1968, he wrote with Licklider a famous article entitled "The Computer as a Communication Device," now considered a classic in the field, in which he prophesied that communication through machines would outperform direct verbal communication.

2.2 The computer as a communication device

This article, which anticipates the birth of the ARPANET by a year, is a sort of manifesto for a new type of informatics with a unique role in society, in contrast to the agendas, concepts and reference models of communication engineering. Licklider and Taylor's idea was very simple, though revolutionary, considering the practicalities of the machines of the time: it was for a device that could be used not just for the transfer of information but for interaction. Some elements, such as communication itself and its agents, would have to be rethought if the benefits of these next-generation devices were to be realized.

Taylor had repeatedly claimed to have been influenced by Bush and Wiener (Aspray 1989, 5), and of course by Licklider, in reformulating the relationship between humans and machines. Already Bush, but more especially Wiener, introduced the idea of communication as something more complex than a mere stream of bits encoded to reach their destination, where they would be symmetrically decrypted. It also included the idea of a response function to the data stream, which took the form of feedback; it was, so to speak, a naturally interactive and interconnected mechanism that included a relationship between the two parties involved in the exchange. If the feedback was at a high level this meant that interaction was no longer confined to the mere passage of information, but produced a real change in the scenario, involving new rules and even a new model in which to frame the various elements of communication.

The position of Licklider and Taylor pushed beyond the simple transmission of data. They expressly stated that "their emphasis on people is deliberate" while denying the centrality of the machinery responsible for the transmission of the data: "to communicate is more than to send and to receive." They were convinced that

> … we are entering a technological age in which we will be able to interact with the richness of living information — not merely in the passive way that we have become accustomed to using books and libraries, but as active participants in an ongoing process, bringing something to it through our interaction with it, and not simply receiving something from it by our connection to it. (Licklider and Taylor 1968, 21)

The central object of the interests of these two experimental psychologists was the support and help machines could give to the most creative aspects of human communication. To accomplish this end, they needed a medium that could be tailored to the circumstances, a plastic and dynamic medium which everyone could contribute to, and experiment in. They believed that the computer was just such a medium: "a well-programmed computer can provide direct access both to informational resources and to the processes for making use of the resources" (Licklider and Taylor 1968, 22).

If the computer could handle both information and the processes that allow it to be used, it is clear that users of this tool should not only be advanced technical and communication engineers or programmers, but also "creative people in other fields and disciplines who recognize the usefulness and who sense the impact of interactive multiaccess computing upon their work" (Licklider and Taylor 1968, 30–31). The supercommunity of the ARPANET, then, would include, alongside the technicians and engineers, creative people from other areas who were able to exploit the new communication tools for their areas of interest. Part of this community can certainly be identified with as digital humanists, together with engineers and programmers, and groups from other disciplines. All these members of the community had equal priority in the interactive information process as they followed their own research agendas.

The only group excluded from the community interaction, other than those who did not work in a creative or informational context, were those who interpreted the computer as a simple connector and data transmitter, without valuing its potential as a tool to encourage interaction and the building of tools for the externalization of cognitive models.

Starting from these considerations, Licklider and Taylor made it clear that, although computer programs were important because they allowed the raw data to be structured and manipulated at a higher level, they were only a part "of the whole that we can learn to concentrate and share. The whole includes raw data, digested data, data about the location of data — and documents — and most especially models" (Licklider and Taylor 1968, 29).[2]

Finally, they also focused on the future of these online interactive communities and imagined how they would look. This exercise could be seen as a self-fulfilling prophecy, rather than as a vision about the future, because that future was literally "invented" in their 1968 article. That paper imagines a reality of online communities composed of single individuals or organized groups separated geographically but united by interest, along with the use of computers for every information-based transaction, contributing to lower connection costs, and the replacement of letters and telegrams with electronic messages.

Finally, Licklider and Taylor made a prediction that is staggering in its precision and its evocative character:

> When people do their informational work "at the console" and "through the network," telecommunication will be as natural an extension of individual work as face-to-face communication is now. The impact of that fact, and of the marked facilitation of the communicative process, will be very great — both on the individual and on society. (Licklider and Taylor 1968, 40)

This could not be a more precise and detailed description of what happens today in our daily experience of the Internet. Licklider and Taylor, however, recognized a risk that constituted a crack in their enthusiasm for the network-to-come: the *digital divide* (cf. § 2.1). They wondered if "being online" might become a privilege or a right, and this raised the question of whether the network would become a benefit or a risk to society. Subsequent events appear to bear this out, and the communication technology they gave birth to in an excess of optimism has also revealed other problematic areas.

2.2.1 The birth of the ARPANET

The rest of the story of the ARPANET is well known and often recounted.[3] In 1966, Taylor won from Charles Herzfeld, the director of ARPA at the time, an initial funding of one million dollars for the construction of the ARPANET infrastructure. In December 1966 Taylor was finally able to convince Larry Roberts, the communications engineer who set up the project, to collaborate with ARPA. From the technological point of view, there were a few innovations and important choices made in terms of function. The first of these was the use of packet-switching — not a new idea, but one rediscovered at the time — of routing the packets of information one at a time, and allowing each to follow its own path independently of the others. Second, the choice of a network architecture that provided machines at each node dedicated to the management of traffic, which did not keep track of the packets in transit. The original plan, however, involved the creation of a single network to control all the data that passed through it. If this original solution had been chosen, the Internet would surely not have been able to grow. The choice to distribute control over all the nodes not only had an impact on the architectural design of the network, but also on its social and political conception. No one could organize the network at their own behest, but anyone could contribute to its shape, at least in principle.

In 1967, the first nodes to be connected were identified. None of these nodes was military; they were all university research centers. The first node to be connected was the University of California, Los Angeles (UCLA) under the guidance of Leonard Kleinrock. The second was the Stanford Research Institute (SRI), controlled by Douglas Engelbart. The first message was sent on October 29, 1969. Later were added the nodes of the University of California, Santa Barbara and the University of Utah, to where Ivan Sutherland, who had directed the IPTO before Taylor, had moved. By December 5 of that year, the network consisted of these four nodes, and could be considered to be in operation. Since then, the network has not stopped growing and increasing its services with the cooperation of all those connected, without any limitation. Innovations came about through the precepts of learning and producing by using. The intergalactic community, as Licklider called it, met and worked collectively for the improvement of common tools, driven by a sense of belonging to a shared project. Each member contributed without any personal gain, apart from the reputation that he or she was building among their peers through the success of their contributions. No scientist among those who collaborated in the development of the ARPANET would ever file a patent to protect the innovations that spread throughout the network; the diffusion of their solutions was the prize that this group of pioneers most desired: the commercial aspect was totally absent.

A special mention is due to Douglas Engelbart, who was one of the designers not only of the network but also of a set of tools, which came together in the construction of personal computers. It was SRI, the center that he directed, which studied the techniques needed to build a friendlier interface for the computer. It was he who invented the idea of technology as human augmentation, experimenting with a way of interacting with the computer that did not conform to the idea of "simulation," the idea so dear to the promoters of artificial intelligence. For this reason, his research was funded both by Licklider, with whom he had a deep common understanding, and by Bob Taylor, first at NASA and then at IPTO. It was at the SRI that some of the features of graphical operating systems that are still our interface with the machine were invented. Engelbart invented the mouse, windows, icons and the idea of groupware. However, many of his colleagues grew tired of his management style, and followed Bob Taylor to CSL at PARC. It was here that proper graphical user interfaces were created and the *object oriented* programming that made them possible. In this legendary laboratory Taylor, with the expertise of part of a group from SRI, and its undisputed capacity for managing technology, in 1973 built the first prototype of the personal computer, the Alto. Although the Alto was not a commercial success, it was later an inspiration for both Apple and Microsoft.[4]

It is interesting to note that the idea of a computer suitable for personal use and the ARPANET grew out of the development of user-oriented technology and the idea of human augmentation, rather than from a desire to replace humans with self-sufficient machines. This concept that underlay both the personal computer and the ARPANET can be traced in the line of technological thinking from Bush and cybernetics to Licklider, Taylor and Engelbart, and is based on an interpretation of communication as interaction with the machine and with other human beings via properly programmed and organized mechanical devices. It was this common strategy in management and funding that produced the most remarkable achievements in terms of changing the role of technology in all sectors of society. As long as machines were shut up in a few computer science centers, they would never have the social impact that comes from the concept of the terminal or a personal computer accessible from the desk. Only the process, originally symbolic, of concentrating on the needs of the user, had the power to transcend von Neumann's model, in the design of more interactive devices that were symbiotic with the common user: a machine designed for humans. Thus even before the digital humanities, it was realized that the computer should be a humanistic machine.

2.2.2 The WWW: an authoring system in the heart of Europe

When Tim Berners-Lee presented the project for the World Wide Web (WWW) at CERN in Geneva, an institution that dealt in nuclear physics, he could not imagine how great an impact his creation would have not only on the Internet but on the whole world. In the first instance, his work was merely intended to facilitate the passage of information between researchers at CERN, to avoid duplication of efforts within the research center. It was supposed to rationalize resources and organize information to make it easily accessible. The first proposal, a memo submitted in March 1989, made no mention of the Web: it was called simply *Information Management: A Proposal*.[5] In an interview[6] granted on the 20 year anniversary of the proposal, Tim Berners-Lee, recalling the origin of his proposal, stated that CERN was the natural place to invent the Web, because of the need to connect professionals and scientists from all parts of the world, who were using the most diverse systems of hardware and software. The need for integration of all that wealth of information required some means of sharing — at this point, only via an "imaginary" tool — that would allow the information to be held in one place, while making it accessible on alternative platforms. This imaginary system had, however, very realistic features: it would have to combine the function of an authoring system with a mechanism for viewing the pages, as well as a strategy for connecting independently produced documents or their components through the use of links. Tim Berners-Lee said that the invention was easy because each instrument was already there, ready for use. The Internet was already there, an infrastructure designed without assumptions having been made about how it ought to be used. Its protocols already existed: TCP/IP (Transfer Control Protocol/Internet Protocol), a group of rules that regulated the transmission and routing of data, DNS (Domain Name System) that worked to uniformly define the various resources of the Internet by associating each connected server with a unique number. The concept of hypertext was also already present. It had been invented by Vannevar Bush, then independently by Ted Nelson. Douglas Engelbart had also worked on something similar to the Web, limited only by the fact that at the start of the 1960s, the Internet was not yet invented.[7] The idea of the hypertext was simply adapted to the protocols of the Internet.

But a closer examination of the process of invention reveals a situation slightly different from that told by Berners-Lee. As often happens, the main players in the invention of a new technology do not have the clearest idea about what they have done. Above all, the idea that Bush had contributed to the idea of hypertext as an authoring system is problematic. Bush's proposed machine consisted of a personal workstation in which each user could connect to knowledge represented analogically in his Memex (cf. § 1.3).

It was instead Ted Nelson who invented hypertextuality both as a term and as a comment.[8] He had a BA in philosophy and a MA in sociology, and he realized "that, in order to write the essay that he wanted, a hypertextual authoring system that would connect the various parts of his work was needed." It was this that inspired him to launch project Xanadu, historically the first hypertext authoring system (cf. Ch. 3 for more information). It was also to Nelson that Tim Berners-Lee paid homage in the summer of 1992 when, while traveling in California, he made a trip to Sausalito, where Nelson was living on a house-boat. The meeting, as recounted in his book (Berners-Lee and Fischetti 1999, 70–72), was both cordial and a little strained. The fact that it happened at all shows the debt Berners-Lee felt he owed to that volcanic intellect who had given him the idea of hypertextual content. Recognizing this debt in his book, the inventor of the Web also acknowledged how his project depended on Nelson's humanistic background. From this perspective, the meeting between the two pioneers is witness to a constant osmosis between technology and the humanities, necessary to fertilize the fields of interdisciplinary study and make innovation possible. Landow likewise regards hypertext as a humanistic invention, as an embodiment of Barthes' ideal nonlinear text, or as a realization of the digital scholarly edition, where the reader follows from the main text to its variant readings and annotations (Landow 2006, 53–55).

At this point, rather than recount the subsequent history of the Web, which has already been the subject of dozens of books (Gillies and Cailliau 2000), it will suffice here to mention some key ideas that were crucial for its development.

Firstly, the original idea of the Web was that of an authoring system, in which the ability to edit as well as view pages was equally important. The design of this universal Web, in which everyone participated, has only come into being with the proliferation of systems such as wikis, blogs and social software that allow the user to browse pages and, at the same time interact with them editorially, often in a very simple and intuitive way. From this point of view the present development of the Web known as Web 2.0 (which will be discussed shortly) is simply, in some of its characteristics, a completion of the original project of information management from 1989. In all the interviews and lectures on the origins of the Web, its inventor always points out with great solemnity the absolutely central role of the community. And, as Berners-Lee underlines, in order to work like this, the community must feature the most diverse specialists, from scientists to sociologists, from jurists to biologists, etc. The community, therefore, is the same transdisciplinary community described above. One can see how this project is relevant for the digital humanist as a writer, curator and organizer of knowledge, its relationships and connections. It is evident from the original documents about the Web that the objective was to define a mechanism of an editorial nature that would favor content, and the ability to make it easily accessible in an organic fashion through associative pathways.

But while many of the properties of the Web were suggested by the existing infrastructure and ideas circulating at the time, it is also true that Berners-Lee demonstrated his own special skills in connecting the various components. The three elements that distinguish the Web were defined in his book:

> The art was to define the few basic, common rules of "protocol" that would allow one computer to talk to another, in such a way that when all computers everywhere did it, the system would thrive, not break down. For the Web, those elements were, in decreasing order of importance, universal resource identifiers (URIs), the Hypertext Transfer Protocol (HTTP) and the Hypertext Markup Language (HTML). (Berners-Lee and Fischetti 1999, 39)

He insisted, therefore, that the Web was simply a space where information could exist, be produced and connected. The realization of the project thus transformed the status of the most important element of its constituent parts. From the idea of the universal identification of the resources, there came a more prosaic "uniform resource locator"[9] (URL), a label that did not define a given resource in a universal way, but, more concretely, identified its location in a uniform and therefore unique manner. The next two stages of the development date from 1993. On 30 April of that year, CERN signed a document of just two short pages[10] agreeing with the legal department of the institution, in which they promised not to claim any royalties on the technological protocols developed by Tim Berners-Lee and his colleagues that were needed to run the Web. This was a historic step, because the Internet was beginning to take off at that time in the United States and the Web could use that infrastructure as an enabler for its development, provided that anyone could download programs and protocols that allowed machines to understand and connect hypertextual documents. To achieve this goal another small piece of the puzzle was needed: the creation of a browser, a hypertext-page viewer that was platform independent (i.e. independent from any operating system). Berners-Lee first sought European collaborators. He found a group who were working in the INRIA laboratory in France on a hypertext system written in SGML (standard generalized markup language) and asked them to develop software to display pages in HTML. The group, however, before beginning, sought assurances on the presence of European funds to finance the project. "They did not want to risk wasting time…" (Berners-Lee and Fischetti 1999, 49). Finally in February of 1993, the NCSA (National Center for Supercomputing Applications) at the University of Illinois made available a suitable browser, with all the most popular characteristics, easy to download and install and available for all platforms. It was Mosaic (the precursor to Netscape/Firefox), which was developed by a student, Marc Andreessen, and a staff member, Eric Bina. The Web now had almost everything it needed to grow and spread. It lacked only one last thing: governance. Thus, in 1994 Tim Berners-Lee agreed, after careful reflection, to move to MIT, at Princeton in the United States, to found an international institution to guide the Web, the *W3Consortium*, whose purpose was to "govern the World Wide Web to its full potential by developing protocols and guidelines that will ensure the long-term growth of the Web."[11] The consortium still unites institutions, countries and private companies operating on the Web and proposes and establishes rules for civil coexistence. Its objectives include the definition of standards accepted by the entire community to ensure its homogeneity.[12] It also focuses on the preservation of the heart of the Web as a public thing, a collective good, since there are no royalties to pay to the inventors. Tim Berners-Lee is still guiding the consortium, and retains all the authority it had when he established it with a few other collaborators. Even now that the WWW has matured and is autonomous, he continues to believe that it needs the coordination and integration that only an international, non-profit organization can provide. In his speech to the 20th anniversary conference,[13] he pointed out the primacy of democracy and consensus, and confirmed that the W3C was working to increase sharing and joint participation. In his speeches, he recognizes that the activity of the W3C is political, even when it deals with the definition of technical standards. He has stated quite clearly that "The Web, and everything which happens on it, rest on two things: technological protocols, and social conventions. The technological protocols, like HTTP and HTML, determine how computers interact. Social conventions, such as the incentive to make links to valuable resources, or the rules of engagement in a social networking website, are about how people like to, and are allowed to, interact" (Berners-Lee 2007, 7). The relevance of the social and human aspects of the Web cannot, in his view, be underestimated. For this reason, his projects involve, for example, the creation of institutions where scholars can meet to create "an intellectual foundation, educational atmosphere, and resource base to allow researchers to take the Web seriously as an object of scientific inquiry and engineering innovation" (ibid., 8). The Web Science project (http://webscience.org) starts from the premise that, for the future of the Web, it is necessary to adopt a "systemic" approach like that of biology, in which not only technological aspects are taken into account, but also effective strategies for understanding "social machines" in action (Hendler *et al.* 2008). In 2009, he further revised his position regarding the future of the Web and found that Web science was not enough: more than 80% of humanity, to which the Web is dedicated, is not connected; and many of the most popular Web-based tools are culturally focused on habits and customs of the average American, without regard to the rest of the world. To this end he is still advocating the creation of the WWW Foundation (http://webfoundation.org). Its purpose is to study the diversity of conditions in which the Web can be accessed: for example, cultural differences, (it does not make much sense for someone in an African village to interact with others using tools such as social networking, which are designed for young Westerners), technological differences (in developing countries access is, in most cases, via a mobile phone), and language differences.[14] If the *gap* that separates two thirds of the world's population from the Web cannot be bridged, all the effort in favor of humanity will have been wasted, and will end up bringing further discrimination. One aim of the digital humanist should therefore be to address all problems related to social and political traditions, and to help define a common ground for action.

2.3 Web 2.0 and beyond

As Tim Berners-Lee likes to say, the Web is not something that is completed: it is a constantly evolving tool that must be redesigned periodically to remain in the service of humanity. However, it is not always clear whether, and in what way, technology can remain in the service of humanity, rather than serve only one part of it, usually the richest, most efficient and most organized from a certain cognitive point of view, as Wiener had warned, about 60 years ago (cf. § 1.4). At this point it would be useful to temporarily suspend the discussion of history and to look at the present, and try to form hypotheses about what will happen in the near future.

"Web 2.0" is a successful label invented by Tim O'Reilly (2005) (cf. §§ 3.6–8). This was the name he gave to a series of conferences organized by his publishing house in 2004, and it was an unprecedented marketing success. Wherever one roams among Web applications, there is not one that cannot be understood in the framework of Web 2.0. Giving the Web a new title was intended to revive a sector affected by the collapse of the dot-com bubble at the beginning of the century, as a result of the excessive aspirations of Web services companies. So O'Reilly, using the well-known numbering system of major software releases that implies a simplification and resolution of problems, relaunched the businesses of the Web from a new perspective. The thread of his discourse, as traced by the Italian sociologist Carlo Formenti (2008, 248), describes a scenario that differs from the original as imagined by Berners-Lee. O'Reilly's Web seems to be a sort of caricature. The objectives of Web 2.0 were summarized in a few bullet-points: focus on the offering of services rather than software, consider the Web as an architecture of participation, develop strategies for the exploitation of collective intelligence, with particular regard to the possibilities of remixing services in new combinations. Questioned on the subject in 2006, the inventor of the Web said that Web 2.0 was a "piece of jargon," and that besides wikis, blogs and social networks (the focus of the Web 2.0 era) there were many other ways for people to collaborate and share content (Berners-Lee 2006). However, the instrumental nature of Web 2.0 and its commercial interests are fully transparent in O'Reilly's project. User-generated content is presented in different forms and organized to make it more attractive to the advertising market and other related business models. In summary, the common good, as represented by digital content, is put to the service of private business. This is a kind of capitalism 2.0, where whoever owns a platform for sharing information with friends, and can post videos and photos, will not have to worry about paying for content, and can sell advertising on the attention it generates, as well as widen the audience for investors. Online, in fact, you can buy and sell small amounts of advertising space, so that even small advertisers can have their own little place in the sun.

In an article published in *Scientific American* on the anniversary of the launch of the WWW, Tim Berners-Lee (2010) expressed his concern about the risks to which the project had been exposed in recent years. The lines of evolution considered most problematic are related to the development of the more commercial aspects of Web 2.0. He identified some critical areas for the universality of the service due to the tendency of some operators to take ownership of the content published by users on the platform (Facebook, LinkedIn, etc.) and to prevent them from being exported to other sites, even although they were produced by the service's users. Also at risk, in his analysis, is the place of open standards, which he believes are the only way to produce continuous innovation. The use of proprietary protocols such as that of iTunes to sell copyrighted music and videos is thus considered problematic. These proprietary tools not only lock up information and prevent the creation of links to protected data, but they also produce an even more threatening long-term effect: the interruption of investment in services that take advantage of open protocols. Indeed, if technologies become proprietary, they do not become standardized, and nobody wants to risk producing applications for proprietary protocols that are not widespread.

Another critical point raised by the inventor of the Web is the issue of Net neutrality. The original idea that each packet is equal before the network is set to crumble, not so much for the simple Internet connection, but certainly in mobile and broadband, where, as will be seen below, there exist situations of privilege related to the economic strength of various users. The loss of this principle would seriously harm the environment of the network as a place of equal opportunity for all users, if only from the theoretical point of view. In practice it is clear that, due to the topological structure of Internet and the practical facts of visibility, some content and some nodes are privileged above others (cf. Ch. 5). In short, the transparency of Internet access would be threatened not only by the mechanisms of network self-organization, but also by the will of some of its commercial players (see the Google-Verizon agreement, substantially accepted by the American Federal Communications Commission (FCC), [Stelter 2010]).

In conclusion, the phenomenon of Web 2.0 can be described critically as the progressive entrance into the field of Internet services of a new, and sophisticated class of brokers who earn money in their capacity as organizers of collective content. All of this takes place with the blessing of the content producers (i.e., *us*), whose personal details they are using. The era of "zero comments," as defined by the Net critic Geert

Lovink (2007) is fast approaching, when the writer on a network usually does not reach a position of visibility and recognition that would enable him or her to acquire the status of an "author." Web 2.0 is considered the realm of the amateur. There are almost no professionals and, when there are, they are treated as if they were not (i.e. they are not paid for their services) in a sterile celebration of the wisdom of crowds that simply becomes an excuse for a new generation of Web businesses that have no interest in developing ways to finance intellectual production. There is much here for the digital humanist to reflect upon. On the other hand, these new intermediaries have not yet removed the previous ones (the telecom companies, industrial producers of content, the large television networks and publishing groups) but have instead come alongside them, sometimes in conflict, but more often seeking agreements to establish some form of revenue sharing that would secure the interests of both sides.[15] Chapter 3 investigates the positive side of Web 2.0 as a set of tools for creating personal identity through writing and collective sharing. Another possible positive attribute of the phenomenon is the fact that through its instruments, it is now much easier to implement marketing strategies for individuals (the five minutes of fame that everybody can achieve after posting personal performances online) or for small groups who previously had to maintain visibility through traditional media. The first chapter of Clay Shirky's book (2008, 3–20), which tells how the services of Web 2.0 allowed a woman to recover her lost mobile phone, found by someone else on the seat of a taxi, by building a community who sided with her, is an exciting representation of the power of Web 2.0. It is beyond doubt that some of the tools of Web 2.0 can help people (famous, rich, or talented) to bounce off the traditional media, and then provide a springboard for a career in show business or other fields through the publication of suitable information. Indeed, it has been argued that the recent efflorescence of the digital humanities worldwide owes much to the development of Web 2.0 (Jones, 2014). On the other hand, the weak (the unknown, poor, and or less clever) will never be able to use these products to their advantage and their hopes will be easily crushed. And if success still passes through the traditional media, then what is so new about using the Web as a way of penetrating the agenda of the media circus?

An exception in the grand landscape of Web 2.0 can be made for the activities of social tagging or cataloging. These applications, such as Delicious, LibraryThing and Connotea, are services that allow users to establish collective descriptions, in the form of labels, or keywords (tags), for certain components of the Web, (e.g. pages, personal books, or digital resources).

All of these applications for the categorization and sharing of online content take the name of *folksonomy,* the contraction of the words "folk" and "taxonomy."[16] A successful example of a social filter, originating back in June 1998 (in very difficult times compared to Web 2.0), is the Open Directory Project (dmoz.org), which has become the Google directory. A special mention may be made of the use of the social Wiki technology to create a collectively-edited encyclopedia. The result of the project was wikipedia.org,[17] currently one of the most updated and reliable reference works, which has come to rival the most prestigious encyclopedias. There are also plenty of other examples of social software currently in use to access, organize and categorize Web resources.

However, there is another group of tools for archiving and retrieving information useful to the community. These allow the organization and sharing of the knowledge by each member of the community, provided there is some way to propagate trust and distrust (Guha *et al.* 2004). In social data-mining systems (cf. Ch. 4), which represent an advanced version of collaborative categorization, it is not even necessary for users to be explicitly involved in order to contribute. The system exploits the behavior of surfers to find information implicit in the description of their activities (Amento *et al.* 2003). An efficient and successful example of this strategy is the Amazon knowledge management system, the largest online retail shop. This site keeps track of customers' behaviors and uses the information to provide them with advice on any articles of interest, based on similar preferences by other users. This mechanism is particularly effective at producing suggestions consistent with the real interests of customers. The system also takes advantage of the sense of belonging to a virtual community based on the common interests of clients, who sometimes also actively contribute by writing reviews and offering assessments of books and other items for sale. These tools are more problematic because users do not know that they are "serving" the community, often for the profit of the mediators. In the case of Amazon, the aim is to offer those articles that fit as closely as possible to those desired by the user to maximize revenue. Categorization and collective filtering are a resource of the Web, which should be protected and defended from purely commercial interests, because they belong to the commonwealth of the Web.[18]

2.4 Leibniz's *Lingua Characteristica* and the Semantic Web

The machine built in the 1940s was the confluence of various ideas that came from far away. One can recognize in the debate about the intelligence of the machine the discussions of the 17th and 18th centuries between the rationalistic and empiricist approaches to knowledge and its creation. Could perhaps the machine solve all problems by calculating the solution, as Leibniz would have suggested? As mentioned above (cf. § 1.2), according to Gottfried Wilhelm Leibniz (1646–1716), one of the most prominent scientists, politicians and philosophers of his time, the best method to obtain certainty through knowledge was the creation of a system called *Characteristica Universalis,* which would allow all the people who used it to "calculate" the solution for all the scientific and philosophical problems. The system consisted of two modules; one was the *lingua characteristica,* a sort of universal language that permitted the expression in unequivocal form all the necessary and useful ideas in science or philosophy. The second module was called *calculus ratiocinator,* it was a method that allowed everybody to "deduce" via a calculus the correct conclusion for all possible premises that were expressed correctly using the universal language.

The use of this system, according to Leibniz, would avoid all possible mistakes and guarantee that all the conclusions were sound and true. The project was first envisaged when he was only twenty years old, but he kept on thinking of it for the rest of his life. In a letter to one of his many correspondents he declared:

> I am convinced more and more of the utility of this general science, and I see that very few people have understood its extent .... This characteristic consists of a certain script or language ... that perfectly represents the relationships between thoughts. The characters would be quite different from what has been imagined up to now. Because one has forgotten the principle that the characters of this script should serve invention and judgement as in algebra and arithmetic. This script will have great advantages; among others, there is one that seems particularly important to me. This is that it will be impossible to write, using these characters, chimerical notions .... An ignoramus will not be able to use it or, in striving to do so, he himself will become erudite. (Letter to Jean Galloys December 1678, translated from French in Davis 2000, 16)

In this letter, he showed the major advantages of the new "script" to his correspondent. First of all, it offered the guarantee that only the "real" concept could be represented in it, and secondly it forbade ignorant people to use it, or alternatively they would become savant in the effort to master the method. Such a language would also allow perfect correspondence of the relations among thoughts and would also help the user to have clear and correct thoughts, adequate both to the external world and to the true consequences of all axioms. All these results could be obtained by using a calculus similar to algebra or to arithmetic, which meant that once the notions were represented with the language symbols, it would be very easy to "calculate" the right conclusions. This project was only one on the long list of the dreams of reason by which human beings tried to control knowledge creation, by guaranteeing the correctness of every conclusion that was driven by correct assumptions. The birth of the computer and of the consequent "dream" of creating a mechanical intelligence could be considered just another scene of the same drama: the hope that truth and certainty were achievable exclusively by performing the right calculus. It will be shown shortly below that Semantic Web and AI share a lot of common beliefs.

On the other side of the epistemic range lies the work of David Hume (1711–1776), the philosopher who could be considered the champion of the empiricist tradition in the 18th century. He discussed knowledge and its characteristics in the first volume of *A Treatise of Human Nature.* Here, among other crucial questions, he stressed the central role of the association of ideas for knowledge creation, declaring:

> This uniting principle among ideas is not to be consider'd as an inseparabile connexion; for that has been already excluded

> from the imagination … but we are only to regard it as a gentle force, which commonly prevails, and is the cause why, among other things, languages so nearly correspond to each other …. The qualities, from which this association arises … are three viz. Resemblance, Contiguity in time or place, and Cause and Effect. (Hume 1978, 10–11)

The basic characteristics of the association of ideas are the contingency of the connections and the central role of imagination in the creation of the links between them. Both these principles are central in Bush's description of the operation of the mind as it accesses and connects thoughts together. The relationships between the association of ideas and logic are not necessarily rigorous. There are many different reasons why ideas unrelated to the logical inference between concepts cannot be connected. One of the consequences of the use of the association of ideas in the paradigm of the communication machine was the introduction of *hypertextuality* (see § 3.3) as a new writing method that allowed the association of different ideas to each other, without following a linear train of thought. The non-sequential writing model that was used in the human–computer interface research environment had a remarkable impact on the development of information technologies, whose consequences are still difficult to describe and foresee in detail.

According to Michael Mahoney (2005), the design of the computer was based on a confluence of contributions not only by different communities, but also from different philosophical approaches. Identifying these various influences and professional attitudes would be a challenging research project, but it is necessary to explain not only the history of computing but also the actual epistemological status of the various fields of information technology. The digital age did not come about as an achievement only of engineers, but represents the merging of many professional influences and models. Such an investigation may produce surprising results, and would in any case help in determining both the multi-faceted nature of the computing machine, as well as the opportunities, risks, threats and future directions of computer science.

In thinking about the development of the Web, it is important to recognize that different positions were being taken up that undoubtedly related to the general epistemological standpoints of rationalists and empiricists, who had different ideas about computers and the other devices that were being developed from the end of 1960s when the network was still in its infancy.

As described above (cf. § 2.2.2) Tim Berners-Lee's model of hypertextuality was inspired by Ted Nelson and Vannevar Bush, who believed that associations of (especially free-thinking) ideas, would lead to the creation of new knowledge, and enhance our mental capacity in building cognitive links between different scenarios. In spite of this influence, Tim Berners-Lee did not wed himself to the empiricist philosophy that lay behind it. In fact, at the same time as the Web was being launched, he admitted that he wanted to create another project, which was much more ambitious and philosophically quite different. This was the Semantic Web. The idea was, as he explained, to add a layer of logic to the Web (Berners-Lee 1998, Berners-Lee *et al.* 2001) that would identify every single online resource through a set of tags, or metadata, which would then allow the machines to "read and understand" the descriptive "semantic" layer of the Web. Under this scheme, the network would not be able to maintain its characteristic spontaneity in the publication of data and resources, but would have to be built as a database of structured information, organized according to specific types.

Today, this idea of categorization is completely different to the Web that Berners-Lee created. The idea of openness and serendipity is central to its working, founded on the free association of resources. In 2006, through the famous article *The Semantic Web revisited,* Berners-Lee and other authors reflected on how the future of the Web should look, and clarified their position. They admitted that "a Web of data" would be very different from the operation of the actual Web, first because of the presence of structured data, expressed in machine readable and context-sensitive formats, and second because intelligent agents could handle that data independently in a generalized way. They did, however, claim that the W3C and the Internet Engineering Task Force "has directed major efforts at specifying, developing, and deploying language for sharing meaning. These languages provide a foundation for semantic interoperability" (Berners-Lee *et al.* 2006, 97). These languages, such as RDF, SPARQL and OWL, like some languages used in artificial intelligence, can describe objects or events via properties and a function associated with each element (for example, if it was a number, an event, a film or a novel and so on). Each of these typological characteristics had in turn their own properties that had to be described in each instance of the type. To make them usable it was necessary to proceed with standards of composition that would allow the creation of ontologies suitable for any context in which you might want to provide a description. So this was just what Leibniz wanted: a representation of the world through a unique language, and an inference engine capable of extracting all the knowledge implied by each definition. The Semantic Web thus presents a precise vision of science and, more generally, of a system in which all knowledge can be described in a hierarchical manner starting from first principles, following a few simple rules of inference — a sort of a pyramid in which everything is organized, connected and perfectly consistent. However, that is not the way things always actually are. In the first place, the definition of standards for representing objects, events and ideas is not without its own problems. When a tag is chosen, using English of course, it represents a certain concept that in some way describes and therefore affects its content. From that moment on, the specification must be slavishly respected by all other actors in the process. The Semantic Web or Linked Data, as the project came to be known, can be considered as a system for data classification of online resources. At the start of any project of classification, it is assumed that it can guarantee some minimum standards, e.g. that: "there are consistent, unique classificatory principles in operation .... The categories are mutually exclusive .... The system is complete ...." However, one might also agree with the position that "No real-world working classification system that we have looked at meets these 'simple' requirements and we doubt that any ever would" (Bowker and

Leigh Star 2000, 10–11). If this is so, then it is very important that when designing a classification system one takes into account the arbitrary, culturally biased character in any organization. Tim Berners-Lee and his collaborators have strongly underlined the importance of openness in setting standards: "the construction of a standards body that's been able to promote, develop and deploy open standards" (Berners-Lee *et al.* 2006, 100). But one should not forget that they are also subject to the same cultural pressures as anyone else, and that such standards are also the fruit of specific historical and temporal situations, which perfectly reflect classification as a social and cultural act. A standard can be defined as "any set of agreed-upon rules for the production of objects. … It has temporal reach as well, in that it persists over time. … There is no natural law that the best standard shall win … Standards have significant inertia and can be very difficult and expensive to change" (Bowker and Leigh Star 2000, 13–14). These considerations should not be seen as good reasons not to create standards, but they do underline the need to remain aware of social conditions and the collective and constituent dimension they represent. The impression of the authors, however, is that the process of the Semantic Web/Linked Data has been activated without reflecting on the cultural aspects of each binding and limiting decision that produced the classifications that were adopted. The interpretation offered by the team gathered around Tim Berners-Lee of the activities of categorization seems devoid of any problematic aspects. It appears to be objective and free from any reference to the cultural or political issues connected with the social group most strongly represented by the bodies responsible for establishing the agreed names for categories, known as types. "Areas such as epistemology and logic are to some extent operationalized in computers and computer infrastructures. Knowledge representation and ontology engineering are about trying to capture aspects of shared conceptualization" (Berners-Lee *et al.* 2006, 101). Here is where Leibniz's universalizing dream resurfaces, with the idea that it is possible to "operationalize" logic and epistemology without the impact of time, of the history of subjectivity that enables practice. It would be too complicated here to account for the close relationship between the epistemological project of artificial intelligence that emerged and developed in the 1960s to the 1980s and that of the Semantic Web/Linked Data. However, the connections are explicit:

> AI will be one of the contributing disciplines. AI has already given us functional and logic programming methods, ways to understand distributed systems pattern detection and data mining tools, approaches to inference, ontological engineering and knowledge representation. All of these are fundamental to pursuing a Web Science agenda and realizing the Semantic Web. (Berners-Lee *et al.* 2006, 101)

Although the universal dream of the Semantic Web is far from being achieved, one might recall that in Spanish, the same word is used for sleep and dream: *sueño*. In interpreting Francisco Goya's ambiguous phrase "*el sueño de la razón produce mon-*

struos" one could substitute *sleep* by *dream* to suggest that monsters are also produced when the dream of reason comes true.

2.5 Social and cultural inequalities on the Web[19]

The preceding paragraphs attempted a critical assessment of Web 2.0, and argued that humanists should be aware of both its creative applications and its potentially manipulative agenda.[20] Before concluding this second chapter, however, it is time to take a broader view of the Internet, and to consider some of the issues that stand in the way of a more democratic and genuinely multicultural development of digital humanities. This may be termed the *Digital Humanities Divide.* It breaks down into five interconnected problems:

1. A digital divide may exist within or between countries, and possess different internal and external dimensions, e.g. geographical, sociological, economic, cultural, etc.;
2. the governance of digital infrastructures (from local institutions to worldwide organizations, like ICANN, IETF, IAB, W3C, etc.);
3. the development of standards (again, from large organizations like the Unicode consortium to more focused and smaller scholarly communities like TEI[21]);
4. the "code hegemony," i.e. the semiotic and technical dominance of multinational private groups, from Microsoft to Google, from Apple to FaceBook;
5. and finally, how all this relates to problems of governance structure, multicultural and linguistic issues, gender, and the representation of minorities (including alternative methodological views) within current DH organizations.

The main issues regarding point 4 will be discussed in Ch. 5, but the other problems will be dealt with here in order.

2.5.1 The digital divide

As described above, since Licklider and Taylor (1968) introduced the concept of the computer as a communication device, it was clear that, although the potentialities of the tool were tremendous, the central critical issue was how access could be controlled and distributed. Although in the US[22] and in the rest of the Western industrialized world, the access problem seems to be at least reasonably resolved, other criteria must still be taken into account in assessing the digital divide.[23] By the end of 2015, the Internet has an estimated 3.2 billion users: about 40.3% of the world's population, according to UN agency ITU,[24] and 40.7% according to the World Bank. In 2014, approximately 83.8% of people living in the 27 high-income OECD member states had access to the Internet.[25] In comparison, 50.2%, 38.3%, and 19.2% of the populations of Latin America and the Caribbean, the Middle East and North Africa, and Sub-Saharan Africa, respectively, had access to the Internet.[26] The data gathered by the World Bank (see Figure 2.1, overleaf) provides a snapshot of the situation by

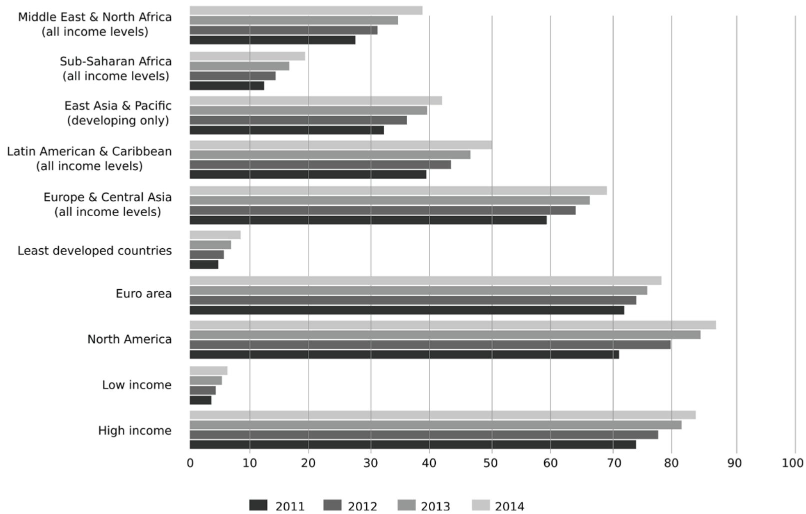

Figure 2.1. Internet users per 100 people in selected geographical areas (http://databank.worldbank.org/data/reports.aspx?source=2&Topic=9)

geographical region and income level, and shows the existence of a persistent and expected digital divide at the global level. However, as already mentioned, the digital divide cannot be reduced to a mere economic inequality between states, but is an internal problem within each individual state, based on age, education, type of government, ethnic group, etc., as is clear in the case of the US. The bitter conclusion of two experts like Witte and Mannon is that a technology "designed to be decentralized and democratic ends up maintaining and even expanding inequality" (2010, 127).

Data released in 2015, provided by the ITU, shows that Internet penetration in developing countries stands at 35%, but is only 10% for least developed countries. In developing countries 34.1% of households have Internet access, in contrast with 81.3% in developed countries and only 6.7% in least developed countries. As regards mobile broadband, the percentages are 86.7%, 39.1% and 12.1% respectively. In the case of Africa, with an estimated population in 2015 of 1.166 billion people (about 15% of the world's population), only 0.5% have a fixed broadband Internet subscription, although 17.4% have access to mobile broadband. However, the digital divide is not simply the result of geo-economical inequalities. According to Eurostat,[27] Europe seems to be a good example of issues relating to both access and exploitation of the Net. In the level of Net penetration, a clear dividing line can be drawn between Nordic countries, including Germany and France, on the one hand, and southern and eastern countries on the other (i.e. Italy, Spain, Greece, Portugal and Romania). Northern countries show a percentage of Internet usage similar if not superior to the US, while southern and some eastern European countries lag seriously behind. The EU Information Society database provides the relevant data.

Starting with Figure 2.2, there is a distinct gap between northern and southern European countries in the number of households with Internet access. In the second

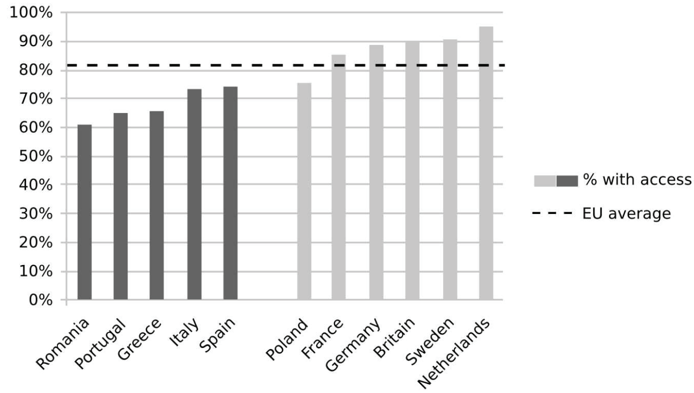

Figure 2.2. Percentage of households with Internet access (http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode=tin00134&plugin=1)

category "Individuals — Internet use in the last 12 months" (Figure 2.3 overleaf), Romania drops to 54%, Italy to 62%, and Poland to 67%, while almost all the other countries mentioned increase their percentages or remain the same. The values for weekly access are lower: Greece and Italy 59%, Portugal 61%, Spain 71%, France 80%, Germany 82%, Britain 89%, etc. The final interesting data worth mentioning are the interaction of citizens with their governments via the Internet, as shown in Figure 2.4 (overleaf).

Here the division between southern and northern Europe increases: all the old continent falls behind the EU average, while the northern countries pull ahead.

In conclusion, not only do the differences among European and OECD countries remain large, but age, income, education, family structure, and gender, in individual countries, all play a role in determining computer adoption, Internet access, and the level of digital literacy (Dobson and Willinsky 2009, 295–298). By "digital literacy" is meant not only basic computing skills, but also what Jeremy J. Shapiro and Shelley K. Hughes in 1996 called *information literacy*:

> A new liberal art that extends from knowing how to use computers and access information to critical reflection on the nature of information itself, its technical infrastructure, and its social, cultural and even philosophical context and impact — as essential to the mental framework of the educated information-age citizen as the trivium of basic liberal arts (grammar, logic and rhetoric) was to the educated person in medieval society. (Shapiro and Hughes 1996)

But in order to realize the ambitious curriculum of Shapiro and Hughes, the multiplicity and pitfalls of those digital divides must be addressed. According to Witte and Mannon "in a country in which some form of Internet access is becoming com-

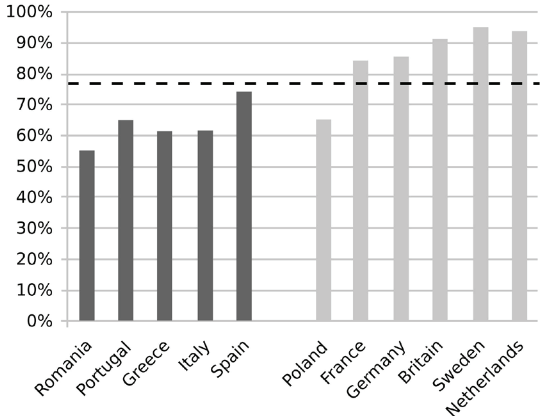

Figure 2.3. Percentage of people who accessed the Internet (http://data.worldbank.org/indicator/IT.NET.USER.P2)

mon, we also need to examine how individuals participate in and benefit from the Internet in distinct ways" (2010, 145). Their analysis seems particularly useful to the DH scenario. The authors combine three different perspectives to define the effects of inequality produced or introduced by the use of the Internet in US society. Only two of the three proposed views will be examined here: the Marxist and cultural perspectives. The former argues that inequality is not only preserved, but increased by the Internet habits of different social and cultural groups. The Marxist vision rests on the idea that in a capitalistic society the dominant class uses its assets to increase and maintain its advantage with regards to production:

> Recent theorists define skills as a kind of asset. In today's information-based economy, Internet access and use can be understood as an asset used to maintain class privilege and power. Second, capitalist relations of production can only be maintained if the inequalities upon which they rest are reproduced from one generation to the next. (Witte and Mannon 2010, 81)

Turning to the cultural perspective (inspired by the German sociologist Max Weber), and the extent to which education and income affect Internet literacy, Witte and Mannon underline that

> … occupational prestige and family background channel individuals into differential lifestyles, which in turn mark culturally enduring social divides. One of those cultural markers is Internet use. … Better-off and better-educated Americans left online footprints many time larger than the poorest and the least educated segments of American society. Moreover the online

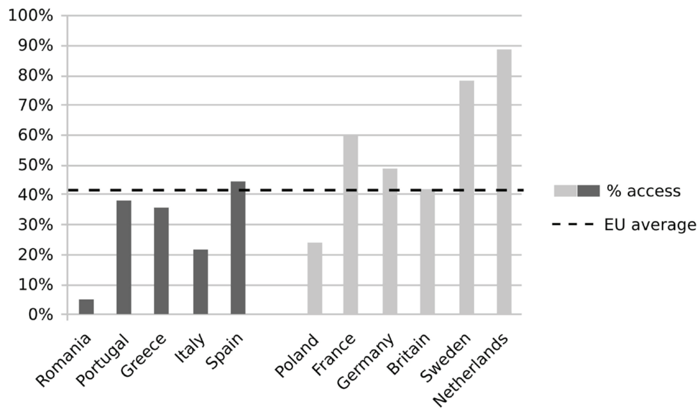

Figure 2.4. Internet use: interaction with public authorities (http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode=tin00105&plugin=1)

> footprints for more privileged members of American society were more extensive, indicating online activities associated with consumption and production, as well as information and communication. (ibid., 113–114)

It can thus be concluded that the cultural and socio-economic characteristics of a population have a direct impact on the visibility, efficacy and pervasiveness of Internet use. If this perspective is accepted, some examples of the theoretical, cultural and political biases lying at the core of the nature and origins of the Internet may be given, since these have the potential to increase the asymmetry of the network and its players, including the world of DH.

2.5.2 Geopolitics of the network

The revelations of former NSA contractor Edward Snowden in 2013 have shown to a global audience, among other things, the *geographical* dimension of cyberspace, in other words that "where technology is located is as important as *what* it is" (Deibert 2015, 10). Governance and standards go hand in hand, and there is always a *symbolic* level implied in a *political* (let alone *technical*) decision: "classifications and standards are material, as well as symbolic," and their control "is a central, often underanalyzed feature of economic life" (Bowker and Leigh Star 2000, 15; 39). In their studies Bowker and Leigh Star show how the classification techniques (and the standards generated from them) have always played a fundamental economic and socio-cultural role. Current digital technologies standards appear to be the result of a double bias: the technical one and the cultural one (geopolitical). These two biases are entangled and it is almost impossible to discern where the technological choice begins and where the cultural prejudice ends. Although the socio-cultural origins of the Internet have often been discussed (cf. Ch. 3), the impact of these origins, and their symbolic and cultural implications are rarely taken into consideration. Thus, the vast cultural consequences of tacit choices made by a group of English-speaking pioneers often pass unnoticed. And once again, *where* we are tells often *what* we do. An example is the structure of the addressing code rules that use the 128 ASCII (American standard code for information interchange) to describe all the servers on the Internet. The same techno-cultural bias affects most of the services and instruments of the network, such as the domain name. In the last forty years it has not been possible to use accented vowels in the URL address, and in spite of recent IETF and ICANN efforts[28] the new internationalizing domain names in applications (IDNA) system can only be implemented in applications that are specifically designed for it, and it is rarely used in Latin alphabet-based URLs. Some of the initial top-level domains can be used only by US institutions. For example, a European university cannot use the top-level domain .edu, which is reserved for US academic institutions.[29] The domain .eu was only added in 2006, and applications for top-level domains using characters outside ISO-Latin were only recently invited (requests were open from 12 January to 12 April 2012). ICANN (the Internet Corporation for Assigned Names and Numbers) finally allowed the opening up of top level domains to Arabic or Chinese characters, included in Unicode, but every decision has so far rested in the hands of an organization under the clear control of Western industries and governments (Hill 2015). The request procedure is very complicated, many of the rules are described only in English, the cost of an application for a top-level domain is $185,000, and the application does not guarantee that it will be granted. The applying institution needs to show a clear technical and financial capability that must be certified at the discretion of ICANN itself. The problem is that ICANN, a "not-for-profit public-benefit" corporation, has always taken decisions of global relevance, but still lacks a clear institutional and multi-stakeholder accountability.[30] According to Richard Hill, president of the Association for Proper Internet Governance and former ITU senior officer, "for the most part the narratives used to defend the current governance arrangements are about maintaining the geo-political and geo-economic dominance of the present incumbents, that is, of the US and its powerful private companies" (Hill 2015, 35).[31]

The Domain Name System (DNS), one of the technical backbones of the network, is another example of the "centralized hierarchy" of the Internet. In March 2014, in the wake of Edward Snowden's revelations, the United States agreed to relinquish control of the DNS root zone. But even now, the system of root name servers, according to the IANA Web site, is still operated by thirteen organizations: ten US-based institutions or private companies (led by VeriSign), two European entities, and one Japanese company. As noticed by Laura DeNardis, "there is a physical geography of the Internet's architecture as well as a virtual one", and "root servers are the gateway to the DNS so operating these servers is a critical task involving great responsibilities in both logical and physical management" (DeNardis 2014: 50). Despite this change, the geopolitical set-up of the DNS root zone is still based on a network governance paradigm of the "West and the rest". Many of the Internet standards were set at the beginning of its history and were obviously conceived by, and made for conformance with the small community involved at the time. The character set standard was clearly designed from a legitimate mono-cultural point of view, but today we are still dependent on those choices as an international and multicultural community. The data for Internet access[32] reveals that users in the Western world (Europe and the US) represent only 34% of the total, while Asian users represent 45%. However, as suggested by Tim Berners-Lee in the presentation of his Web Foundation, "creation of locally-relevant content on the Web is impeded in many places, not by lack of the Internet, but by a lack of knowledge."[33] Figures suggest that these differences are due not only to a lack of skill and competence of people in developing countries, but also to a specific cultural orientation of contents and opportunities. Although it is true that Internet adoption in Africa and Middle East is not comparable to that of developed countries,[34] it has been widely acknowledged that the 2011 Arab spring benefited from the use of social media such as Twitter, particularly in Tunisia (Howard and Hussain 2011; Bettaïeb 2011; Meddeb 2011), and Facebook, particularly in Egypt (Ghonim 2012). It is not possible here to discuss the opposing views of cyber-utopians and cyber-sceptics about the so-called "Facebook revolution in the Middle East,"[35] however it is clear that the role of social networking was a crucial element (although not the only cause) in organizing and informing people about what was going on in the streets, even if the same tools were also used by antidemocratic political powers to trace and repress their opponents.[36] Social networks certainly cannot be ignored after what happened in 2011 in North Africa or in 2013 in Turkey (Durdağ 2015), but what should be kept in mind is that the role of technology is always mediated by the people and their capabilities to transform the potentialities of the tool (Etling, Faris and Palfrey 2010). It is still a matter of skill to exploit the medium for revolutionary purposes, not a property of the technology itself.

2.5.3 The value of cultural and linguistic diversity

Although today Chinese and Spanish are increasingly used on the Web,[37] access and control of the Internet are firmly in the hands of select Western (and mainly anglophone) authorities. Discussions on identity, ethnicity, gender, etc. on the Internet abound (Siapera 2010, 183–197), but the mix of technical, methodological and linguistic biases of Internet resources and tools defy current analyses. José Antonio Millán is a linguist, net analyst and Spanish blogger who left university twenty years ago to dedicate himself entirely to the study of digital textuality and digital media. His blog "Libros y bitios" (http://jamillan.com/librosybitios) is known as one of the best online resources in the Hispanic DH world. Millán in 2001 published an important book, which is still a valuable source of information, and at the same time an effective manifesto of the "digital margins" of the world.[38] His work helps to substantiate with researched examples the geopolitical scenario outlined above, by closely analyzing the production and spread of all Internet technologies that concern language. According to Millán, there are many products and services which derive from these technologies, all of them of strategic value, and all in "alien" hands: operating systems, search engines, intelligent agents, distance learning, electronic commerce, the copyright industry, etc., Each of these areas presupposes or stimulates specific research sectors. These range from automatic translators to syntactic parsers, from terminological databases to software for speech recognition, etc. Even though the estimated burden of linguistic technology for each product and service analyzed is low (see Table 2.1), the result is astonishing: for an audience of 61 million Spanish speakers, the annual business turnover was estimated in 2001 to be something like 91 million Euros (Millán 2001, 148–149).

Product/ service	**Weight in linguistic technology**
Electronic Commerce	0.01
Copyright industry	0.01
Tourist information services	0.03
Operating Systems	0.05
Distance Education	0.07
Word processing	0.10
Teaching material for Spanish as SL	0.10
Information services (non-touristic)	0.10
Editorial platforms	0.20
Search engines	0.30
Information managers	0.50
Intelligent agents	0.80
Teaching software	0.80–0.90
Terminology assistant	0.90
Translation software	0.90

Table 2.1. Linguistic technologies and products: weight per product or service. In this list the author omits the technology of voice recognition, which is dealt with in Ch. 9 (Millán 2001, 134)

The author concludes by saying: "while networks are the highways of digital goods and service flows, technologies linked to the user's language are their compulsory tolls" (Millán 2001, 140). Thus, at the roots of economic, social, political primacy one does not find "just" technology, but rather the mix of copyrighted algorithms and protocols that manipulate and control languages. Presiding over both natural and artificial codes has become a profitable business: not investing in this sector presently means being forced to pay to be able to use one's own language. Unfortunately, the problem of cultural primacy overflows linguistic boundaries: the pervasiveness of cultural representations and metaphors belonging to the Anglo-American context in all technological appliances and computing tools is a well-known tendency since at least the 1960s. Many familiar elements borrowed from everyday US and Western life were exported to the computer world. Beyond programming languages or algorithms, where deep semiotic and cultural biases are intrinsically evident (An-

Figure 2.5. "Kiss my butt" gesture in Second Life

dersen 1997; Chun 2011; Kittler 2008), ideologies extend to the "superficial" (and not less subtle) world of icons and graphical interfaces (Selfe and Selfe 1994; Ford and Kotzé 2005; Galloway 2012). One example is the manila folder, a ubiquitous object used in all American offices that owes its name to a fiber (manila hemp) commonly used in the Philippines for making ropes, paper products and coarse fabrics. An object coming from a removed colonial past suddenly, thanks to the Xerox Star desktop,[39] became later the metaphor for any computing content: a symbol that conceals the bureaucratic origins of the desktop computer and its unique ties to the cultural imagery of the average US customer. Examples of symbolic digital colonization are Second Life facial expressions and user-playable animations, where we can find body language gestures which can be only deciphered by expert American native speaker.[40] Take for example the famous "kiss my butt" animation (see Figure 2.5), where both the verbal expression and the body posture would suggest (at best) deceptive or vaguely alluring meanings to most of Latino or Mediterranean cultures.

This list of aggressive US iconic settlements in the global world could continue, but a more important example of representational bias that directly affects the work of humanists, is the important work carried out by the Unicode consortium.

First, two words about the organization itself. Unicode is a non-profit organization "devoted to developing, maintaining, and promoting software internationalization standards and data, particularly the Unicode Standard, which specifies the representation of text in all modern software products and standards."[41] The Board

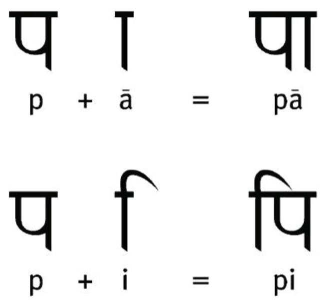

Figure 2.6. Two graphames of Devanagari Indic script as shown in Perri 2009, 735

of Directors of Unicode is currently made up of people from Intel, Google, Microsoft, Apple, IBM, OCLC, and IMS Health.[42] Not very different is the make-up of the Executive Officers (the president is a Google engineer since 2006). Apart from one or two exceptions, universities, public or research institutions are not represented. As a matter of fact, Unicode is an industrial standard made and controlled by industry. And claims about the geopolitical neutrality or impartiality of this organization appear to be at least questionable.[43] Localization still matters, and the researchers of the Language Observatory Project (http://www.language-observatory.org/) noted that, although Unicode is recognized as a step forward for multilingualism, "many problems in language processing remain":

> The Mongolian language, for example, is written either in Cyrillic script or in its own historical and traditional script, for which at least eight different codes and fonts have been identified. No standardisation of typed fonts exists, causing inconsistency, even textual mistranslation, from one computer to another. As a result, some Mongolian web pages are made up of image files, which take much longer to load. Indian web pages face the same challenge. On Indian newspaper sites proprietary fonts for Hindi scripts are often used and some sites provide their news with image files. These technological limitations prevent information from being interchangeable, and lead to a digital language divide. (Mikami and Kodama 2012, 122–123)

The Italian linguist and anthropologist Antonio Perri has offered convincing examples of the bias inherent in the Unicode system for representing characters, showing the concrete risks of oversimplification and erasure of the "phenomenological richness of human writing practices" (Perri 2009, 747). Perri analyzed a number of encoding solutions proposed by the Unicode consortium for different problems relating to Indian sub-continental scripts, to Chinese, Arabic and Hangul (Korean writing). In all these cases, in addition to being excessively dependent on visualization software, which raises problems of portability, he showed that the Unicode

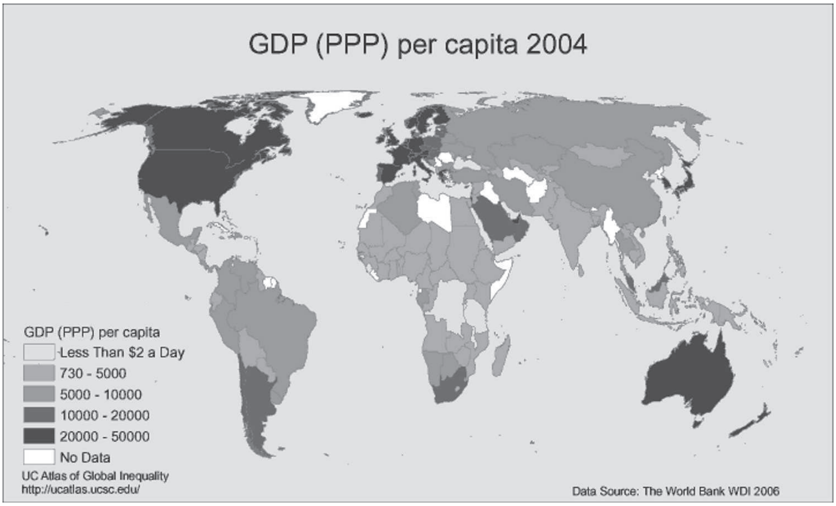

Figure 2.7. World Gross Domestic Product in 2004. (http://ucatlas.ucsc.edu/)

solutions were based on a "hypertypographic" concept of writing, i.e. Western writing embodied in its printed form and logical sequencing. By neglecting the visual features of many writing systems this view overlooks their important functional aspects. Perri gives a striking example of this bias when discussing Unicode treatment of ligatures and the position of vowel characters in the Devanagari Indic script.[44] Often in Indian systems aspects of a graphic nature prevail over the reading order of the graphemes.[45] As showed in Figure 2.6, in the second glyph the order pronunciation/graphic sequence is reversed. Unicode experts, however, argue that Indic scripts are represented in its system according to a "logical scheme" that ignores "typographic" details.[46] Perri concludes:

> But why on earth should the order of characters corresponding to the phonetic segment be considered logical by an Indian literate? Who says that the linearity of Saussure's alphabetic signifier should play a role in his writing practices? ... It is therefore all too evident that the alphabetic filter, the rendering software and the automatic process of normalization of Indic scripts are the result of a choice that reflects the need for structural uniformity as opposed to the *emic* cultural practices of the real user. (Perri 2009, 736; our transl.)

One last example is a comparative experiment based on two graphic representations. The first image (Figure 2.7, overleaf) is a map of world income inequalities from the University of California Atlas of Global Inequality database. The second world map (Figure 2.8, overleaf), a Wikipedia image based on Ethnologue.com sources, repre-

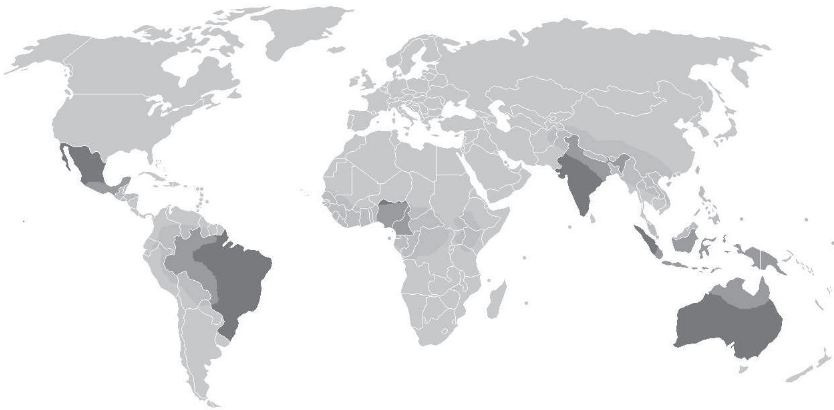

Figure 2.8. Linguistic diversity in the world (http://en.wikipedia.org/wiki/Linguistic_diversity#Linguistic_diversity)

sents linguistic diversity in the world: in red are shown the 8 megadiverse countries that together represent more than 50% of the world's languages, and in blue, areas of great diversity. By overlaying these two maps, one can notice that — excluding Australia, where the linguistic diversity is in any case that of the aboriginal inhabitants before settlement by the British in the 18th century[47] — the lower income countries of the first map in many cases fit the areas of greater linguistic diversity. Over the past decade the idea began to emerge that linguistic diversity is inextricably linked to biodiversity (Maffi 2001). The Index of Biocultural Diversity adopted in 2004 by a group of interdisciplinary scholars, who later would found the international NGO terralingua.org, shows that three areas of the world "emerge as 'core areas' of exceptionally high biocultural diversity: the Amazon Basin, Central Africa and Indomalaysia/Melanesia" (Maffi 2010: 6).

> Instead, a lowering of both cultural and biological diversity has been found to correlate with the development of complex, stratified and densely populated societies and of far-reaching economic powers. ... From ancient empires to today's globalized economy, these complex social systems have spread and expanded well beyond the confines of local ecosystems, exploiting and draining natural resources on a large scale and imposing cultural assimilation and the homogenization of cultural diversity. (Maffi 2010, 8)

In other words: cultural and biological richness does not necessarily match material wealth.[48]

Unsurprisingly, the world income map also overlaps with the "Quantifying Digital Humanities" infographics flatland produced by the UCL Centre for Digital

Humanities.[49] Although that survey was generated from spontaneous inputs from anglophone scholars, and not from a systematic research and data collection, it is interesting to note how powerful is the tendency of this community to self-represent itself as The Digital Humanities world, without any kind of geographical, linguistic or cultural restriction. As observed by Chan (2014), these kinds of universalizations are very common in the West.[50] As already pointed out, regions and countries with ubiquitous Internet access tend to think of the rest of the world as having similar coverage, and this "first world" outlook contributes to the reassurance of representing the entire globe.

The issue of the over-representation of anglophone institutions and people in the DH international organizations has been discussed in Fiormonte 2012 and Dacos 2013 (cf. Conclusions). Most of the "international" organizations are monolingual, and the rhetorical structure of their websites and official documents does not leave space for anything except the "inner" Anglo-American rhetoric and academic narrative (Canagarajah 2002, 109–27). All this seems to confirm Millán's hypothesis of the strict relation between economic hegemony, technological concentration and linguistic impoverishment, and raises the as yet untackled question of the internal and external *digital humanities divide* in Western countries.

Although confirmed by recent studies (Amano *et al.* 2014), the comparison between the two maps proposed here does not intend to suggest easy conclusions. However, it is legitimate to hold that in some of the poorest areas of the world, in the deserts, jungles, and mountains at the margins of our globalized society, a handful of communities continue to cultivate the last resource still entirely in their own hands: biological and cultural diversity. The significance of language diversity and of its present loss in a digital world cannot be underestimated (Harrison 2010). Digital humanists cannot ignore this multiform dimension if they want to build inclusive digital resources and tools, and become more conscious about their role as knowledge gatekeepers and producers.

2.6 The challenge of open knowledge

The issues of inclusivity, access and diversity call for possible solutions. Is the variegated galaxy of *open knowledge* a good candidate for addressing some of those problems? As we will see, many DH projects embraced it successfully, and from many points of view the idea of knowledge as commons (Hess and Ostrom 2007) can be a powerful antidote to some of the negative effects of globalization. But let's start from the beginning of the practice, as abstract definitions will not help here.

In 2006 Ilaria Capua, an Italian veterinarian and researcher, refused to share her research about the H5N1 avian influenza virus with a private database in Los Alamos (NM). Instead, she decided to release her findings to the public domain, and later helped to launch the database *Global Initiative on Sharing Avian Influenza*

Data (GISAID) (Armstrong Moore 2009). Her decision attracted the attention of the world's media, and in 2008 she was included among the five "Revolutionary Minds" by the American magazine Seed. Today she is considered one of the leaders of the open science movement, based on the idea that all scientific data relevant to society (including but not limited to public health) should be openly accessible. This chapter will sketch a brief overview of the influence of this movement on the digital humanities, and show how some aspects of the open science/open data remain controversial. One problem is that there are many overlaps between open knowledge, open data and open science, and sometimes it is hard to separate the genuine impulse of the scientific community to share its knowledge from the commercial interests of companies that have embraced the open content model.

Web 2.0 may be a problematic term, but many of its positive aspects can still be welcomed, as agreed even by the once skeptical Tim Berners-Lee (cf. § 2.4). Previously, only documents were shared online. Now it is also possible to share other forms of data. The open data movement proposes making data from all areas, from chemistry to genetics, from medical trials to physics experiments, from grammar to geography available to all, open and non-proprietary.[51] There are already some shared databanks, and others are being developed.[52] The project's aim is to produce quality content socially, to make it freely available to all, and to allow the data to be updated and cross-checked with other archives using the same architecture.[53] As mentioned earlier (cf. § 2.4), Linked Data is the new brand name for the Semantic Web, but it also reduces the idea to a form that is easier to implement. It is "a set of best practices for publishing and connecting structured data on the Web" (Bizer, Heath and Berners-Lee 2009, 1). While the Semantic Web proposed the addition of a semantic layer to make the Web "readable or understandable" by machines, Linked Data only considers data on the Web that is already in structured form. However, this change in perspective has not altered the basic principle, which is to continue to increase the amount of data, to label it consistently with "meaning," and then to find ways to make it interoperable. That Linked Data is a more practical and limited version of the Semantic Web can be seen from the official declarations of the project's working group:

> The first step is putting data on the Web in a form that machines can naturally understand, or converting it to that form. This creates what I call a Semantic Web — a Web of data that can be processed directly or indirectly by machines. Therefore, while the Semantic Web, or Web of Data, is the goal or the end result of this process, Linked Data provides the means to reach that goal. By publishing Linked Data, numerous individuals and groups have contributed to the building of a Web of Data, which can lower the barrier to reuse, integration and application of data from multiple, distributed and heterogeneous sources. Over time, with Linked Data as a foundation, some of the more sophisticated proposals associated with the Semantic Web

> vision, such as intelligent agents, may become a reality. (Bizer, Heath and Berners-Lee 2009, 17)

The bar thus appears to have been lowered. There is no more talk about adding a layer of logic to the entire Web, but only of making structured data already on the network interoperable, or working on the development of an appropriate vocabulary for schemas, or query tools. All this is evidently based on good will and the ability to build reliable networks for shared data (Bizer, Heath and Berners-Lee 2009, 19–20).

The availability of so much data on the Web, including personal information made available by social network users, has led to the development of other methods of data processing — for example, the analysis of user sentiment in relation to a certain brand or idea. These analyses, which are becoming less common, because of simpler tools such as the "like" button on Facebook or the +1 of Google+, make use of sophisticated linguistic computational analysis to evaluate the emotional orientation of the public to a new product or idea (Wilson *et al.* 2009). The management of user data derived from clickstream analysis, from active network participation, from personal information supplied about religious beliefs, political or sexual inclinations, literary and personal tastes, makes up the rich mine of Web 2.0 data. It can be exploited in various ways, but always with surprising accuracy by the biggest data collectors, who are also the most successful service providers. This data-mining business is a key activity of the social dimension of the Web, and one of the most interesting and risky novelties of business in the Web 2.0 world.

2.6.1 Big Data[54]

A new hype surrounds the "Big Data" phenomenon. According to some scholars (Barabási 2010; Mayer-Schönberger and Cukier 2013) access to such a huge amount of information will revolutionize the way scientific results are obtained, particularly in the field of social sciences and humanities. This promise is very attractive to media companies that store all the data, but alarming for users, whose freedom is threatened, not only in terms of privacy. In fact, the data mining techniques used to manage and interpret information derived from the digital footprints of users are based largely on the same techniques adopted to analyze corpora or to interpret texts by literary scholars in the digital humanities.[55] So digital scholars face a dilemma: do they want to participate in the epistemologically and ethically problematic activities related to the use of Big Data? This is still an open question with no easy answer.

The collected digital traces left by almost any human activity, such as organizing a trip abroad, or starting a love affair, will allow researchers to manage not only statistical data on a population, but people's real lives. According to other scholars (Boyd 2010, Chun 2011, Gitelman 2013), however, the excitement around the change of perspective of human sciences due to the manipulation of Big Data is completely overestimated. According to Boyd and Crawford "Big Data offers the humanistic disciplines a new way to claim the status of quantitative science and objective method. It makes many more social spaces quantifiable. In reality, working with Big Data is still subjective, and what it quantifies does not necessarily have a closer claim on objective truth — particularly when considering messages from social media sites" (Boyd and Crawford 2012, 667). So it is imperative for digital humanists to maintain their critical attitude towards those quantification techniques that appear to give their disciplines the appearance of objectivity. From the epistemological point of view, humanistic studies offer a privileged perspective to assess the awareness that "raw data" does not exist (Gitelman 2013). The reasons for the critical approach to big data are complex and various. One argument relates to the incompleteness and dirtiness of the data that form the basis of data-mining procedures. People are unaware that they are recording data on themselves when they participate in social networks, and, as a result, they may record false or incomplete information about themselves or their friends, which are then stored in the database and considered true. One of the reasons for this, according to Wendy Chun (2011, 93–94) is that people are always inconsistent in describing themselves, and any self-produced data design can only provide a misleading understanding of the subject and an inadequate prediction of his/her future preferences and actions. Others, like Jaron Lanier (2013), critique our current digital economy, making a case that links rising income inequality to the spread of what he calls "Siren Servers," or data-gathering companies:

> … progress is never free of politics … new technological syntheses that will solve the great challenges are less likely to come from garages than from collaboration by many people over giant computer networks. It is the politics and the economics of these networks that will determine how new capabilities translate into benefits for ordinary people. (Lanier 2013, 17)

In fact, Big Data raises a lot of critical issues relating to control and access to private information, as is clearly shown by the data protection saga unleashed by the publication of documents by whistle-blower Edward Snowden and the New York Times, Guardian and other media during the summer of 2013. The details of the multi-million dollar programs managed by NSA (National Security Agency) and its British equivalent GCHQ (Government Communications Headquarters) show that PRISM and other tools are used with the (overt or covert) help of the "big four" (Hotmail/Microsoft, Google, Yahoo! and Facebook). From email to texts, from mobile traffic to social network data, everything is collected and processed to prevent the potential risk of terroristic activities. The approach of English speaking intelligence agencies is based on the theory that it is better to know everything, than to miss information that may be relevant to a potential enemy action. The normal balance between the right to privacy and the right of executive power to protect society against violence has been completely subverted, given both the commonly misperceived level of risk in social networks, and the power of new brute-force decryption technologies to decipher formerly secret information. The opportunities offered by technology, together with the social perception of risks, has already changed the boundary between the permitted and illicit public exercise of power.

All these new applications of Big Data and intensive information extraction techniques raise enormous social and political issues that should be addressed by digital humanities scholars, considering that they are among the few experts who possess both a humanistic and a technological background.[56] Digital humanities skills provide a unique opportunity to assess the ethical and social constraints placed on information technologies by a society that aims to maintain a consistent and plausible sense to the term "democracy" (cf. §§ 5.5, 5.6 and 5.9).

It is hard to accept a tool for understanding social behavior that forces us to accept that "the current ecosystem around Big Data creates a new kind of digital divide: the Big Data rich and the Big Data poor. Some company researchers have even gone so far as to suggest that academics should not bother studying social media data sets" (Boyd and Crawford 2012, 674). Big Data sets out to change the rules of what was once the reproducibility of experimental results (regarded as one of the foundations of the scientific method) and the accessibility of data, which would provide the controls needed to validate those results. If the data in question are controlled by a few companies who do not need to share it, or who can decide who to share it with, then an aristocracy of data-owners will be able to define the rules and hypotheses of social knowledge beyond the scrutiny of any external authority. According to Kristene Unsworth, what is needed is an "Ethics of Algorithms" to ensure their neutrality with respect to responsible data use and interpretation.[57] As suggested by Wendy Chun, Big Data represents a challenge for digital humanists "not because they are inherently practical, but rather because they can take on the large questions raised by it, such as: given that almost any correlation can be found, what is the relationship between correlation and causality? Between what's empirically observable and what's true?" (Chun and Rhody 2014, 21). In his work Databasing the world, Geoffrey Bowker underlines that "the most powerful technology ... in our control of the world and each other over the past two hundred years has been the development of the database" (Bowker 2005, 108). Computers and networks have shaped the knowledge and standards of database design. But the greater the need to put together data from different fields and contexts, the greater the need to find an agreement, or a standard for the relevant form of that data. "The point is that there is no such a thing as pure data" (Bowker 2005, 116). With Big Data the problem is no different, although the discussion shifts to the different methods needed to process such a large amount of data. It is always necessary to find and define a context, a standard, and an interpretation hypothesis that makes sense of the data. According to Culturomics[58] and distant reading of books, the fact that we have millions of books that can be read together changes the perspective and the possibilities for interpretation. For example, Ngrams (an application that can be used to find the statistical frequency of terms within the corpus of Google Books), which is one of the tools used within the Culturomics project,[59] enables the distant reading of many books arranged in historical or linguistic categories. Although it seems necessary to keep a critical eye on this project, the applications proposed by researchers in the field of Culturomics are new and promising.

2.6.2 Open data and the humanities

Open data is, according to the Open Knowledge Foundation,[60] "data that can be freely used, reused and distributed by anyone." But for this to work in the digital humanities there must be freely available tools to operate on that data in typically used formats.

One of these is TaPoR,[61] an open-source set of text-analysis services, which can be used either online or in-house. Another suite of tools for reading and analyzing texts is Voyant,[62] which can be used on any corpus of textual data to obtain a quick summary of the rhetorical characteristics of the corpus, including word-frequencies or word-trends within a text. The interesting perspective adopted by this tool-suite is that it focuses on representing texts not only at their intrinsic textual level, but also at the metalevel. One of its objectives is to study how computer-assisted text analysis works.

But because the format of data is often dictated by the needs of the research that gave rise to it, even open data is not always as open in practice as it is supposed to be. One example is the so-called e-Government 2.0 delivery model, which represents the results of transparency and collaboration by public institutions and governments with its citizens. The US federal administration in 2009 launched a portal for participation in the life of Congress, where it became possible to leave comments and discussion on the laws under debate (http://www.regulations.gov). In the summer of 2009 the then prime minister of the United Kingdom, Gordon Brown, asked Tim Berners-Lee to collaborate on the creation of an instrument to increase the transparency of government activities, and facilitate communication with its citizens.[63] However, one may still question whether such public data is truly open. Although software can be developed to understand it, only citizens with a technical education will be able to use the programs needed to analyze the data and obtain suitable results. In spite of the transparency rhetoric, many open government initiatives thus still risk transforming public information archives into private data mining exercises (Birchall 2011a, 2011b, 2014).

2.6.3 Open access

The open data movement is also connected with open access. According to one of its main exponents, Peter Suber, "Open-access (OA) literature is digital, online, free of charge, and free of most copyright and licensing restrictions" (Suber 2007). The basic idea is that science and knowledge should be accessible to everyone without restriction and, therefore, it is necessary to break down barriers such as the cost of printing. The cost of scientific journals since the 1980s has increased by at least four times the rate of inflation, increasing the strain on university libraries, which has led to a reduction in the budget for monographs (Van Orsdel and Born 2009; Pochoda 2013).

The main argument in favor of open access for scientific articles is the fact that much of the research is publicly financed, with the result that the institution pays twice: once to finance the delivery of results and again to fund libraries to buy the journals in which the results are published.[64] Since the 2003 Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities,[65] many things have changed. Some publicly funded institutions in the US and abroad are now adopting a policy of requesting that studies published with their contributions must be available in open-access journals and repositories.[66] In times of economic crisis, it is natural that public research institutions will lose some of their financing. The reduction in funds available to libraries, along with the indiscriminate increase in the costs of journals, is making open access an almost obligatory choice,[67] though also fairer and more democratic. Even paywalled journals are gradually changing their strategies. Some are now maintaining their own repositories of preprints that authors have allowed to be open access. Others have started to convert themselves into open access journals. It is not at all the case that open access journals should side-step the peer-review process, or host any kind or quality of content. Nor is it the case that an open access journal cannot be an efficient business, since the costs of digital editing and publication are much lower than those for traditional print media.[68] But in spite of these advantages, a full conversion to the open access model still seems unlikely. The vast majority of high-prestige journals are still subscription-based, and for the authors, once they pass the peer review process, publication is essentially free. Laakso (2011) has charted the increasing popularity of open access publishing, but progress has been modest, and by 2009 only 7.7% of articles were published using the open access model. The reason for this slow take-up is that open-access journals simply shift the cost of publication from the libraries through subscription fees onto the researchers. *Digital Scholarship in the Humanities,* the flagship journal of the field, currently charges authors $3000 to publish an article with open access.[69] Fees for other journals vary greatly, from nothing to $3,900, the average being $660. The cost to the publisher is harder to determine and varies widely, though it tends be around $200–$300 per article and in some cases is much higher (Van Noorden 2013). But the open-access movement has at least made it possible for authors to directly compare costs, which should lead to greater competition that will drive down prices.

Public archives are another component of open access. These are either theme-based or belong to the university where the researcher works.[70] From an author's point of view, authoritative studies show that open access articles are cited more often than those behind a paywall.[71] This is important because in the world of royalty-free publication (scientific literature in which the authors are not paid for what they produce) what matters is only the impact and dissemination of research. Even in terms of the authoritativeness of content, an open-access channel running parallel to paywalled journals is possible, without surrendering control and evaluation of articles.

Gloria Origgi and Judith Simon in 2010 edited "Scientific Publications 2.0. The End of the Scientific Paper?" a special issue of Social Epistemology where authors try to analyze and challenge the current practices and systems of scientific publishing and scholarly communication. Their perspective is that, while the official strategic evaluation of scientific works is based on the need for peer-review to increase the value of scientific publications, "epistemic vigilance" tells a different story. Digital open knowledge practices weaken the distinction between "certified" academic entities and the public, reduce the time lapse between when an idea is created and when it is publicly accessible and citable, allow a more relaxed conversation among interested parties, and remove the boundaries between scientific and non-scientific subjects. As a result, "a system of norms for publishing that encourages today's researchers to submit to peer-reviewed journals while discouraging them to write for Wikipedia is an epistemically irresponsible system that should be challenged by researchers" (Origgi and Simon 2010, 146). However one should be aware that the open access and Web 2.0 scientific practices cannot alone resolve the problem of cultural and linguistic inequalities in the evaluation and transmission of knowledge. As pointed out recently by the Indian economist C.P. Chandrasekhar, the "distribution of recognized knowledge tends to be extremely uneven, with that unevenness being geographically stark," and open access does not yet challenge "either the for-profit framework or the problem of North Atlantic domination. The result could be the magnification rather than attenuation of currently prevailing distortions in the control over the nature of knowledge" (Chandrasekhar 2013). Scientific citation indexes are biased by many different variables (not least a linguistic bias, see Gazzola 2012), tend to increase popularity of already popular authors and institutions, and depend on the centrality of the paper positioning within mainstream knowledge.[72] So why rely on such evaluation methods while disregarding other social recognition of scholarly work? According to the scientist Michael Nielsen (2011), collective intelligence activity is definitely changing the methods, evaluations and results of science. In his book on networked science he describes how the practice of producing science in an open environment is transforming science itself, its objectives, and its accepted methods.

While this is true in general, and especially for the humanities, the value of humanistic research is intrinsically difficult to quantify or measure. The critical faculty and the diversity of approaches that form the basis of the humanities disciplines require independent judgment and a change in perspective from mainstream beliefs. DH should take into account such changes resulting from the open access model as applied to the study of culture and science, and support discussion of such delicate issues as measuring the validity of scientific outputs. As argued by Roopika Risam, "Rethinking peer review in the age of digital academe is a task that goes beyond the question of medium or platform to a question of epistemology" (Risam 2014). A new approach to digital scholarship is reflected by projects like DHCommons (http://dhcommons.org/) whereby DH scholars put together resources, data and projects in order to cooperate and benefit from a mutual collaboration in research and the sharing of results. However, as in the rest of the sciences, much still needs to be done to overcome the problem of the under-representation of DH research produced in less visible languages, and by more marginal institutions and countries.[73]

The funding of digital humanities projects and their assessment is rather complicated because it involves different competences and skills that need to cooperate together in order to achieve desired goals. For this reason, the presence of institutions eager to fund and to sponsor DH projects is fundamental to consolidation of the field. Open access projects can be funded by various institutions, particularly within the Anglophone countries. Looking at the projects funded in 2014 from the various grants of the National Endowment for the humanities[74] (http://www.neh.gov, the US agency which funds humanities projects) it is clear that many of the projects belong to the area of open access digital humanities research. Moreover, an increasing number of DH journals are now published as open access: Digital Humanities Quarterly, (http://www.digitalhumanities.org/dhq/, published since 2007), *Digital Studies/Le Champ Numérique* (http://www.digitalstudies.org, published since 1992, though irregular until 2009), and the Journal of the Text Encoding Initiative (http://journal.tei-c.org/journal/, since 2011).

The digital humanities community has another open space to discuss and assess DH projects: the already mentioned DHCommons project (http://dhcommons.org), sponsored by CenterNet — an international network of DH centers.[75] DHCommons is a hub for people and institutions to find collaborators and a space to discuss ongoing projects. The DHCommons also publishes an open access journal, the *DHCommons journal.* The scope of this journal is to represent and give voice to the multilingual, multidisciplinary activities of the DH community; "to certify the scholarly contributions made by digital projects-in-progress, helping scholars articulate the interventions of their digital work; ... to foster an innovative, truly developmental model of peer review." (http://dhcommons.org/journal).

The objectives of this open access DH journal underline the difficulties faced by the scholars in the field in obtaining a fair assessment, and academic acknowledgement of, their multidisciplinary work.

The permanent crisis of the humanities has pushed these disciplines towards new strategies for communicating with the general public, finding new sources of funding, and assessing the value of the projects produced within their departments. The humanities, in fact, as opposed to science, medicine or defense, do not enjoy the same access to private or national agencies for funding resources. This permanent underfunding condition forces every scholar to justify his or her engagement with research. There is also a problem in the lack of communication infrastructure and in the technical abilities of humanities researchers, as suggested by Alan Liu (2012, 496). The skills of digital humanists could help the humanities develop digital tools such as blog or content management systems that merge platforms for publication as Open Journal Systems with other devices, such as text analysis and extraction tools, as in the case of Simile Exhibit and Timeline (Liu 2012, 497).

It is important, however, that DH maintains a critical attitude and does not become ancillary or subsidiary to the humanities. This is the essence of Liu's heartfelt claim about the need for DH to become a leading advocate for the humanities (Liu 2012, 495–498).[76] But the risk is that DH scholars may fail in their goal to deliver tools that increase awareness in the use of digital technologies by only automating existing processes. As suggested by Liu "... the appropriate, unique contribution that the digital humanities can make to cultural criticism at the present time is to use the tools, paradigms, and concepts of digital technologies to help rethink the idea of instrumentality. … The goal is to rethink instrumentality so that it includes both humanistic and STEM (science, technology engineering and mathematics) fields in a culturally broad, and not just narrowly purposive, ideal of service" (2012, 501). The critical attitude that characterized the study of the humanities for centuries should also be directed inwardly towards their own methods and results.

Wendy Chun's approach is similar: "the blind embrace of DH … allows us to believe that the problem facing our students is our profession's lack of technical savvy rather than an economic system that undermines the future of our students. … The humanities are sinking … because they have capitulated to a bureaucratic technocratic logic" (Chun and Rhody 2014, 3–4). She underlines the desire to give humanities students a technical education, as if this could let them escape the crisis of the humanities, without any critical attitude towards which technologies are needed to face the crisis and to find satisfying job opportunities.

Open journals could thus be one of the key tools to escape the old, traditional procedures for assessing the merits of the scholars. However, tension still exists because "the position of being a scholar and that of being a blogger" (Fitzpatrick 2012, 452) are too different to permit a single, adequate publication policy for ideas. Open publishing journals do not offer a definitive solution to this dilemma. The peer review policy is not in line with "publish, then filter" proposed by Clay Shirky. According to Fitzpatrick, "the self-policing nature of peer review, coupled with its reliance on the opinions of a very small number of usually well-established scholars, runs the risk of producing an ingrained conservatism, a risk-averse attitude toward innovation, and a resistance to new or controversial approaches" (Fitzpatrick 2012, 454). This imposes a strong tension between the traditional peer-reviewed practices of academic journals (including the open publishing selection strategies) and the kind of authority definition advocated by Internet blogging and posting practices, which relates mainly on the effect of visibility and success through readership, not retrospective filtering. Hence digital scholars find themselves in a continuous struggle between two contrasting trends. On one hand, they need to get an academic position, like all other scholars, according to the traditional authority-oriented evaluation practices of departments; on the other hand, they are eager to participate in the on-going digital debate, whose engagement rules are based on a scarcity of attention, rather than limited printing space.

Although this kind of new digital authority has its own bias and positive reinforcement effects, these are mainly due to the readily adopted hegemonic practices typical of online publication. However, online communication uses measurements and parameters that are completely different from the traditional publishing methods of the academic press. So it is important to find a balance and to openly discuss the changes that digital publishing are bringing to assessment policies. A suggested by Gary Hall, one of the founders of Open Humanities Press (OHP), the Open Publishing movement, which until now has mainly worked just in the field of Science, Technology and Medicine, must engage with the practices of the humanities. "… the humanities could help prevent the OA movement from becoming even more moralistically and dogmatically obsessed with maximizing performance, solving technical problems and eliminating inefficiencies than it already is .... The humanities could help the OA community to grow, precisely by forcing scholars to confront issues of politics and social justice" (Hall 2010).

In turn, the humanities can offer added value to the open publishing movement, by limiting the rhetoric of efficiency, and at the same time discussing the meaning of "openness" in a more critical perspective than is normally adopted by scholars of the "hard sciences," who generally support this publishing practice.

The expertise of the humanities thus has a lot to offer the discussion about open press and the open access movement as a whole. It can play a leading role in an open and dynamic discussion of the crucial issues of this new way of assessing knowledge and enhancing science. It can say something relevant, for example, in the debate about the establishment of a new system for evaluating authority in open and traditional presses, and it can make a crucial contribution to the loss of author-recognition, and to the risks related to the lack of responsibility for assertions made in a scientific environment.

Summary of Part I

In the introduction of his famous paper History of Science and its Rational Reconstructions *Imre Lakatos remarks: "Philosophy of science without history of science is empty; history of science without philosophy of science is blind" (Lakatos 1971/1978, 102). Lakatos' apophthegm may serve as the underlying theme to the first part of this book, which investigated the historical background of the digital humanities from a different perspective. One of the objectives of this first chapter (*Technology and the humanities: a history of interaction*) was to show that, from its inception, computer science was built around a new form of symbol manipulation, which led in turn to a renewed representation and organization of knowledge. The various components of the electronic revolution: its foundations in logic, computability theory, cybernetics, artificial intelligence, human–machine interaction, and the theory of communication and information, all contributed to the discourse and rhetoric that informed and guided the birth of the digital humanities.*

The central issues of our new discipline were also, interestingly, major points of discussion at the origins of computer science. But the pioneers of that early era cannot be so clearly discerned as humanists or computer scientists as they would be today. As suggested by Wiener (1948, 2), it was a discipline that started to colonize the "no man's land" between the various established fields that contributed to its birth. Looking back at the key events that made the computing and communication technologies possible, it is clear that many of the important scientific problems their inventors faced were already being discussed openly within the humanities. These included the organization of information, the ethical responsibilities arising from the creation of machines for processing it, the development of databases for scientific or everyday use, communication between humans and machines, the definition of intelligence, and the representation of interactions between the human mind and what Licklider called the "fund of knowledge" (1965).

The answers to these crucial and substantially "pre-scientific questions"[1] *not only shaped the core of computer science, but also informed and directed the birth of humanities computing, and later digital humanities. Our aim was to describe the debates that not only animated the first years of computer science, but also cast their shadows on the present discussions about the digital humanities (Gold 2012).*

The possibility of describing the human activity of computing as a succession of elementary steps that could be written down (Turing 1937) lay the foundation for the invention of first the Turing Machine, and then the Universal Machine. When, almost ten years later, the computer was realized by the combined effort of a group of engineers, linguists, physicists and mathematicians, the Turing Machine came to embody a certain notion of computability as the manipulation of symbols in accordance with a set of definitive and precise written rules, which would allow the machine to be controlled.

The next step was unforeseen: that symbols manipulated according to precise rules could be used to represent any kind of content, whether linguistic or mathematical. The victory of the digital over the analog representation of information was based on the argument that the digital approach allowed for a more precise handling of content that would be more suited to obtaining a rigorous set of responses from the same inputs. The *concept of the black box, borrowed from cybernetics, was central to the definition of this new kind of machine: a device whose internal function need not be understood by its users so long as it guaranteed a certain behavior within a given limit of reliability. The end-users were thus not in complete control of their machine, but were still able to communicate precise orders and illicit the semblance of a precise and adequate response. The outcome of the struggle for technological success was thus a new paradigm for a machine based on the concepts of communication, control, feedback and information-processing.*

The "computer as a communication device" (Licklider and Taylor 1968) became the new paradigm for this next generation of machines. Its definition started with the work of Licklider at the beginning of the 1960s, including some of the analog criteria within the digital computing environment, which was the key to its sudden success. Licklider was the leading figure who was able to create a precise and convincing agenda for the emerging paradigm: the human–machine symbiosis. A human operator, who could perform his/her[2] *gestures within the constraints of the digital device, represented the analog component of the symbiosis, and could exploit all the capabilities of the machine within the new frame of communication. This scenario, based on the possibility of interaction, and on the rhetoric of "augmenting" human intelligence, differed from that of artificial intelligence, whose goal was the complex simulation of human capabilities by a machine (see Franchi, Güzeldere 2005, I).*

The perspective supplied by Norbert Wiener differs from that of the other pioneers, by including ethical, political and social issues raised by the new control and communication paradigm, which had been completely disregarded by Licklider and his colleagues at the Information Processing Technologies Office (IPTO), and by the Arpanet project. Wiener was unhappy that new solutions, based partly on his work, were being used in military or private sector projects. His behavior suggests an opposition to the major transformation of science that was taking place at the end of World War II, as suggested in one of his books, published posthumously, Invention: The Care and Feeding of Ideas *(1993).*

According to Wiener, the Human use of human beings *(1950/1954) would include a different attitude to the use of technology. His critical attitude can also be regarded as a recommendation for the digital humanities: that we should not lose sight of the true "soul" of the humanities disciplines that that were being transformed by the introduction of computing technology. Wiener's inheritance is that he considered technology only as a means to an end, and not as a goal in itself.*

Another interesting topic for the digital humanist is Wiener's insistence that machines should be under the control of human operators, even when they could operate autonomously. Wiener's perspective was based on an equal understanding of the fields of technology/science on the one hand, and of humanities on the other, and of the potential risks and benefits for interaction between the two. He did not want to oversee a reduction of the human element in decision making. The machine was incapable of making responsible choices, even though it was technically able to choose between alternative behaviors. In other words, one cannot ask the machine to decide on our behalf, even though it may be consulted when solving problems.

These are exactly the kinds of issues being discussed by digital humanists today. The historical debate thus highlights not only the importance of these events, but also their *relevance to current debates within the digital humanities. Beyond technical arguments about which methods should be adopted in the processes of digitization, the scope of the discipline should also include the ethical, political and social consequences of using digital techniques to reassess our critical vision of literature, the visual representation of poetry, or the analysis of Big Data and social media (Kitchin 2014). What cannot be denied is the crucial role played by representing, organizing, archiving, retrieving and classifying content for the creation of new knowledge, or in defining priorities for research. If digital humanists truly want to be key players in the transformation of the new "texts" (in the broader sense of the term), they cannot avoid questions like these, which are their most natural research objectives.*

*The second chapter (*Internet, or the humanistic machine*) offered a critical account of the history of the Internet from a humanistic perspective. It used examples of trends and tools that exemplify the social, economic and cultural biases inherent in current technological choices, and how they shape the use, organization and practical rules governing whether the Internet will become a common pluralistic space, or a controlled and elitist environment. In order to extend and strengthen its positive social and economic effects, the issue of a genuinely democratic governance of the Net, and the creation of tools and opportunities to support and strengthen cultural differences must be addressed. Digital humanists can play a strategic role in meeting this challenge because they have historically been trained to recognize, understand and empower cultural diversities. Digital representations — both visual interfaces and textual — embody and reflect the spatial and temporal layers of cultural difference:*

> *... cultural diversity, which includes identities, experiences of, and encounters with difference, is always* mediated, *that is, constructed, (re)presented, and experienced through the media of communication. (Siapera 2010, 5)*

Digital humanists are responsible for managing these representations. DH institutions and international organizations should renew their agenda on issues of multicultural diversity. But simply opening conferences to foreign projects and authors is not enough. They must also demonstrate that research in this field can be independent of the interests of industry, from the governmental obsession with control, and that the DH community is able to propose new critical solutions, and not merely reproduce or utilize state-of-the-art software and databases. It is clear what ducking this responsibility might mean. At a time when Western cultures are setting about translating (and transforming) their own knowledge into digital formats, the need today, as five centuries ago, is clearly that of elaborating a new paideia: the creation of a multicultural and multilinguistic community able to train the trainers.

Among other issues, this chapter also traced a map of the opportunities and risks provided by open knowledge from a theoretical, critical and social standpoint. It discusses the challenges in overcoming and discussing various kinds of resistance, from the strictly organizational to the political. It focused on the relationship between open access and open publishing within the DH community, and described the positives and negatives *in the use of open tools. It also highlighted the possible epistemological, critical and economical outcomes of an overly enthusiastic endorsement of open Big Data.*

The last part of this chapter attacked the familiar characterization of the origins of the Internet as a lost paradise of freedom and opportunity, destroyed by the interests of commercial, marketing, military and intelligence organizations. It offers instead a multifaceted reconstruction of its history, in which the opportunities and risks of the technologies involved were present from the very beginning.

One important goal of this book is to argue that technological innovation is never neutral. A "way back machine" cannot be applied to return society to a former state once innovation and transformation have taken root. But history shows that the consequences of technology are not always inevitable. Everyone involved in the process of transformation should be aware of the potential risks, and can accept or reject the new tools. Managing the effects that new technologies have involves political and social choices as well as an accurate assessment of the potential benefits. As digital humanists, we have a responsibility to discuss how we use the new tools, and not to blindly accept the latest technical fad without any critical assessment of its "meaning," or of the epistemological and cultural biases it introduces.

Notes

Preface

1. Fish, Stanley. "The Old Order Changeth." *Opinionator, New York Times* blog, December 26, 2011. <http://opinionator.blogs.nytimes.com/2011/12/26/the-old-order-changeth/>
2. See also Pannapacker, William. "The MLA and the Digital Humanities." *Brainstorm, The Chronicle Review* blog, Dec. 28, 2009.
3. See <http://tcp.hypotheses.org/411>
4. See Svensson, Patrick, "Humanities Computing as Digital Humanities," *DHQ,* 4:3, 2009, <http://www.digitalhumanities.org/dhq/vol/3/3/000065/000065.html> and "The Landscape of Digital Humanities," *DHQ,* 4:1, <http://www.digitalhumanities.org/dhq/vol/4/1/000080/000080.html>
5. Wang, Xiaoguang and Mitsuyuki Inaba, "Analyzing Structures and Evolution of Digital Humanities Based on Correspondence Analysis and Co-word Analysis," *Art Research,* 9, 2009, pp. 123–134.
6. Other prominent figures in the digital humanities in Italy include Tito Orlandi in Rome and Dino Buzzetti in Bologna.
7. This Preface is based on a review I wrote of the previous Italian edition in 2010.
8. See for example Willard McCarty's *Humanities Computing* (New York: Palgrave, 2005) which goes much deeper than *The Digital Humanist* into issues of what we do, but ultimately presents the field as an interdisciplinary commons for methods without any political agenda of its own. For that matter, my forthcoming book *Hermeneutica* (MIT Press) avoids political/cultural issues.

Introduction

1. See http://www.globaloutlookdh.org/working-groups/491–2/.
2. Also because resources are not stored according to the geographical provenance of the project's server (and many projects on Asiatic culture, as is well known, are based in North American and European institutions).
3. Among the collective aims to define Digital Humanities should be mentioned two different manifestos that appeared in the past few years. In the Anglo-American context see http://www.humanitiesblast.com/manifesto/Manifesto_V2.pdf and for a European and international vision see the Paris 2011 Manifesto: http://tcp.hypotheses.org/411.
4. Social and ethical issues are raised by the document collectively assembled at the Bern 2013 DH Summer School: "This document is aimed as a contribution to the current debates in the digital humanities about how digital humanists conduct themselves as professionals ethically, and as a reflection of their core values" (https://docs.google.com/document/d/1A4MJ05qS0WhNlLdlozFV3q3Sjc2kum5GQ4lhF0NKcYU/edit?pli=1#heading=h.fbfb3vwicb5).
5. The organization around "clusters of problems" has inspired the authors' interdisciplin-

ary project http://www.newhumanities.org/, built around seven projects, each of which is carried out by a mixed team of humanists and scientists.

6. An impressive epistemological and historical account of the concepts and ideas behind the digital computer is offered by Luigi Borzacchini in his monumental three-volume work (Borzacchini 2008, 2010, and 2015).
7. Although the book is the result of almost five years shared research and discussion that started with the revision of our previous book (*L'umanista digitale,* Il Mulino, 2010), chapters 1, 2 and 5 are the work of Teresa Numerico, chapter 3, section 2.5 and the Conclusions are the work of Domenico Fiormonte, and chapter 4 is the work of Francesca Tomasi. Domenico Fiormonte edited all chapters with the help of Desmond Schimdt who revised the previous translation of Chris Ferguson, but also suggested important corrections and additions in all sections. Finally, we are grateful to Giorgio Guzzetta for his help and material assistance throughout the completion of this project.

Chapter 1

1. The ability to deal with formulas and numbers in the same way had already been realized with the method invented by Gödel (1931), known as arithmetization, which was used to demonstrate his incompleteness theorem.
2. See V. Bühlmann, "The idea of a Characteristica Universalis between Leibniz and Russell, and its relevancy today" http://monasandnomos.org/2012/12/05/the-idea-of-a-characteristica-universalis-between-leibniz-and-russell-and-its-relevancy-today/.
3. Andrew Pickering focuses on the influence of practice and its role in the reassessment of theoretical models in science. He uses the expression "mangle of practice" to define the social dimension of research – yet another humanistic aspect that is often neglected: "This temporal structuring of practice as a dialectic of resistance and accommodation is, in the first instance, what I have come to call the mangle of practice" (Pickering 1995, xi).
4. In *How Much Information?* a study published in 2003, it was estimated that the information stored on paper, film and magnetic and optical media has doubled since 1999 (http://www2.sims.berkeley.edu/research/projects/how-much-info-2003). An update to these estimates is provided by the more systematic work of Hilbert and López 2011, claiming to be "the first study to quantify humankind's ability to handle information and how it has changed in the last two decades".
5. See Bardini (2000) for more details on the Engelbart's contributions to friendly interfaces. Most of Engelebart's original papers and reports are available on his website: http://www.dougengelbart.org/pubs/augment-3906.html.
6. For a more detailed account on the relationships between computer science and linguistics see Fiormonte and Numerico 2011, and Hajič 2004.
7. For a discussion on the relationship between computer science and humanities computing see McCarty 2005, 177–198.

Chapter 2

1. On User-Centric Design see e.g. Ch. 9 of Norman 1998.
2. The concept of model is now central also in the field of DH, cf. McCarty 2005, Orlandi 2010, Buzzetti 2002, 2009, Fiormonte 2009.
3. The birth of the ARPANET is one of the most studied events in the history of technology. Good starting points are, for example, Hafner and Lyon 1996 and Naughton 1999. For a more sociological point of view on the Internet's origins see Castells 2001.
4. For more details on the role of Engelbart see Bardini 2000, and for the role of Xerox PARC in the development of Apple (329ff) and

Microsoft (358–360) see Hiltzik 1999.

5. The original proposal can be downloaded from the "history" section of the W3Consortium: http://www.w3.org/History/1989/proposal.html.
6. The interview, divided into two parts, can be found on YouTube: http://www.youtube.com/watch?v=TkOpzbTsDJE&feature=PlayList&p=4B2E3AC7440A2CD1&index=63
7. See *Answers for Young People,* accessible through Berners-Lee's pages on the W3C, http://www.w3.org/People/Berners-Lee/Kids.html.
8. The first traceable reference to "hyper-text" by Nelson is in the Vassar Miscellany News of February 3, 1965: http://faculty.vassar.edu/mijoyce/MiscNews_Feb65.html.
9. Tim Berners-Lee originally wanted to speak explicitly of documents (UDI = Universal Document Identifier) but then chose to accept the compromise both to replace the document with the more general resource and replace the philosophy of universality with that of uniformity (Berners-Lee 1999, 67).
10. The document can be read here http://tenyears-www.web.cern.ch/tenyears-www/Declaration/Page1.html.
11. http://www.w3.org/Consortium/.
12. For examples of the W3C's role in the affirmation of the standard markup languages cf. below Ch. 4.
13. Web Science Conference 2009, held in Athens in March 2009. The video of Berners-Lee's speech is available on: http://www.youtube.com/watch?v=ol_Y_MPDc4E&feature=related.
14. See for details the Webfoundation page dedicated to web challenges: http://www.webfoundation.org/programs/challenges/.
15. See, for example, Google One Pass, which offers a perfect system for allowing publishers to offer paid-for content via a search engine platform.
16. For more details on this method of storing content cf. below Ch. 4.
17. See http://en.wikipedia.org/wiki/Wikipedia:About.
18. Similar conclusions can be found in the 2014 UNESCO's Fostering Freedom Online report (MacKinnon et al. 2014). The report discusses policies and practices of companies representing three intermediary types (internet service providers, search engines, and social networking platforms) across ten countries.
19. Except for the first paragraph, written by T. Numerico, this section was authored by Domenico Fiormonte.
20. Or its elitist tendency: "In the end, Web 2.0 works best for the Internet everyman or everywoman, who tends to be educated and well-off. Thus, in addition to restricting information to those who have access to the Internet, Web 2.0 restricts relevant information to those who are most similar to that typical Internet user." (Witte and Mannon 2010, 19).
21. For reasons of space problems related to cultural and technological biases of the TEI and in general of markup systems will be not tackled here. For a detailed discussion see Schmidt 2010 & 2014; Fiormonte and Schmidt 2011; Fiormonte et al. 2010.
22. According to the OECD Factbook 2013 (http://dx.doi.org/10.1787/factbook-2013–67-en), in 2011 about 71% of US adults had Internet access. According to Pew Research Center in 2015 the share of all US adults who use the internet has reached 84% (http://www.pewinternet.org/2015/06/26/americans-internet-access-2000-2015/). The US Census Bureau 2013 survey on Computer and Internet use showed that age, gender, ethnicity, income and educational level can have a significant impact on Internet access and household computer ownership (http://www.census.gov/content/dam/Census/library/publications/2014/acs/acs-28.pdf).
23. See for example the reflection on the digital divide by a Mexican researcher: "Una tesis ocupa un lugar central de este tipo de investigación de la brecha digital en los EE.UU. y postula que el 48% de la población que no utiliza el Internet se abstienen de ello porque no lo encuentran relevante para su vida diaria o no les interesa (Pew Research Center, 2013, NTIA/FCC, 2013; estos estudios encuentran eco en estudios de caso mexicanos: WIP, 2013, INEGI, 2013). No obstante, dicha tesis deja sin desarrollar la pregunta de ¿Cómo es que este segmento de la población se ha inclinado por responder de esa manera? es decir ¿Cuáles son las prácticas que median la relación entre personas y tecnologías de la información?" (Sánchez 2014: 89).
24. http://www.itu.int.
25. "High-income economies are those in which 2012 GNI per capita was $12,616 or more" (http://databank.worldbank.org/data/).
26. Source: World Bank Indicators: http://databank.worldbank.org/data. See also the OECD

2001 survey on the digital divide: http://www.oecd.org/internet/ieconomy/understandingthedigitaldivide.htm.

27. See http://epp.eurostat.ec.europa.eu/portal/page/portal/information_society/data/database.
28. http://www.icann.org/en/news/announcements/announcement-30oct09-en.htm.
29. This wasn't always the case: e.g. www.monash.edu, an Australian academic institution. Those who obtained .edu names in the past are allowed to keep them, but no new names for non-US academic institutions will be allowed. See http:// net.educause.edu/edudomain/eligibility.asp.
30. Until 2009 ICANN was essentially controlled by the US Department of Commerce: http://www.readwriteweb.com/archives/commerce_department_loosens_grip_on_icann.php. Until July 2012 the CEO and President of ICANN was Rod Beckstrom, former Director of the National Cybersecurity Center (NCSC) at the US Department of Homeland Security, who has been replaced by US-based entrepreneur and former IBM manager Fadi Chehadé (http://www.icann.org/en/groups/board/chehade-en.htm). With the arrival of Chehadé ICANN increased its efforts for evolving in a "bottom-up, consensus-driven, multi-stakeholder model" (see http://www.icann.org/en/about/welcome). In December 2013 the Strategy Panel on ICANN Multistakeholder Innovation was launched and began a collaboration with thegovelab.org. The objective of this new course will be, among other things, to "open ICANN to more global participation in its governance functions" and "proposing new models for international engagement, consensus-driven policymaking and institutional structures" (http://thegovlab.org/the-brainstorm-begins-initial-ideas-for-evolving-icann/).
31. UN agencies promoted and sponsored a number of events as well as produced many reports on Internet Governance; for a synthesis see, for example, the NetMundial Multistakeholder initiative in 2014: http://netmundial.br.
32. See http://www.internetworldstats.com/stats.htm.
33. Web Foundation, Challenge page http://www.webfoundation.org/programs/challenges/.
34. According to World Stats (a private company), at the end of 2011 Africa represents 15% of the world population, but only 6.2% of Internet world users with a penetration of 13.5% of the population. However, Egypt and Tunisia are among the 10 African countries in which the Internet is more developed.
35. The dark sides of Facebook and other social media are explored in a collection edited by Geert Lovink and Miriam Rash (2013). See especially articles by Langlois, Gehl, Bunz and Ippolita and Mancinelli. For a recent overview on the role of social media in Arab countries see also Jamali 2014. A critical perspective on the "the celebratory hype about online activism" during the Arab spring is provided by Aouragh 2012.
36. Cases of cyber-repression are constantly being reported in most Asian and Arabic countries where different forms of online political activism exist. A source of information in English is the database http://cyberdissidents.org.
37. For a detailed description of the most spoken languages on the Internet see http://www.internetworldstats.com/stats7.htm, where it is clear that English is the most spoken language, but Chinese and Spanish, respectively second and third languages of the Web, are growing fast.
38. Some of the themes of the 2001 book were taken up and upgraded in a recent post: http://jamillan.com/librosybitios/espimprered.htm.
39. "By far its most striking feature was its graphical user interface, ... The arrangement of folders and icons built around what the Star engineers called the 'desktop metaphor' is so familiar today that it seems to have been part of computing forever." (Hiltzik 1999, 364).
40. A complete list of these animations is available from http://wiki.secondlife.com/wiki/Internal_Animations#User-playable_animations.
41. http://www.unicode.org/consortium/consort.html.
42. http://www.unicode.org/consortium/directors.html.
43. "Even if Unicode does not exactly 're-map' real life politics onto the virtual realm, such technical solutions do point to the ideological, political, and economic forces that promote and serve to benefit from attempts at universal language." (Pressman 2014, 2).
44. "For example, the game cricket in Hindi is क्रिकेट *krikeṭ*; the diacritic for /i/ appears before the consonant cluster /kr/, not before the /r/." (http://en.wikipedia.org/wiki/

Abugida).

45. The difficulties of encoding some Indic scripts according to the Unicode model are highlighted also by the SIL's Non-Roman Script Initiative: "Indic scripts have combining vowel marks that can be written above, below, to the left or to the right of the syllable-initial consonant. In many Indic scripts, certain vowel sounds are written using a combination of these marks …." (http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&id=IWS-Chapter04b#fb2c362c). In conclusion, "there are over 16,000 characters defined in Unicode that in one way or another go against the basic design principles of the Standard." (http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&id=IWS-Chapter04b#bb06c97e).
46. "Because Devanagari and other Indic scripts have some dependent vowels that must be depicted to the left of their consonant letter (*although they are pronounced after the consonant*), the software that renders the Indic scripts must be able to reorder elements in mapping from the logical (character) store to the presentational (glyph) rendering." (Aliprand 2003, 228; quoted in Perri 2009, 736. Italics added by Perri).
47. The invasion of Europeans in the 19th and 20th centuries changed the living standards of the linguistically diverse aboriginal population. However, according to the Australian Bureau of Statistics "Today, there are approximately 22 million Australians, speaking almost 400 languages, including Indigenous languages" (Source: http://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/1301.0Feature%20Article32009%E2%80%9310?opendocument&tabname=Summary&prodno=1301.0&issue=2009%9610&num=&view=).
48. "The dominating effect of a single socioeconomic factor, GDP per capita, on speaker growth rate suggests that economic growth and globalization … are primary drivers of recent language speaker declines (mainly since the 1970s onwards), for instance, via associated political and educational developments and globalized socioeconomic dynamics. This conclusion is also supported by the positive effect of GDP per capita on range size and many language extinctions in economically developed regions, such as the USA and Australia." (Amano et al. 2014, 7).
49. See http://melissaterras.blogspot.in/2012/01/infographic-quanitifying-digital.html.
50. "In the book, I argue that the kind of thinking expresses a kind of Digital Universalism that disguises the means which elite designers and entrepreneurs of the IT world's leading corporations work to promote and circulate it – whether in the pages of Wired magazine or across any number of TED conference stages. It also disguises the diverse imaginaries and investments around the digital that are cropping up all over the world, including in Peru, from diverse civil society actors – but that are easy to overlook when we focus our attention only on those coming out of just a handful of innovation centers." (Chan and Jenkins 2015).
51. http://en.wikipedia.org/wiki/Open_Data.
52. From the most famous, Wikipedia http://wikipedia.org to wikibooks http://wikibooks.org a bank of texts editable by the community, and geonames http://www.geonames.org/ to Wordnet http://wordnet.princeton.edu/, an English lexicon that includes a complex system for grouping terms according to semantic affinities, hosted by MIT. There is a map of all data accessible and connected according to the norms of Linked Data.
53. The W3C's wiki dedicated to the Semantic Web shows the resources for creating open data archives: http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData.
54. There are some unavoidable overlappings between this section and Ch. 5. Other aspects of Big Data are analyzed in § 5.8.
55. More information and reflections on big data and DH can be found on Geoffrey Rockwell's blog theoreti.ca. In December 2014 Rockwell gave a lecture on *Big Data in the Humanities*: http://vimeo.com/114389377.
56. "[W]here are the digital humanists critiquing the growing surveillance state? … The synthesis of our training in traditional humanistic modes of inquiry with our proficiency in network analysis, text-mining, metadata, and the other methods the US government uses to monitor its own people in a (I would argue, misguided) search for threats should lead to a proliferation of analyses, arguments, and action. To date, however, I have heard only quiet murmurings in the community." (Widner 2013).
57. Cf. Geoffrey Rockwell's note on Kristene Unsworth's talk: "The justice of big data and global citizenship" (http://philosophi.ca/pmwiki.php/Main/InternationalEthicsRoundtable2014InformationEthicsAndGlo-

balCitizenship).

58. For a description of the project and the first results of this research see Michel et. al. 2011.
59. There is not space here for a critical discussion of Culturomics. We would like just to underline that this initiative comes from a group of scientists and the Google Book Search team who seem to have a marginal interest in the humanities. It is also interesting to note that they claim to define what culture is, using a corpus of 4% of the total number of books.
60. http://opendatahandbook.org/en/what-is-open-data/index.html.
61. For a description of the tools set see the www.tapor.ca.
62. For more information see http://voyeurtools.org and Rockwell's blog http://hermeneuti.ca.
63. Tim Berners-Lee (2009) outlined the guidelines of its contribution to the project to reorganize communication with digital users of English institutions (http://www.w3.org/DesignIssues/GovData.html). See also Berners-Lee's talk at TED http://www.ted.com/index.php/talks/tim_berners_lee_on_the_next_web.html.
64. For these reasons, a Cambridge (UK) mathematician, Tim Gower, in January 2012 has launched a boycott of the international publisher Elsevier (http://thecostofknowledge.com/). Academics are protesting against Elsevier's business practice, such as exorbitantly high prices, and its support for measures such as SOPA, PIPA, and the Research Works Act "that aim to restrict the free exchange of information".
65. http://openaccess.mpg.de/Berlin-Declaration.
66. From April 2008, the National Institute of Health (NIH) has made research it has helped finance available through PubMed Central http://www.pubmedcentral.nih.gov, its own open access archive of scientific literature on medicine and life.
67. A recent overview of the rapidly growing scenario of OA publishing is given in Laakso et al. 2011.
68. For more details on the business model of open source journals see Suber 2007 and its sources, in particular Crow 2009. For a list of open access journals see the *Directory of Open Access Journals* (http://www.doaj.org).
69. http://dsh.oxfordjournals.org/for_authors/index.html.
70. For a list of institutional repositories by location, see *Directory of Open Access Repositories – Open DOAR* (http://www.opendoar.org/); for other information on repositories and their policies: Open access directory (http://oad.simmons.edu/oadwiki/Main_Page). For a list of repositories by subject, there is a section on disciplinary repositories (http://oad.simmons.edu/oadwiki/Disciplinary_repositories). The majority of Open Access Archives are compliant with the protocols of the Open Archive Initiative (OAI http://www.openarchives.org/) in their use of metadata.
71. For an up-to-date bibliography on this topic, see the *Open Citation Project* for evaluating analysis of citations of articles in open access archives. http://opcit.eprints.org/oacitation-biblio.html.
72. Mainstream knowledge that is often linked to well-known geopolitical issues: "US accounted for 26% of the world's total Science & Engineering (S&E) articles published in 2009 and the European Union for 32%. In 2010, the US share in total citations of S&E articles stood at 36% and the EU's share at 33%, whereas that of Japan and China remained at 6% each." (Chandrasekhar 2013).
73. Not entirely convincing is also the policy of "discounted rates available for authors based in some developing countries". See http://www.oxfordjournals.org/our_journals/litlin/for_authors/index.html.
74. The complete list of funded projects can be found at http://www.neh.gov/files/press-release/july2014grantsstatebystate.pdf.
75. http://centernet.adho.org/about. See Conclusions.
76. Alan Liu is one the founders of the 4humanities.org project. The aim of this platform is to assist in advocacy for the humanities, creating a starting point to help and support humanities disciplines to display their potential within the digital environment.

Summary of Part I

1. See Keller 2010 for a discussion of the prescientific questions that inform scientific

discourse.
2. Women, however, were not included in Licklider's description of the symbiosis, which was explicitly called a "man-machine symbiosis".

Chapter 3

1. Petrucci defines *dominus* as the owner of the cultural and economic processes that shape and transform textual products (Petrucci 1986, xxi).
2. Some of the concepts included in the "Text Wheel" elaborated by Patrick Sahle in the realm of the scholarly digital edition partially overlap (and can integrate) with our model. Sahle identifies six dimensions of the text: Text as Idea, Intention; Text as (structured) Work; Text as Linguistic Code; Text as Version; Text as Document; Text as (visual) Sign (cf. Sahle 2013, 1–98).
3. "[W]e have conceived ourselves and the natural entities in terms of data and information. We have flattened both the social and the natural into a single world so that there are no human actors and natural entities but only agents (speaking computationally) or actants (speaking semiotically) that share precisely the same features. It makes no sense in the dataverse to speak of the raw and the natural or the cooked and the social: to get into it you already need to be defined as a particular kind of monad." (Bowker 2013, 169).
4. The scenario described hereby can also be interpreted in the apocalyptic way of Gilles Deleuze: "there is no need of science fiction for conceiving a control mechanism able to provide every single moment the position of an element in an open environment, an animal in a reserve, a man in a company (electronic collar). Félix Guattari has imagined a city where one would be able to leave one's apartment, one's street, one's neighborhood, thanks to one's (own) electronic card that raises a given barrier. But that card could just as easily be rejected on a given day or between certain hours; what counts is not the barrier itself but the computer that tracks each person's position–*licit* or *illicit,* and effects a universal modulation." (Deleuze 1990, authors' transl.)
5. Reference is not being made here to the "biological machines", such as the cyber beetle created by Michel Maharbiz (http://maharbizgroup.wordpress.com). When the beetle is at the pupal stage electrodes are implanted into the nervous and muscular system of the insect. A receiver and a battery is then integrated into the system on the back of the adult, so that its movements can be controlled remotely. Undoubtedly, such hybrids pose important ethical, social, aesthetic, semiotic and even political questions (the research is co-financed by the Pentagon).
6. Cf. the international online anthology available at http://www.hermeneia.net/eng/espais/literatura.html.
7. Baron (2000) remains the best general introduction to the digital dimension of language.
8. In addition to video, speech and writing, Skype can also integrate other tools, for example TalkAndWrite (http://www.talkandwrite.com), an application that allows two or more users to write and draw together on the same board. Skype is also among the first pieces of software that lets you correct the text after it is typed and sent, with the interesting effect that the temporal synchronicity of the chat can be reversed.
9. Peter Boot suggested calling certain forms of digital annotations *mesotext* (from ancient Greek *mesos,* the middle): "it is text that can be located somewhere in between the primary texts of scholarship (the sources that scholarship is based on), and its secondary texts" (Boot 2009: 203). Therefore, in Genette's terms "mesotext is a metatext" (Boot 2009, 207). However this is too fortuitous an expression to be confined to the realm of annotations, so I think it would be a good candidate for describing an intermediate level or specific kind of digital metatext.
10. See also Bolter and Grusin's concept of *Remediation* (1999); George Landow (1992) said similar things at the beginning of Web 1.0.
11. "Post-modern ethnography privileges 'discourse' over 'text', it foregrounds dialogue as opposed to monologue, and emphasizes the cooperative and the collaborative nature of the ethnographic situation in contrast to the ideology of the transcendental observer." (Tyler 1986, 126).

Birchall, C. (2011b), “Introduction to ‘Secrecy and transparency’: The politics of opacity and openness”, *Theory, Culture & Society,* 28, 7–8, pp. 7–25.

Birchall, C. (2014), “Radical transparency?”, *Cultural Studies <=> Critical Methodologies* 14, 1, pp. 77–88.

Bizer, C., Heath, T., Berners-Lee, T. (2009), “Linked data: The story so far”, *International Journal on Semantic Web and Information Systems (IJSWIS),* 5, 3, pp. 1–22. http://tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf.

Blair, D.C. (1984), “The data-document distinction in information retrieval”, *Communication of the ACM,* 1994, 27, 4, pp. 369–374.

Bohem, G. (2007), *Wie Bilder Sinn erzeugen: Die Macht des Zeigens,* Berlin, Berlin University Press.

Bolter, J.D. (2001), *Writing space: Computers, hypertexts, and the remediation of print,* Mahwah (NJ), Lawrence Erlbaum.

Bolter, J.D., Grusin, R. (1999), *Remediation: Understanding new media,* Cambridge (MA), MIT Press.

Boot, P. (2009), *Mesotext,* Amsterdam, Amsterdam University Press.

Borràs Castanyer, L. (ed.) (2005), *Textualidades electrónicas: nuevos escenarios para la literatura,* Barcelona, Editorial UOC.

Borzacchini, L. (2008), *Il computer di Platone: alle origini del pensiero logico e matematico,* Bari, Dedalo.

Borzacchini, L. (2010), *Il computer di Ockham: genesi e struttura della rivoluzione scientifica,* Bari, Dedalo.

Borzacchini, L. (2015), *Il computer di Kant: struttura della matematica e della logica moderne,* Bari, Dedalo.

Bowker, G. C. (2005), *Memory practices in the Sciences,* Cambridge (MA), MIT Press.

Bowker, G.C. (2013), “Data Flakes: An afterword to ‘Raw Data’ is an oxymoron”, in Gitelman (2013), pp. 167–171.

Bowker, G.C., Leigh Star, S. (2000), *Sorting things out: Classification and its consequences,* Cambridge (MA), MIT Press.

Bowker, G. C., Baker, K., Millerand, F., and Ribes, D. (2010), “Toward information infrastructure studies: Ways of knowing in a networked environment”, in Hunsinger, J., Klastrup, L., Allen, M. (eds.), *International Handbook of Internet Research,* Dordrecht–New York, Springer, pp. 97–117.

Boyd, B. (2009), *On the origin of stories: Evolution, cognition and fiction,* Cambridge (MA), Belknap Press of Harvard University Press.

Boyd, D. (2010), “Privacy and publicity in the context of Big Data.” Raleigh (NC), April 29. http://www.danah.org/papers/talks/2010/WWW2010.html.

Boyd, D., Crawford K. (2012), “Critical questions for Big Data”, *Information, Communication & Society,* 15, 5, pp. 662–679.

Brandwood, L. (1992), “Stylometry and chronology”, Kraut, R. (ed.), *The Cambridge companion to Plato,* New York, Cambridge University Press.

Brin, S., Page, L. (1998), “The anatomy of a large-scale hypertextual web search engine”, *Computer Networks and ISDN Systems,* 30, pp. 107–117. http://citeseer.ist.psu.edu/brin98anatomy.html.

Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., Wiener, J. (2000), "Graph structure in the Web", in *Proceedings of the 9th International World Wide Web Conference* (Amsterdam, May 15–19), New York, ACM Press. http://www.cis.upenn.edu/~mkearns/teaching/NetworkedLife/broder.pdf.

Bruner, J.S. (1990), *Acts of meaning,* Cambridge (MA), Harvard University Press.

Bruner, J.S. (1998), "Celebrare la divergenza: Piaget e Vygotskij", in Liverta Sempio, O. (ed.), *Vygotskij, Piaget, Bruner: concezioni dello sviluppo*, Milan, Raffaello Cortina, pp. 21–36.

Bühlmann, V. (2012), "The idea of a *Characteristica Universalis* between Leibniz and Russell, and its relevancy today", *Monas, Oikos, Nomos.* http://monasandnomos.org/2012/12/05/the-idea-of-a-characteristica-universalis-between-leibniz-and-russell-and-its-relevancy-today/.

Bunz, M (2013), "As you like it: Critique in the era of an affirmative discourse", in Lovink, G., Rasch, M. (eds.) (2013), pp. 137–145.

Burda, H., Kittler, F., Sloterdijk, P. (2011), *The digital Wunderkammer: 10 chapters on the Iconic Turn,* Munich, Wilhelm Fink Verlag.

Bush, V. (1931), "The differential analyzer: A new machine for solving differential equations," *Journal of the Franklin Institute,* 212, July–December, pp. 447–488.

Bush, V. (1945), "As we may think", *Atlantic Monthly,* 176, 1, pp. 101–108, in Nyce, J., Kahn, P. (1991). http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/3881/.

Bush, V. (1959/1991), "Memex II", *Bush Papers,* MIT Archive, in Nyce, J., Kahn, P. (1991), pp. 165–184.

Buzzetti, D. (2002), "Digital representation and the text model", *New Literary History,* 33, 1, pp. 61–88.

Buzzetti, D. (2006), "Biblioteche digitali e oggetti digitali complessi: esaustività e funzionalità nella conservazione", in *Archivi informatici per il patrimonio culturale.* Proceedings of the International Conference organized by the Accademia Nazionale dei Lincei in collaboration with Erpanet and Fondazione Franceschini, Rome, November 17–19, 2003, Rome, Bardi, pp. 41–75.

Buzzetti, D. (2009), "Digital edition and text processing", in Deegan, M., Sutherland, K., (eds.) *Text editing, print and the digital world,* Ashgate Aldershot, pp. 45–61.

Canagarajah, A. S. (2002), *A geopolitics of academic writing,* Pittsburgh, University of Pittsburgh Press.

Carey, J.W. (1992), *Communication as culture: Essays on media and society,* London–New York, Routledge.

Castells, M. (2001), *Internet galaxy,* Oxford , Oxford University Press.

Castells, M. (2009), *Communication power,* Oxford , Oxford University Press.

Chan, A. S. (2014), *Networking peripheries: Technological futures and the myth of digital universalism,* Cambridge (MA), MIT Press.

Chan, A. S., Jenkins, H. (2015), "Peru's digital futures: An interview with Anita Say Chan (part one)", February 17. http://henryjenkins.org/2015/02/perus-digital-futures-an-interview-with-anita-say-chan-part-one.html.

Chandrasekhar, C.P. (2013), "The role of open access in challenging North Atlantic domination in the social sciences". Paper presented at panel *Knowledge as a Commons: Open Access and Digital Scholarship in the Social Sciences,* sponsored by CLACSO, CODESRIA, IDEAs and INASP, World Social Science Forum, Montreal, 14 October 2013.

Chandrasekhar, C. P. (2014), "Open access vs academic power", *Real-world Economics Review,* 66, January 13, pp. 127–130. http://www.paecon.net/PAEReview/issue66/Chandrasekhar66.pdf.

Changeux, J.-P. (2003), *Gènes et culture: Enveloppe génétique et variabilité culturelle,* Paris, Odile Jacob.

Chartier, R. (1991), "Textes, formes, interprétations", preface to Mckenzie, D. F., *La bibliographie et la sociologie des textes,* Paris, Éditions du Cercle de la Librairie.

Cho, J., Roy, S. (2004), "Impact of search engines on page popularity", in *Proceedings of the WWW 2004,* New York, ACM, pp. 20–29.

Chomsky, N. (1957), *Syntactic Structures,* The Hague, Mouton.

Chun, H.K.W. (2006), *Control and freedom. Power and paranoia in the age of fiber optics,* Cambridge (MA), MIT press.

Chun, H.K.W. (2011), *Programmed visions: Software and memory,* Cambridge (MA), MIT Press.

Chun, W.H.K., Rhody, L. M. (2014), "Working the digital humanities: Uncovering shadows between the dark and the light", *Ada: A Journal of Feminist Cultural Studies,* 25, n.1, pp.1–26.

Clavaud, F. (2012), "Digital Humanities as university degree: A brief overview of the situation in France". http://siacre.enc.sorbonne.fr/~fclavaud/presentation/DH2012-DHDegrees-France-FClavaud.pdf

Clavert, F. (2013), "The digital humanities multicultural revolution did not happen yet." http://histnum.hypotheses.org/1546.

Clifford, J. (1986), "Introduction: Partial truths", in Clifford, J., Marcus, G.E. (eds.), *Writing culture: The poetics and politics of ethnography,* Berkeley, University of California Press, pp. 1–26.

Cohen, R. (2014), "What's driving Google's obsession with artificial intelligence and robots?", *Forbes,* January 28. http://www.forbes.com/sites/reuvencohen/2014/01/28/whats-driving-googles-obsession-with-artificial-intelligence-and-robots/print/.

Conway, F., Siegelman, J. (2004), *Dark hero of the information age: In search of Norbert Wiener, the father of cybernetics,* New York, Basic Books.

Coombs, J.H., Renear, A.H., DeRose, S.J. (1987), "Markup system and the future of scholarly text processing", *Communications of the ACM* 30, 11, pp. 933–47. http://www.oasis-open.org/cover/coombs.html.

Copeland, B.J. (ed.) (2004), *The essential Turing,* Oxford, Clarendon Press.

Cosenza, G. (2008), *Semiotica dei nuovi media,* Rome–Bari, Laterza.

Craig, H. (2004), "Stylistic Analysis and Authorship Studies", Schreibman, S., Siemens, R., Unsworth, J. (eds.), *A companion to digital humanities,* Oxford, Blackwell, pp. 273–288. http://www.digitalhumanities.org/companion.

Crane, G. (2006), "What do you do with a million books", *D-Lib Magazine,* March, 12, 3.
Crane, G., Bamman, D., Jones, A. (2008), "ePhilology: When the books talk to their readers", in Siemens, R., Schreibman, S. (eds.), *A companion to digital literary studies,* Oxford, Blackwell. http://www.digitalhumanities.org/companionDLS/.
Cristofori, A. (2005), "Informatica umanistica e obiettivi didattici", *Comunicare storia,* 1. http://www.storicamente.org/04_comunicare/archivio_1.htm.
Crow, R. (2009), "Income models for Open Access: An overview of current practice", Washington, Scholarly Publishing & Academic Resources Coalition. http://www.sparc.arl.org/sites/default/files/incomemodels_v1.pdf.
Cultural Research in the Context of Digital Humanities (2013), Herzen State Pedagogical University (Saint-Petersburg), October, 3–5.
Dacos, M. (2013), "La stratégie du sauna finlandais: les frontières de Digital Humanities. Essai de géographie politique d'une communauté scientifique." https://hal.archives-ouvertes.fr/hal-00866107/document.
Dacos, M., Mounier, P. (2014), *Humanités Numériques: état des lieux et positionnement de la recherche française dans le contexte internationalle,* Paris, Institut Français. http://www.institutfrancais.com/sites/default/files/if_humanites-numeriques.pdf.
Dattolo, A., Duca, S., Tomasi, F., Vitali, F. (2009), "Towards disambiguating social tagging systems", in S. Murugesan (ed.), *Handbook of research on Web 2.0, 3.0 and X.0: Technologies, business and social applications,* Hershey (PA), IGI-Global, pp. 349–369.
Davis, M. (2000), *The universal computer: The road from Leibniz to Turing,* New York, Norton.
Day, M. (1996), "Mapping between metadata formats", *UKOLN: The UK Office for Library and Information Networking.* http://www.ukoln.ac.uk/metadata/interoperability.
De Beaugrande, R.A., Dressler, W.U. (1981), *Einführung in die Textlinguistik,* Tübingen, Niemeyer.
Deibert, R. (2015), "The geopolitics of Cyberspace after Snowden", *Current History. A Journal of Contemporary World Affairs,* 114, 768, pp. 9–15. http://www.currenthistory.com/Article.php?ID=1210.
Deleuze, G. (1990) "Post-scriptum sur les sociétés de contrôle", *L'autre journal,* 1, May. http://1libertaire.free.fr/DeleuzePostScriptum.html.
Deleuze, G. (1992), "Postscript on the societies of control", *October,* 59, Winter, pp. 3–7.
DeNardis, L. (2014), *The Global War for Internet Governance,* New Haven and London, Yale University Press.
Denny, P.J. (1991), "Rational thought in oral culture and literate decontextualization", in Olson, D.R., Torrance, N. (eds.), *Literacy and orality.* Cambridge, Cambridge University Press, pp. 66–89.
Derrida, J. (1996), *Archive fever: A Freudian impression* (trans. by Prenowitz, E.), Chicago, University of Chicago Press.

Descartes, R. (1985), *The philosophical works of Descartes,* vol. 1 (trans. by Cottingham, J., Stoothoff, R., Murdoch, D.), Cambridge–New York, Cambridge University Press.

Dobson, T.M., Willinsky, J. (2009), "Digital Literacy", in Olson, D., Torrance, N. (eds.), *Cambridge handbook of literacy,* Cambridge, Cambridge University Press, pp. 286–312.

Duranti, A. (1997), *Linguistic anthropology,* New York, Cambridge University Press.

Duranti, A. (2007), *Etnopragmatica: la forza nel parlare,* Rome, Carocci.

Durdağ, B. (2015), "Creating alternative communication spaces: Resistance, technology and social change". *5th ICTS and Society Conference 2015,* Vienna, June 3–7. http://fuchs.uti.at/wp-content/icts5/4_2.pdf.

Durkheim, E. (1982), *The rules of sociological method* (trans. by Halls, W.D.), New York, Free Press.

Dyer, R. (1969), "The new philology: An old discipline or a new science", *Computer and the humanities,* 4, 1, pp. 53–64.

Eduards, P.N. (1997), *The closed world,* Cambridge (MA), MIT Press.

Engelbart, D. (1962), "Augmenting human intellect: A conceptual framework", *SRI Summary Report AFOSR-3223. Prepared for Director of Information Sciences, Air Force Office of Scientific Research, Washington 25, DC, Contract AF 49(638)-1024, SRI Project 3578 (AUGMENT, 3906).* http://www.dougengelbart.org/pubs/augment-3906.html.

Etling, B., Faris, R., Palfrey, J. (2010), "Political change in the digital age: The fragility and promise of online organizing", *SAIS Review,* 30, 2, pp. 37–49.

European Science Foundation (2011), *Changing publication cultures in the humanities,* Young Researchers Forum, ESF Humanities Spring 2011, June 9–11, Maynooth. http://www.esf.org/fileadmin/Public_documents/Publications/Changing_Publication_Cultures_Humanities.pdf.

Fabbri, P., Marrone, G. (eds.) (2000), *Semiotica in nuce,* vol. 1: *I fondamenti e l'epistemologia strutturale,* Rome, Meltemi.

Falk, D. (2009), *Finding our tongue,* New York, Basic Book.

Faloutsos, C. (1985), "Access methods for text", *ACM Computing Surveys,* 17, 1, March, pp. 49–74.

Faloutsos, M., Faloutsos, P., Faloutsos, C. (1999), "On Power-Law relationships of the Internet Topology", *Proceedings of ACM SIGCOMM,* August, pp. 251–262. http://citeseer.ist.psu.edu/michalis99powerlaw.html.

Fiormonte, D. (2008a), "Pragmatica digitale: paratesti, microtesti e <metatesti> nel web", in *Testi brevi. Atti del convegno internazionale di studi,* Università Roma Tre, June 8–10, 2006, Dardano, M., Frenguelli, G., De Roberto, E. (eds.), Rome, Aracne, pp. 65–84.

Fiormonte, D. (2008b), "Il testo digitale: traduzione, codifica, modelli culturali", in Piras, P.R., Alessandro, A., Fiormonte, D. (eds.), *Italianisti in Spagna, ispanisti in Italia: la traduzione.* Atti del Convegno Internazionale (Rome, October 30–31, 2007), Rome, Edizioni Q, pp. 271–284.

Fiormonte, D. (2009), "Chi l'ha visto? Testo digitale, semiotica, rappresentazione. In margine a un trittico di Dino Buzzetti", *Informatica umanistica,* 2, pp. 21–63. http://www.ledonline.it/informatica-umanistica/Allegati/IU-02-09-Fiormonte.pdf.

Fiormonte, D. (ed.) (2011), *Canoni liquidi,* Napels, ScriptaWeb.

Fiormonte, D. (2012), "Towards a cultural critique of digital mumanities", *Historical Social Research / Historische Sozialforschung,* 37, 3, 2012, pp. 59–76.

Fiormonte, D. (2013), "Humanités digitales et sciences sociales: un mariage conclus dans les cieux", January 31. http://claireclivaz.hypotheses.org/249.

Fiormonte, D. (2014), "Digital humanities from a global perspective", *Laboratorio dell'ISPF,* 11. DOI: 10.12862/ispf14L203.

Fiormonte, D. (2015), "Lenguas, códigos, representación: márgenes de las humanidades digitales", in Priani, E., Galina,I. (eds.), *Las humanidades digitales.* Forthcoming.

Fiormonte, D., Matiradonna, V., Schmidt, D. (2010), "Digital encoding as a hermeneutic and semiotic act: The case of Valerio Magrelli", *Digital Humanities Quarterly,* 4, 1. http://digitalhumanities.org/dhq/vol/4/1/000082/000082.html.

Fiormonte, D., Numerico, T. (2011), "Le radici interdisciplinari dell'informatica: logica, linguistica e gestione della conoscenza", in Perilli, L., Fiormonte, D. (eds.), *La macchina nel tempo: studi di informatica umanistica in onore di Tito Orlandi,* Florence, Le Lettere, pp. 13–38.

Fiormonte, D., Schmidt, D. (2011), "La rappresentazione digitale della varianza testuale", in Fiormonte, D. (ed.), "Introduzione. Senza variazione non c'è cultura", *Canoni liquidi,* Naples, ScriptaWeb, pp. 161–180.

Fiormonte, D., Schmidt, D., Monella, P., Sordi, P. (2015), "The politics of code. How digital representations and languages shape culture", *ISIS Summit Vienna 201: The Information Society at the Crossroads,* June 3–7, Vienna.

Fitzpatrick, K. (2012), "Beyond metrics: Community authorization and open peer review", in Gold, M. K. (ed.), *Debates in the digital humanities,* Minneapolis, University of Minnesota Press, pp. 452–459.

Flanders, J., Fiormonte, D. (2007), "Markup and the digital paratext", *Digital Humanities,* June 2–8, Urbana–Champaign, University of Illinois. http://www.digitalhumanities.org/dh2007/dh2007.abstracts.pdf, pp. 60–61.

Ford, G., Kotzé, P. (2005), "Designing usable interfaces with cultural dimensions", *INTERACT'05. Proceedings of the 2005 IFIP TC13 international conference on human–computer interaction,* Berlin, Springer-Verlag, pp. 713–726.

Formenti, C. (2008), *Cybersoviet,* Milan, Raffaello Cortina Editore.

Fortunato, S., Flammini, A, Menczer, F., Vespignani, A. (2006), "The egalitarian effect of search engines", *WWW2006,* May 22–26, Edinburgh. http://arxiv.org/pdf/cs/0511005v2.pdf.

Foucault, M. (1969/1982), *The archaeology of knowledge,* London, Vintage.

Foucault, M. (1970), "What is an author?" (trans. by Bouchard, D.F., Simon, S.) (1977), *Language, Counter-Memory, Practice,* Ithaca (NY), Cornell University

Press, pp. 124–127. http://foucault.info/documents/foucault.authorFunction.en.html.

Franchi, S. Güzeldere, G. (2005), *Mechanical bodies, computational minds,* Cambridge (MA), Bradford Books.

Fuller, M. (2008), *Software studies: A lexicon,* Cambridge (MA), MIT Press.

Galina, I. (2013), "Is there anybody out there? Building a global digital humanities community". http://humanidadesdigitales.net/blog/2013/07/19/is-there-anybody-out-there-building-a-global-digital-humanities-community/

Galina, I. (2014), "Geographical and linguistic diversity in the Digital Humanities", *Literary and Linguistic Computing,* 29, 3, pp. 307–316.

Galloway, A.R. (2004), *Protocol: How control exists after decentralization,* Cambridge (MA), MIT Press.

Galloway, A.R. (2012), *The interface effect,* London, Polity Press.

Garret, J. R. (1995), "Task Force on Archiving Digital Information", *D-Lib Magazine,* September, http://www.dlib.org/dlib/september95/09garrett.html.

Garskova, I. (2014), "The past and present of digital humanities: A view from Russia", *H-Soz-Kult. Forum: The Status Quo of Digital Humanities in Europe,* October and November 2014. http://geschichte-transnational.clio-online.net/forum/type=diskussionen&id=2409.

Garzone, G. (2002), "Describing e-commerce communication: Which models and categories for text analysis?", *Textus,* 15, pp. 279–296.

Gazzola, M. (2012), "The linguistic implications of academic performance indicators: General trends and case study", *International Journal of the Sociology of Language,* 216, pp. 131–156.

Genet, J.-P. (1993), "La formation informatique des historiens en France: une urgence", *Mémoire vive,* 9 (1993), pp. 4–8.

Genette, G. (1997), *Paratexts: Thresholds of interpretation,* Cambridge–New York, Cambridge University Press.

Geri, L. (2007), *Ferrea Voluptas. Il tema della scrittura nell'opera di Francesco Petrarca,* Rome, Edizioni Nuova Cultura.

Ghonim, W. (2012), *Revolution 2.0: The power of the people is greater than the people in power. A memoir,* Boston (MA), Houghton Mifflin Harcourt.

Gillies, J., Cailliau, R. (2000) *How the Web was born: The story of the World Wide Web,* Oxford, Oxford University Press.

Gitelman, L. (ed.) (2013), *"Raw Data" is an oxymoron,* Cambridge (MA), MIT Press.

Gödel, K. (1931), "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I", *Monatshefte für Mathematik und Physik,* 38, pp. 173–198 [Eng. trans. "On formally undecidable propositions of the principia mathematica and related systems I", in Davis, M. (ed.), *The Unidecidable,* New York, Raven Press; reprint with corrections: London, Dover Publications, 2004, pp. 41–73.

Gold, M.K. (ed.) (2012), "The digital humanities moment", *Debates in The Digital Humanities,* Minneapolis–London, University of Minnesota Press, pp. ix–xvi.

Goldfarb, S. (1996), *The Roots of SGML: A personal recollection.* http://sgmlsource.com/history/roots.htm.

Golumbia, D. (2009), *The cultural logic of computation,* Cambridge (MA), Harvard University Press.

Golumbia, D. (2013), "Postcolonial studies, digital humanities, and the politics of language", *Postcolonial Digital Humanities,* May 31. http://dhpoco.org/blog/2013/05/31/postcolonial-studies-digital-humanities-and-the-politics-of-language/.

González-Blanco, E. (2013), "Actualidad de las humanidades digitales y un ejemplo de ensamblaje poético en la red: ReMetCa", *Cuadernos Hispanoamericanos,* 761, pp. 53–67.

Goodrum, A. A., O'Connor, B.C., Turner, J. M. (1999), "Introduction to the special topic issue of *Computers and the Humanities*: 'Digital Images'", *Computers and the Humanities,* 33, 4, pp. 291–292.

Goody, J. (1987), *The interface between the written and the oral,* Cambridge, Cambridge University Press.

Grisham R. (1994), Tipster Phase II Architecture Design Document (Strawman Architecture) Version 1.10. ftp://cs.nyu.edu/pub/nlp/tipster/100.tex.

Guha, R., Kumar, R., Raghavan, P., Tomkins, A. (2004), "Propagation of trust and distrust", *Proceedings of the WWW2004,* New York, ACM, pp. 403–412. http://www.w3.org/DesignIssues/Semantic.html.

Guo, J. (2014), "Electronic literature in China", *CLCWeb: Comparative Literature and Culture,* 16, 5. http://dx.doi.org/10.7771/1481-4374.2631.

Gyöngyi, Z., García- Molina, H. (2005), "Web spam taxonomy", *First International Workshop on Adversarial Information Retrieval on the Web (AIRWeb).* http://citeseer.ist.psu.edu/gyongyi05web.html.

Hafner, C., Lyon, M. (1996), *Where wizards stay up late,* New York, Simon & Schuster.

Hajič, J. (2004), "Linguistics Meets Exact Science", in Schreibman S., Siemens R., Unsworth J. (eds.), *A companion to digital humanities,* Oxford, Blackwell, pp. 79–82

Hall, S. (1980), "Encoding/decoding", *Culture, Media, Language: Working Papers in Cultural Studies 1972–79*, London, Hutchinson, pp. 128–138.

Hall, G. (2010), "'Follow the money': The political economy of open access in the humanities", http://garyhall.squarespace.com/journal/2010/11/19/on-the-limits-of-openness-i-the-digital-humanities-and-the-c.html.

Hand, M. (2014), "From cyberspace to the dataverse: Trajectories in digital social research", in Hand, M., Hillyard, S. (eds.), *Big data? Qualitative approaches to digital research,* Bingley, Emerald Group Publishing, pp. 1–27. http://dx.doi.org/10.1108/S1042-319220140000013002

Harel, D. (2000), *Computers Ltd.: What they really can't do,* Oxford, Oxford University Press.

Harris, R. (2000), *Rethinking writing,* London–New York, Continuum.

Harrison, K. D. (2010), *The last speakers. The quest to save the world's most endangered languages,* Washington, National Geographic.

Hauben, J (2007), "Libraries of the future 1945–1965", in Fuchs-Kittowski, K., Umstätter, W., Wagner-Döbler, R. (eds.), *Wissensmanagement in der Wissenschaft,* Berlin, Gesellschaft für Wissenschaftsforschung, pp. 103–117.

Hayles, N.K. (2008), *Electronic literature: New horizons for the literary,* Notre Dame (IN), University of Notre Dame Press.

Heims, S.J. (1991), *The cybernetics group,* Cambridge (MA), MIT Press.

Held, G. (2005), "A proposito di una nuova testualità. Osservazioni semiotiche e linguistiche sulla base dei testi multimodali nella stampa odierna", *Italienisch,* 54, pp. 46–63.

Hendler, J., Shadbolt, N., Hall, W., Berners-Lee, T., Weitzner, D. (2008), "Web science: An interdisciplinary approach to understanding the web", *Communication of the ACM,* 2008, 51, 7, pp. 60–69.

Hess, C., Ostrom, E. (2007), *Understanding knowledge as a commons,* Cambridge (MA), MIT Press.

Hilbert, M., López, P. (2011), "The world's technological capacity to store, communicate, and compute information", *Science,* 332, no. 6025, pp. 60–65. DOI: 10.1126/science.1200970

Hill, R. (2015), "The true stakes of Internet governance", in Buxton, N., Bélanger Dumontier, M. (eds.), *State of Power 2015: An annual anthology on global power and resistance,* Amsterdam, The Transnational Institute, pp. 28–37. http://www.tni.org/stateofpower2015.

Hiltzik, M.A. (1999), *Dealers of lightning: Xerox Parc and the dawn of the computer age,* New York, Harper Business.

Hjelmslev, L. (1943), *Omkring sprogteoriens Grundlaeggelse,* Copenhagen, Bianco Lunos Bogtryk.

Holmes, D.I. (1994), "Authorship attribution", *Computers and the Humanities,* 28, 2, pp. 87–106.

Honn, J. (2013), "Never neutral: Critical approaches to digital tools & culture in the humanities", October 17. http://joshhonn.com/?p=1.

Hoover, D. (2008), "Quantitative analysis and literary studies", in Schreibman, S., Siemens, R., Unsworth, J. (eds.), *A companion to digital literary studies,* Oxford, Blackwell. http://www.digitalhumanities.org/companionDLS/.

Howard, P. N., Hussain, M. M. (2011), "The upheavals in Egypt and Tunisia: The role of digital media." *Journal of Democracy,* 22, 3, pp. 35–48

Hume, D. (1739/1978), *Treatise on human nature* (Ed. by Selby-Bigge, L. A., Nidditch, P. H.), Oxford, Clarendon Press; New York, Oxford University Press.

IAB – GfK Media & Entertainment (2014), "Original Digital Video Consumer Study", April. http://www.iab.net/media/file/GfKIAB2014OriginalDigitalVideoReport.pdf

Ide, N. (2004), "Preparation and Analysis of Linguistic Corpora", in Schreibman, S., Siemens, R., Unsworth, J. (eds.), *A companion to digital humanities,* Oxford, Blackwell, pp. 298–305 http://www.digitalhumanities.org/companion/.

Innis, A.H. (1951), *The bias of communication,* Toronto, University of Toronto Press.
Jablonka, E., Lamb, M. J. (2005), *Evolution in four dimensions: Genetic, epigenetic, behavioral, and symbolic variation in the history of life,* Cambridge (MA), MIT Press.
Jamali, R. (2014), *Online Arab spring: Social media and fundamental change,* London, Chandos Publishing.
Jameson, F. (1991), *Postmodernism, or the cultural logic of late capitalism,* Durham, Duke University Press.
Jaynes, J. (1976), *The origin of consciousness in the breakdown of the bicameral mind,* Boston (MA), Houghton Mifflin.
Jenkins, H. (2006), *Convergence culture: Where old and new media collide,* New York, New York University Press.
Jobin, A. (2013), "Google's autocompletion: Algorithms, stereotypes and accountability", October 22. http://sociostrategy.com/2013/googles-autocompletion-algorithms-stereotypes-accountability/.
Johnson, S. (1997), *Interface culture,* New York, Basic Books.
Jones, S.E. (2014), *The emergence of the digital humanities,* New York and London, Routledge.
Kahn, R., Wilensky, R. (1995/2006), "A framework for distributed digital object services", *International Journal on Digital Libraries,* 6, 2, pp. 115–123. http://doi.info/topics/2006_05_02_Kahn_Framework.pdf.
Keller, E.F. (1995), *Refiguring life,* New York, Columbia University Press.
Keller, E.F. (2010), *The mirage of a space between nature and nurture,* Durham–London, Duke University Press.
Kelly, K. (2005), "We Are the Web", *Wired,* 13, 8. http://www.wired.com/wired/archive/13.08/tech.html.
Kirschenbaum, M.G. (2002), "Editor's introduction: Image-based humanities computing", *Computer and Humanities,* 36, 1, pp. 3–6.
Kirschenbaum, M.G. (2008), *Mechanisms: New media and the forensic imagination.* Cambridge, (MA), MIT Press.
Kitchin, R. (2014), "Big Data, new epistemologies and paradigm shifts", *Big Data and Society,* April–June, pp. 1–12.
Kleinberg, J. (1999), "Authoritative sources in a hyperlinked environment", *Proceedings of the ACM-SIAM symposium on discrete algorithms.* Extended version published in *Journal of the ACM* (1999), 46, pp. 604–632.
König, R., Rasch, M. (eds.) (2014), *Society of the query reader: Reflections on Web search,* Amsterdam, Institute of Network Cultures. http://networkcultures.org/wp-content/uploads/2014/04/SotQreader_def_scribd.pdf.
Kress, G., van Leeuwen, T. (2006), *Reading images: The grammar of visual design,* London–New York, Routledge.
Laakso, M., Welling, P., Bukvova, H., Nyman, L., Björk, B-C., *et al.* (2011), "The development of open access journal publishing from 1993 to 2009", *PLoS ONE,* 6, 6.

Lafuente, A., Alonso, A., Rodríguez, J. (2013), *¡Todos sabios!: ciencia ciudadana y conocimiento expandido,* Madrid, Cátedra.
Lakatos, I. (1971/1978) "History of Science and its Rational Reconstructions", in Buck, R. C. and Cohen, R. S. (eds.), P.S.A. (1970) *Boston Studies in the Philosophy of Science,* 8, pp. 91–135. Dordrecht: Reidel. Republished as chapter 2 of *The methodology of scientific research programmes,* Vol. 1, Cambridge, Cambridge University Press, 1978, pp. 102–138.
Landow, G.P. (1991), "The rhetoric of hypermedia: Some rules for authors", Delany, P., Landow, G.P. (eds.), *Hypermedia and literary studies,* Cambridge (MA), MIT Press, pp. 81–103.
Landow, G.P. (1992), *Hypertext: The convergence of contemporary critical theory and technology,* Baltimore, John Hopkins University Press.
Landow, G.P. (2006), *Hypertext 3.0,* Baltimore, John Hopkins University Press.
Lanham, R. (2006), *The economics of attention: Style and substance in the age of information,* Chicago, University of Chicago Press.
Lanier, J. (2010), *You are not a gadget,* London, Allen Lane.
Lanier, J. (2013), *Who owns the future?,* New York, Simon & Schuster.
Latour (1988), *Science in action: How to follow scientists and engineers through society,* Cambridge (MA), Harvard University Press.
Lawrence, S., Giles, C.L. (1998), "Searching for the World Wide Web", *Science,* 280, pp. 98–100.
Leerssen, J. (2012), "The rise of philology: The comparative method, the historicist turn and the surreptitious influence of Giambattista Vico", in Bod, R., Maat, J., Weststeijn, T. (eds.), *The making of the humanities,* vol. 2: *From early modern to modern disciplines.* Amsterdam, Amsterdam University Press, pp. 23–35.
Lessig, L. (2004), *Free culture,* New York, Penguin.
Lévi-Strauss, C. (1960), "Le champe de l'anthropologie", *Annuaire du Collège de France,* January 5, col. n. 31.
Lew, M.S., Sebe, N., Djeraba, C., Jain, R. (2006), "Content-based multimedia information retrieval: State of the art and challenges", *ACM Transactions on Multimedia Computing, Communications and Applications,* 2, 1, February, pp. 1–19.
Licklider, J.C.R. (1960), "Man-computer symbiosis", *IEEE Transactions on Human Factors in Electronics,* HFE-1, March, pp. 4–11. http://memex.org/licklider.pdf.
Licklider, J.C.R. (1962), "The computer in the university", in Greenberger, M., (ed.), *Computers and the world of the future,* Cambridge (MA), MIT Press, pp. 203–209.
Licklider, J.C.R. (1963), "Memorandum for Members and Affiliates of the Intergalactic Computer Network", April 23. http://www.chick.net/wizards/memo.html.
Licklider, J.C.R. (1965), *Libraries of the future,* Cambridge (MA), MIT Press.
Licklider, J.C.R. (1988), "An interview with J.C.R. Licklider conducted by W. Aspray and A. Norberg on 28 October 1988", Cambridge (MA). http://conservancy.umn.edu/bitstream/handle/11299/107436/oh150jcl.pdf.
Licklider, J.C.R., Taylor, R.W. (1968), "The computer as a communication device", *Science and Technology,* April. http://memex.org/licklider.pdf.

LIENS (2008), "Position du LIENS au sujet de la bibliométrie". http://www.di.ens.fr/users/longo/files/Data/lettre-bibliometrie.pdf.
Liu, A. (2012), "Where is cultural criticism in the digital humanities?", in Gold, M.K. (ed.), *Debates in the digital humanities,* Minneapolis, University of Minnesota Press, pp. 490–509.
Liu, A. (2013), "Why i'm in it" x 2: Antiphonal response to stephan ramsay on digital humanities and cultural criticism". http://liu.english.ucsb.edu/why-im-in-it-x-2-antiphonal-response-to-stephan-ramsay-on-digital-humanities-and-cultural-criticism/.
Longo, G. (2009a), Incompletezza, in *La Matematica,* vol. 4, Turin, Einaudi (in print).
Longo, G. (2009b), "Critique of computational reason in the natural sciences", Gelenbe, E., Kahane, J.-P. (eds.), *Fundamental concepts in computer science,* London, Imperial College Press/World Scientific. Text originally written in Italian as "Lezione Galileana", Pisa, October 25, 2006, pp. 43–69. ftp://ftp.di.ens.fr/pub/users/longo/PhilosophyAndCognition/CritiqCompReason-engl.pdf.
Lothian, A., Phillips, A. (2013), "Can digital humanities mean transformative critique," *Journal of E-Media Studies,* 3, 1. http://journals.dartmouth.edu/cgi-bin/WebObjects/Journals.woa/xmlpage/4/article/425.
Lotman, J.M. (2006), *Tesi per una semiotica delle culture,* Rome, Meltemi.
Lovink, G. (2007), *Zero comments,* London, Routledge.
Lovink, G., Rasch, M. (eds.) (2013), *Unlike us reader: Social media monopolies and their alternatives,* Amsterdam, Institute of Network Cultures. http://networkcultures.org/blog/publication/unlike-us-reader-social-media-monopolies-and-their-alternatives/.
Lucía Megías, J.M. (2003), "La informática humanística: notas volanderas desde el ámbito hispánico", *Incipit,* 23, pp. 91–114.
Lucía Megías, J.M. (2012), *Elogio del texto digital. Claves para interpretar el nuevo paradigma,* Madrid, Fórcola.
Lughi, G. (2006), *Culture dei nuovi media: teorie, strumenti, immaginario,* Milan, Guerini e Associati.
Luria, A.L. (1979), "Cultural differences in thinking", *The making of mind: A personal account of Soviet psychology,* Cambridge (MA)–London, Harvard University Press. http://www.marxists.org/archive/luria/works/1979/mind/ch04.htm.
Lutoslawski, W. (1897), *The origin and growth of Plato's logic,* London, Longmans.
Luzzatto, M.T. (1988), "L'oratoria, la retorica e la critica letteraria dalle origini a Ermogene", Montanari, F. (ed.), *Da Omero agli Alessandrini: Problemi e figure della letteratura greca,* Rome, La Nuova Italia Scientifica, pp. 208–256.
MacKenzie, D.A., Wajcman, J. (1999), *The social shaping of technology,* Buckingham–Philadelphia, Open University Press.
MacKinnon, R., Hickok, E., Bar, A., Lim, H. (eds.) (2014), *Fostering freedom online: The role of Internet intermediaries,* Paris, United Nations Educational, Scientific and Cultural Organization and Internet Society. http://unesdoc.unesco.org/images/0023/002311/231162e.pdf.

Madden, M., Smith, A. (2010), "Reputation management and social media", *Pew Internet and American Life Project,* May 26. http://pewinternet.org/Reports/2010/Reputation-Management.aspx.
Maffi, L. (2001) (ed.), *On biocultural diversity: Linking language, knowledge, and the environment.* Washington–London, Smithsonian Institution Press.
Maffi, L. (2010), "What is biocultural diversity?", in Maffi, L., Woodley, E. (eds.), *Biocultural diversity conservation. A global sourcebook,* Washington–London, Earthscan.
Mahoney, M.S. (2000), "The structures of computation", in Rojas, R., Hashagen, U. (eds.), *The first computers: History and architecture,* Cambridge (MA), MIT Press, pp. 17–32.
Mahoney, M.S. (2005), "The histories of computing(s)", *Interdisciplinary Science Reviews,* 30, 2, pp. 119–135.
Manovich, L. (2001), *The language of new media,* Cambridge (MA), MIT Press.
Manovich, L. (2013), *Software takes command: Extending the language of new media,* London, Bloomsbury Publishing.
Maragliano, R. (2004), "Insegnare a scrivere con il computer", *Quaderni di didattica della scrittura,* 1, pp. 49–57.
Marchiori, M. (1997), "The Quest for Correct Information on the Web: Hyper Search Engines", *Proceedings of the Sixth International World Wide Web Conference (WWW6),* Santa Clara (CA), pp. 265–276. http://www.w3.org/People/Massimo/papers/WWW6/.
Marcos Marín, F. (1985), "Computer-assisted philology: Towards a unified edition of Osp. Libro de Alexandre", *Proceedings of the European Language Services Conference on Natural-Language Applications,* section 16, Copenhagen, IBM Denmark.
Marcos Marín, F. (1996), *El comentario filológico con apoyo informático,* Madrid, Editorial Síntesis.
Marino, M.C. (2006), "Critical code studies", *Electronic Book Review,* December 4. http://electronicbookreview.com/thread/electropoetics/codology.
Marino, M.C. (2014), "Field report for critical code Studies", *Computational Culture,* November 9. http://computationalculture.net/article/field-report-for-critical-code-studies-2014.
Mason, R.O. (1986), "Four ethical issues of the information age", *Management Information Systems Quarterly,* 10, 1, pp. 5–12.
Markham, A.N. (2013), "Undermining 'data': A critical examination of a core term in scientific inquiry", *First Monday,* 18, 10, October. http://firstmonday.org/ojs/index.php/fm/article/view/4868/3749.
Mateas, M., Stern, A. (2005), "Structuring content in the façade interactive drama architecture", *Proceedings of Artificial Intelligence and Interactive Digital Entertainment (AIIDE 2005),* Marina del Rey, June. http://www.aaai.org/Papers/AIIDE/2005/AIIDE05-016.pdf.
Mattozzi, A. (ed.) (2006), *Il senso degli oggetti tecnici,* Rome, Meltemi.
Maturana, H., Varela, F. (1980), *Autopoiesis and cognition: The realization of the living,* Dordrecht, Reidel.

Mayer-Schönberger, V., Cukier, K. (2013), *Big data: A revolution that will transform how we live, work, and think,* London, John Murray.

McCarty, W. (1999), "Humanities computing as interdiscipline", *Is humanities computing an academic discipline?,* October 22. http://www.iath.virginia.edu/hcs/mccarty.html.

McCarty, W. (2005), *Humanities computing,* Basingstoke, Palgrave.

McGann, J.J. (2001), *Radiant textuality: Literature after the World Wide Web,* New York, Palgrave.

McKenzie, D.F. (1986), *Bibliography and the sociology of texts,* London, The British Library.

McLuhan, M. (2006), *The classical trivium: The place of Thomas Nashe in the learning of his time* (ed. T. Gordon), Corte Madera (CA), Gingko Press.

McPherson, T. (2012), "Why are the digital humanities so white? Or thinking the histories of race and computation", in Gold, M. (ed.), *Debates in the digital humanities,* Minneapolis, University of Minnesota Press, 139–160.

Meddeb, A. (2011), *Printemps de Tunis: la métamorphose de l'histoire,* Paris, Éditions Albin Michel.

Metitieri, F. (2009), *Il grande inganno del Web 2.0,* Rome–Bari, Laterza.

Michel, J.-B., Shen, Y. K., Presser Aiden, A., Veres, A., Gray, M.K., The Google Books Team, Pickett, J.P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M.A., Aiden, E.L. (2011), "Quantitative analysis of culture using millions of digitized books", *Science,* 331, 6014, January 14, pp. 176–182.

Mijksenaar, P. (1997), *Visual function: An introduction to information design,* New York, Princeton Architectural Press.

Mikami, Y., Kodama, S. (2012), "Measuring linguistic diversity on the Web", in Vannini, L., Le Crosnier, H. (eds.), *Net.Lang: Towards the multilingual cyberspace,* Caen, C&F Éditions, pp. 121–139.

Miles, A. (2002), "Hypertext structure as the event of connection", *Journal of Digital Information,* 2, 3. http://journals.tdl.org/jodi/article/view/48/51.

Millán, J.A. (2001), *Internet y el español,* Madrid, Retevision.

Mitchell Waldrop, M. (2002), *The dream machine: J.C.R. Licklider and the revolution that made computing personal,* New York, Penguin.

Mnih V., Kavukcuoglu, K., Hassabis, D., *et al.* (2015), "Human-level control through deep reinforcement learning", *Nature,* February 26, 518, pp. 529–533.

Morais, S. (2013), "Theorizing the digital subaltern", August 2. http://cis-india.org/raw/theorizing-the-digital-subaltern.

Mordenti, R. (2003), "L'insegnamento dell'Informatica nelle Facoltà Umanistiche", *XI Meeting on Humanities Computing/Informatica Umanistica,* University of Verona and Fondazione Ezio Franceschini, February 27–28. http://infolet.it/files/2009/09/mordenti_2003.pdf.

Mordenti, R. (2011), "Parádosis: a proposito del testo informatico", *Atti della Accademia Nazionale dei Lincei,* Anno CDVIII, Classe di Scienze Morali, Storiche e Filologiche. Memorie, Serie IX, Volume XXVIII, Fascicolo 4, pp. 617–692.

Moretti, F. (2013), *Distant reading,* London, Verso.

Mounier, P. (ed.) (2012), *Read/Write, Book 2: une introduction aux humanités numériques,* Marseille, OpenEdition Press. http://books.openedition.org/oep/240.
Munari, B. (1968), *Design e comunicazione visiva,* Rome–Bari, Laterza.
Munari, B. (1971), *Artista e designer,* Rome–Bari, Laterza.
Musil, R. (1996), *The man without qualities 1: A sort of introduction and pseudo reality prevails,* Vintage, New York.
National Science Foundation (2011), *Changing the conduct of science in the information age,* summary report of workshop held on November 12, 2010, Stanford University, National Science Foundation. http://www.nsf.gov/pubs/2011/oise11003/index.jsp.
Naughton, J. (1999), *A brief history of the future: The origins of the Internet,* London, Weifenfeld & Nicolson.
Negroponte, N. (1995), *Being digital,* New York, Random House.
Nelson, T.H. (1965), "A File structure for the complex, the changing and the indeterminate, association for computing machinery", in Winner, L. (ed.), *Proceedings of the 20th National Conference,* pp. 84–100. Reprinted in Wardrip-Fruin, N., Montfort, N. (eds.) (2003), *The new media reader,* Cambridge (MA), MIT Press, pp. 133–146.
Nelson, T.H. (1987), *Literary machines: Edition 87.1,* Sausalito Press.
Nicolaescu, M., Mihai, A. (2014), "Teaching digital humanities in Romania", *CLCWeb: Comparative Literature and Culture,* 16, 5. http://dx.doi.org/10.7771/1481-4374.2497.
Nielsen, J. (1996), "Inverted pyramids in cyberspace". http://www.useit.com/alertbox/9606.html.
Nielsen, J. (1999), *Designing web usability,* Indiana, New Riders.
Nielsen, M. (2011), *Reinventing discovery: The new era of networked science,* Princeton, Princeton University Press.
Niero, I. (2013), "DH in Russia: prove di dialogo epistemologico". http://infolet.it/2013/12/11/dh-in-russia-prove-di-dialogo-epistemologico/.
Nietzsche, F.W. (1980), *On the advantage and disadvantage of history for life* (trans. by Preuss, P.), Indianapolis, Hackett.
Norman, D. (1998), *The invisible computer,* Cambridge (MA), Mit Press.
Norman, D. (2007), *The design of future things,* New York, Basic Books.
Nowotny, H., Testa, G. (2011), *Naked Genes,* Cambridge (MA), MIT Press.
Numerico, T. (2005), *Alan Turing e l'intelligenza delle macchine,* Milan, FrancoAngeli.
Nybom, T. (2003), "The Humboldt legacy: Reflections on the past, present, and future of the european university", *Higher Education Policy,* 16, pp. 141–59.
Nyce, J., Kahn, P. (eds.) (1991), *From Memex to hypertext: Vannevar Bush and the mind's machine.* Boston, Academic Press.
O'Donnell, D. P. (2013), "The true north strong and hegemonic: Or, why do Canadians seem to run DH". http://people.uleth.ca/~daniel.odonnell/Blog/the-true-north-strong-and-hegemonic-or-why-do-canadians-seem-to-run-dh.

O'Gorman, M. (2000), "You can't always get what you want: Transparency and deception on the computer fashion scene", *CTheory,* June 12. http://www.ctheory.net/articles.aspx?id=227.
Olson, D. R. (2012), "Literacy, rationality, and logic. The historical and developmental origins of logical discourse", in Joyce, Terry and David Roberts (eds.), *Units of language – Units of writing.* Special issue of *Written Language and Literacy,* 15, 2, pp. 153–164.
Ong, W. J. (1982), *Orality and literacy: The technologizing of the word,* London–New York, Methuen.
O'Reilly, T. (2005), "What is Web 2.0: Design patterns and business models for the next generation of software", September 30. http://oreilly.com/pub/a/web2/archive/what-is-web-20.html.
Origgi, G. (2010), "Epistemic vigilance and epistemic responsibility in the liquid world of scientific publications", *Social Epistemology,* 24, 3, pp. 149–159.
Origgi, G., Simon, J. (2010), "Scientific publications 2.0. The end of the scientific paper?", *Social Epistemology,* 24, 3, pp. 145–148.
Orlandi, T. (2002), "Is humanities computing a discipline?", *Jahrbuch für Computerphilologie,* 4. http://www.computerphilologie.uni-muenchen.de/jg02/orlandi.html.
Orlandi, T. (2010), *Informatica testuale,* Rome–Bari, Laterza.
Pajares Tosca, S. (2000), "A Pragmatics of Links", *Journal of Digital Information,* 1, 6, Article 22, June 27. https://journals.tdl.org/jodi/index.php/jodi/article/view/23/24.
Pariser, E. (2011), *The filter bubble: What the internet is hiding from you,* New York, Penguin Press.
Pasolini, P.P. (1967), *Empirismo eretico,* Milan, Garzanti; Eng. trans. *Heretical empiricism* (1988), Bloomington, Indiana University Press.
Pellizzi, F. (1999), "Per una critica del link". *Bollettino '900,* 2. II. http://www3.unibo.it/boll900/numeri/1999-ii/Pellizzi.html.
Pérez Álvarez, S. (2007), *Sistemas CBIR: recuperación de imágenes por rasgos visuales,* Gijón, Ediciones Trea.
Perri, A. (2009), "Al di là della tecnologia, la scrittura. Il caso Unicode", *Annali dell'Università degli Studi Suor Orsola Benincasa,* 11, pp. 725–748.
Petöfi, J.S. (2005), "Approcci semiotico-testologici ai testi Multimediali", in Tursi, A. (ed.), *Mediazioni: spazi, linguaggi e soggettività delle reti,* Genua, Costa & Nolan, pp. 94–109.
Petrucci, A. (1986), *La scrittura: ideologia e rappresentazione,* Turin, Einaudi.
Petrucci, A. (1998), "Scritture marginali e scriventi subalterni", in Simone, R., Albano Leoni, F., Gambarara, D., Gensini, S., Lo Piparo, F. (eds.), *Ai limiti del linguaggio: vaghezza, significato, storia,* Rome–Bari, Laterza, pp. 311–318.
Petzold, C. (2008), *The annotated Turing,* Indianapolis, Wiley Publishing.
Pickering, A. (1995), *The mangle of practice,* Chicago, University of Chicago Press.

Piez, W. (2008), "Something called 'digital humanities'", *Digital Humanities Quarterly* 2, 1. http://www.digitalhumanities.org/dhq/vol/2/1/000020/000020.html.

Piez, W. (2013) "Markup beyond XML" *Conference Abstracts Digital Humanities 2013*. Lincoln, Nebraska, July. http://dh2013.unl.edu/abstracts/ab-175.html.

Pinto, R. (2005), "Genealogías postmodernas: Jameson y Pasolini", in Borràs Castanyer, L. (ed.), *Textualidades electrónicas: nuevos escenarios para la literatura,* Barcelona, Editorial UOC, pp. 247–273.

Plaisant, C., Rose, J., Yu, B., Auvil, L., Kirschenbaum, M.G., Smith, M.N., Clement, T., Lord, G. (2006), "Exploring erotics in Emily Dickinson's correspondence with text mining and visual interfaces", *JCDL'06*, June 11–15, Chapel Hill (NC). http://hcil2.cs.umd.edu/trs/2006-01/2006-01.pdf.

Plato (1949), *Meno* (trans. by Jowett, B.), London, Bobbs-Merrill Co.

Pochoda, P. (2013), "The big one: The epistemic system break in scholarly monograph publishing", *New Media & Society,* 15, pp. 359–378.

Pons, Anaclet (2013), *El desorden digital: guía para historiadores y humanistas.* Madrid, Siglo XXI España.

Potter, M. (2002), "XML for digital preservation: XML implementation options for e-mails", *Reports on progress at the Digital Preservation Testbed (Testbed Digital Bewaring) of the Netherlands in using XML as a preservation approach.* http://www.digitaleduurzaamheid.nl/bibliotheek/docs/email-xml-imp.pdf.

Prada, M. (2003), "Lingua e web", Bonomi, I. *et al.* (eds.), *La lingua italiana e i mass media,* Rome, Carocci, pp. 249–289.

PREMIS (2008), "PREMIS Data Dictionary for Preservation Metadata". http://www.loc.gov/standards/premis/v2/premis-2-0.pdf.

Presner, T. (2012), "Critical theory and the mangle of digital humanities". http://www.toddpresner.com/wp-content/uploads/2012/09/Presner_2012_DH_FINAL.pdf.

Pressman, J. (2014), *Digital modernism: Making it new in new media,* Oxford University Press, 2014.

Procter, J. (2004), *Stuart Hall,* London–New York, Routledge.

Purcell, K., Rainie, L., Heaps, A., Buchanan, J., Friedrich, L., Jacklin, A., Chen, C., Zickuhr, K. (2012), "How teens do research in the digital world", *Pew Research Center,* November 1. http://www.pewinternet.org/2012/11/01/how-teens-do-research-in-the-digital-world/.

Purcell K., Rainie, L. (2014), "Americans feel better informed thanks to the Internet", *Pew Research Center,* December 8. http://www.pewinternet.org/2014/12/08/better-informed.

Quinnell, S.L. (2012), "Digital social science vs. digital humanities: Who does what & does it matter?" http://www.socialsciencespace.com/2012/07/digital-social-science-vs-digital-humanities-who-does-what-does-it-matter/.

Rabinow, P. (1986), "Representations are social facts: Modernity and post-modernity in anthropology", in Clifford, J., Marcus, G.E. (eds.), *Writing culture: The poetics and politics of ethnography,* Berkeley, University of California Press, 234–259.

Renear, A. (1997), "Out of praxis: Three (meta)theories of textuality", in K. Sutherland (ed.), *Electronic text,* Oxford, Clarendon Press, pp.107–126.

Renear, A., Mylonas, E., Durand, D. (1992/1993), *Refining our notion of what text really is.* http://cds.library.brown.edu/resources/stg/monographs/ohco.html.

Renear, A., McGann, J.J., Hockey, S. (1999), "What is text? A debate on the philosophical and epistemological nature of text in the light of humanities computing research", Panel with Renear, A., McGann, J., Hockey, S. (Chair), *ACH/ALLC '99,* June, Charlottesville (VA).

Renear, A., Dubin, D., Sperberg-McQueen, C.M., Huitfeldt, C. (2003), "Towards a semantics for XML markup", in Furuta, R., Maletic, J.I., Munson, E. (eds.), *Proceedings of the 2002 ACM symposium on document engineering,* New York, ACM Press, pp. 119–126.

Risam, R. (2014), "Rethinking peer review in the age of digital humanities", *Ada: A Journal of Gender, New Media, and Technology,* 4. DOI: 10.7264/N3WQ0220

Risam, R., Koh, A. (2013), "Open thread: The digital humanities as a historical 'refuge' from race/class/gender/sexuality/disability?". http://dhpoco.org/blog/2013/05/10/open-thread-the-digital-humanities-as-a-historical-refuge-from-raceclassgendersexualitydisability/.

Roburn, S. (1994), "Literacy and the underdevelopment of knowledge", *Mediatribe: Concordia University's Undergraduate Journal of Communication Studies,* 4, 1. http://web.archive.org/web/19970504003533/http://cug.concordia.ca/~mtribe/mtribe94/native_knowledge.html.

Rockwell, G. (2003), "What is text analysis really?", *Literary and Linguistic Computing,* 18, 2, pp. 209–219.

Rodríguez Ortega, N. (2013), "Humanidades Digitales, Digital Art History y cultura artística: relaciones y desconexiones", in Alsina P. (ed.), *Historia(s) del arte de los medios. Artnodes,* 13, pp. x–xx. http://journals.uoc.edu/ojs/index.php/artnodes/article/view/n13-rodriguez/n13-rodriguez-es.

Röhrs, H. (1987), "The classical idea of the university. Its origin and significance as conceived by Humboldt", in Röhrs, H. (ed.), *Tradition and reform of the university under an international perspective,* New York, Peter Lang, pp. 13–27.

Rojas Castro, A. (2013), "El mapa y el territorio: una aproximación histórico-bibliográfica a la emergencia de las Humanidades Digitales en España", *Caracteres: estudios culturales y críticos de la esfera digital,* 2, 2, pp. 10–52. http://revistacaracteres.net/revista/vol2n2noviembre2013/el-mapa-y-el-territorio/.

Rojas Castro, A. (2014), "¿Global DH? Hablemos de dinero", August 27. http://www.antoniorojascastro.com/global-dh-hablemos-de-dinero.

Romero Frías, E. (2014), "Ciencias Sociales y Humanidades Digitales: una visión introductoria", in Romero Frías, E., Sánchez González, M. (eds.), *Ciencias Sociales y Humanidades Digitales. Técnicas, herramientas y experiencias de e-Research e investigación en colaboración,* CAC: Cuadernos Artesanos de Comunicación, 61, 2014, pp. 19–50. http://www.cuadernosartesanos.org/2014/cac61.pdf.

Romero Frías, E., Del Barrio García, S. (2014), "Una visión de las Humanidades Digtiales a través de sus centros", *El profesional de la Información,* 23, 5, September–

October, pp. 485–492. http://www.elprofesionaldelainformacion.com/contenidos/2014/sept/05.html.
Ronchi, R. (2008), *Filosofia della comunicazione,* Turin, Bollati Boringhieri.
Rorty, R. (ed.) (1967), *The linguistic turn: Recent essays in philosophical method,* Chicago, University of Chicago Press.
Rosati, L., Venier, F. (eds.) (2005), *Rete e retorica: prospettive retoriche della Rete,* Perugia, Guerra.
Rose, S. (1998), *Lifelines: Biology, freedom, determinism,* London, Penguin.
Ross, S. (2014), "In Praise of overstating the case: A review of Franco Moretti, *Distant reading* (London, Verso, 2013)", *Digital Humanities Quarterly,* 8, 1, 2014. http://www.digitalhumanities.org/dhq/vol/8/1/000171/000171.html.
Rothenberg, J. (1999), "Avoiding technological quicksand: Finding a viable technical foundation for digital preservation", Washington (DC), Council on Library and Information Resources. http://www.clir.org/pubs/reports/rothenberg/pub77.pdf.
Ruiz, É., Heimburger, F. (2011), "Faire de l'histoire à l'ère numérique: retours d'expériences", *Revue d'histoire moderne et contemporaine,* 58, 4bis, (2011), pp. 70–89.
Russell, S., Norwig, P. (1995/2009), *Artificial intelligence: A modern approach,* Upper Saddle River (NJ), Prentice Hall.
Sahle, P. (2013), *Digitale Editionsformen: Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels.* Teil 3: *Textbegriffe und Recodierung*. Schriften des Instituts für Dokumentologie und Editorik – Band 9, Norderstedt, BoD. http://kups.ub.uni-koeln.de/5013/.
Salton, G. (1989), *Automatic text processing: The transformation, analysis, and retrieval of information by computer,* Boston (MA), Addison-Wesley Longman Publishing.
Sánchez, G. D. (2014), "Hacia una mirada cualitativa de la brecha digital. Retos y oportunidades en dos estudios de caso en México y EUA", in Matus Ruiz, M. (ed.), *El valor de la etnografía para el diseño de productos, servicios y políticas TIC. Memoria del Seminario,* Centro de Investigación e Innovación en Tecnologías de la Información y Comunicación (INFOTEC), México (DF), pp. 85-91.
Sanderson, R., Ciccarese, P., Van de Sompel, H. (2013), *Open annotation data model.* http://www.openannotation.org/spec/core.
Sano-Franchini, J. (2015), "Cultural rhetorics and the digital humanities: Toward cultural reflexivity in digital making", in Ridolfo, J., Hart-Davidson, W. (eds.), *Rhetoric and the digital humanities,* Chicago, University of Chicago Press, pp. 49–64.
Sánz, A. (2013), "Digital humanities or hypercolonial studies?", *RICT: Responsible Innovation,* 2013. http://observatory-rri.info/sites/default/files/obs-technology-assessment/Final%20_sanz_hypercolonial_sent3.pdf.
Schmidt, D. (2010), "The inadequacy of embedded markup for cultural heritage texts", *Literary and Linguistic Computing,* March 25, pp. 337–356.

Schmidt, D. (2012), “The role of markup in the digital humanities,” *Historical Social Research Special Issue,* 37, 3, 125–146.
Schmidt, D. (2014) “Towards an interoperable digital scholarly edition”, *Journal of the Text Encoding Initiative,* 7. http://jtei.revues.org/979.
Schmidt, D., Colomb, R. (2009), “A data structure for representing multi-version texts online”, *International Journal of Human–Computer Studies,* 67, 6, pp. 497–514.
Scolari, C. (2004), *Hacer clic: hacia una sociosemiótica de las interacciones digitales,* Barcelona, Gedisa.
Segre, C. (1988), *Introduction to the analysis of the literary text,* Bloomington (IN), Indiana University Press.
Selfe, C.L., Selfe, R.J. Jr. (1994), “The Politics of the interface: Power and its exercise in electronic contact zones”, *College Composition and Communication,* 45, 4 pp. 480–504.
Shah, N. (2010), “Internet and society in Asia: Challenges and next steps”*, Inter-Asia Cultural Studies,* 11, 1, pp. 129–135.
Shannon, C., Weaver, W. (1949), The mathematical theory of communication, Urbana (IL), University of Illinois Press.
Shapiro, J.J., Hughes, S.K. (1996), “Information literacy as a liberal art: Enlightenment proposals for a new curriculum”, *Educom Review,* 31, 2. http://net.educause.edu/apps/er/review/reviewArticles/31231.html.
Shirky, C. (2008), *Here comes everybody: The power of organizing without organizations,* London, Penguin Press.
Shiva, V. (1993), *Monocultures of the mind: Perspectives on biodiversity and biotechnology,* London–Atlantic Highlands (NJ), Zed Books.
Shiva, V. (2013), *Making peace with the earth,* London, Pluto Press.
Siapera, E. (2010), *Cultural diversity and global media: The mediation of difference,* Chichester, West Sussex–Malden (MA), Wiley-Blackwell.
Slack, J.D., Wise, J.M. (2005), *Culture + technology: A primer,* New York, Peter Lang.
Sobrero, A.M. (2009), *Il cristallo e la fiamma: antropologia fra scienza e letteratura,* Rome, Carocci.
Solomon, J. (1993), *Accessing antiquity: The computerization of classical studies,* Tucson, University of Arizona Press.
Solomon, M., Ilika, A. (2007), “Introduction: New Media and Hispanic Studies”, *Hispanic Review,* 75, 4, Autumn, pp. 327–329.
Spence, P., González-Blanco, E. (2014), “A historical perspective on the digital humanities in Spain”, *H-Soz-Kult. Forum: The Status Quo of Digital Humanities in Europe,* October and November. http://geschichte-transnational.clio-online.net/forum/type=diskussionen&id=2449.
Sperber, D. (1996), *Explaining culture: A naturalistic approach,* Oxford–Cambridge (MA), Blackwell.
Stelter, B. (2010), “FCC is set to regulate net access”, *The New York Times,* December 20. http://www.nytimes.com/2010/12/21/business/media/21fcc.html?_r=2&ref=juliusgenachowski.

Stodden, V. (2010), "Open science: Policy implications for the evolving phenomenon of user-led scientific innovation", *Journal of Science Communication,* 9, 1, A05. http://jcom.sissa.it/archive/09/01/Jcom0901%282010%29A05/Jcom0901%282010%29A05.pdf.

Storey, H.W. (2004), "All'interno della poetica grafico-visiva di Petrarca", in Belloni, G., Brugnolo, F., Storey, H.W., Zamponi, S. (eds.), *Rerum Vulgarium Fragmenta. Cod. Vat. Lat. 3195. Commentario all'edizione in fac-simile,* Rome–Padua, Editrice Antenore, pp. 131–171.

Suber, P. (2007), *Open access overview.* http://www.earlham.edu/~peters/fos/overview.htm.

TEI P5 (2008), "Guidelines for electronic text encoding and interchange", in Bournard, L., Bauman, S. (eds.), *TEI Consortium,* Oxford, Providence, Charlottesville, Nancy, 2 vols.

Thibodeau, K. (2012), "Wrestling with shape-shifters: Perspectives on preserving memory in the digital age", in Duranti, L., Shaffer, E. (ed.), *The memory of the world in the digital age: Digitization and preservation, Conference proceedings of the international conference on permanent access to digital documentary heritage,* September 26–28, 2012, Vancouver, UNESCO, pp. 15–23. http://www.ciscra.org/docs/UNESCO_MOW2012_Proceedings_FINAL_ENG_Compressed.pdf.

Thomas, S. (2006), "The end of cyberspace and other surprises", *Convergence,* 12, 4, November, pp. 383–391.

Tomasi, F. (2005), "Il paratesto nei documenti elettronici", in Santoro, M., Tavoni, M.G. (eds.), *I dintorni del testo: approcci alle periferie del libro,* pp. 712–722.

Tomasi, F., Vitali, F. (eds.) (2013), *DH-CASE 2013 – collaborative annotations in shared environments: Metadata, vocabularies and techniques in the digital humanities,* New York, ACM, pp. 1–113.

Turing, A.M. (1937), "On computable numbers with an application to the *Entscheidungsproblem*", *Proceedings of the London Mathematical Society,* 2, 42, pp. 230–265, in Copeland, B.J. (2004), pp. 58–90.

Turing, A.M. (1947/1994), "Lecture to the London Mathematical Society on 20 February 1947", in Copeland, B.J. (2004), pp. 378–394.

Turing, A.M. (1948/2004), "Intelligent Machinery" Report, National Physics Laboratory, 1948, in Meltzer, B., Michie, D., (eds.), *Machine Intelligence,* 5, 1969, pp. 3–23; reprinted in Copeland, B.J. (2004), pp. 410–432.

Turing A.M., (1950/2004), "Computing machinery and intelligence", *MIND,* 59 (1950), pp. 433–460; reprinted in Copeland, B.J. (2004), pp. 441–464. http://mind.oxfordjournals.org/content/LIX/236/433.full.pdf.

Twilley, N. (2015), "Artificial Intelligence goes to the arcade", *The New Yorker,* February 25, http://www.newyorker.com/tech/elements/deepmind-artificial-intelligence-video-games.

Tyler, S.A. (1986), "Post-modern ethnography: From document of the occult to occult document," in Clifford, J., Marcus, G. E. (eds.), *Writing culture: The poetics and politics of ethnography,* Berkeley, University of California Press, pp. 122–140.

Unsworth, J. (2009), "Text-Mining and Humanities Research", *Microsoft Faculty Summit,* July, Redmond (WA).

Uspenskij, B. A.; Ivanov, V. V., Piatigorskij, A., Lotman, J. M. (1973), "Tezisy k semiotičeskomu izučeniju kul'tur (v primenenii k slavjanskim tektstam)", in M. R. Mayenowa (ed.), *Semiotyka i struktura tekstu: Studia święcone VII międz.* Kongresowi Slawistów, Warsaw, pp. 9–32. Eng. trans. "Theses on the Semiotic Studies of Culture (As applied to Slavic Texts)", in Van der Eng, J., Grygar M. (eds.), *Structure of texts and semiotics of culture,* The Hague, Mouton, 1973, pp. 1–28.

Vaidhyanathan, S. (2011), *The googlization of everything (and why we should worry),* Berkeley–Los Angeles, University of California Press.

Valeri, V. (2001), *La scrittura: storia e modelli,* Rome, Carocci.

Vanhoutte, E. (2014), "The Journal is dead, long live The Journal!". http://dsh.oxfordjournals.org/long_live_the_journal.

Van Noorden, R., (2013), "Open Access: The true cost of science publishing", *Nature International Weekly Journal of Scienc*e 485, 7442. http://www.nature.com/news/open-access-the-true-cost-of-science-publishing-1.12676.

Van Orsdel, L.C., Born, K. (2009), "Reality bites: Periodicals price survey", *Library Journal,* April 15. http://www.libraryjournal.com/article/CA6651248.html.

Van Zundert, J. (2012), "If you build it, will we come? Large scale digital infrastructures as a dead end for digital humanities", *Historical Social Research. Controversies around the Digital Humanities,* 37, 3, pp. 165–186.

Veen, J. (2001), *The art and science of Web design,* Indianapolis, New Riders Press.

Vinck, D. (2013), "Las culturas y humanidades digitales como nuevo desafío para el desarrollo de la ciencia y la tecnología en América latina", *Universitas Humanística,* 76, 2013. http://revistas.javeriana.edu.co/index.php/univhumanistica/article/view/5906.

Vise D.A. Malseed M. (2005), *The Google story,* New York, Delacorte Book.

Von Neumann, J. (1948/1963), "General and logical theory of automata", in Hixon Symposium, Taub, A.H. (ed.), *Collected Works,* Vol. 5, Oxford, Pergamon Press, pp. 288–328.

Von Neumann, J. (1958), *The computer and the brain,* New Haven, Yale University Press.

Vygotskij, L.S. (1934), *Myšlenie i reč': Psichologičeskie issledovanja,* Moscow–Leningrad, Gosudarstvennoe Social'no-Ekonomičeskoe Izdatel'stvo.

Vygotskij, L.S. (1978), *Mind in society: The development of higher psychological processes,* Cambridge (MA)–London, Harvard University Press.

Waldrop, M.M. (2001), *The dream machine: J.C.R. Licklider and the revolution that made computing personal,* New York, Penguin.

Wardrip-Fruin, N. (2007), "Playable media and textual instruments", in Gendolla, P., Schäfer, J. (eds.), *The aesthetics of Net literature: Writing, reading and playing in programmable media,* Bielefeld, Transcript Verlag, pp. 211–256. http://www.noahwf.com/texts/nwf-playable.pdf.

Waters, D., Garrett, J.R. (1996), “Preserving Digital Information”, *Report of the Task Force on Archiving of Digital Information.* http://www.clir.org/pubs/reports/pub63watersgarrett.pdf.

Webb, C. (2003), “Guidelines for the preservation of digital heritage”, *National Library of Australia.* http://unesdoc.unesco.org/images/0013/001300/130071e.pdf.

Weinberger, D. (2007), *Everything is miscellaneous: The power of the new digital disorder,* New York, Holt Paperbacks.

Weinrich, H. (2001), *Tempus: Besprochene und erzählte Welt,* Munich, Beck.

Widner, M. (2013), “The digital humanists' (lack of) response to the surveillance state”, August 20. https://people.stanford.edu/widner/comment/183#comment-183.

Wiener, N. (1948/1961), *Cybernetics: Or control and communication in the animal and the machine,* Cambridge (MA), MIT Press.

Wiener, N. (1950/1954), *The human use of human beings,* Boston (MA), Houghton Mifflin.

Wiener, N. (1993), *Invention: The care and feeding of ideas,* Cambridge (MA), MIT Press.

Wieviorka, M. (2013), *L'impératif numérique,* Paris, CNRS.

Wilson, T., Wiebe, J., Hoffmann, P. (2009), “Recognizing contextual polarity: An exploration of features for phrase level sentiment analysis”, *Computational Linguistics,* 35, 3, pp. 399–433.

Witte, J.C., Mannon, S.E. (2010), *The Internet and social inequalities,* New York, Routledge.

Witten, I., Gori, M., Numerico, T. (2007), *Web dragons: Inside the myths of search engine technologies,* New York, Morgan Kaufmann.

Woolf, V. (2005), *A room of one's own* (Ed. by Hussey, M.), New York, Mariner Books.

Yang, J. (2012), “[Focus] Digital humanities research in China”, October 16. http://dh101.ch/2012/10/16/focus-digital-humanities-research-in-china/.

Yen, E., Lin, S.C. (2011), “Uptakes of e-science in Asia”, in Lin, S. C., Yen, E., (eds.), *Data driven e-science,* New York, Springer, pp. 45–50.

Yeo, R. (2007), “Before Memex: Robert Hooke, John Locke, and Vannevar Bush on external memory”, *Science in Context,* 20, 1, pp. 21–47.

Zeldman, J. (2009), “The vanishing personal site”, April 27. http://www.zeldman.com/2008/04/27/content-outsourcing-and-the-disappearing-personal-site/.

Zinna, A. (2004), *Le interfacce degli oggetti di scrittura: teoria del linguaggio e ipertesti,* Rome, Meltemi.

Zorich, D. M. (2008), “A survey of digital humanities centers in the United States”, Washington (DC): Council on Library and Information Resources. http://www.clir.org/pubs/abstract/pub143abst.html.