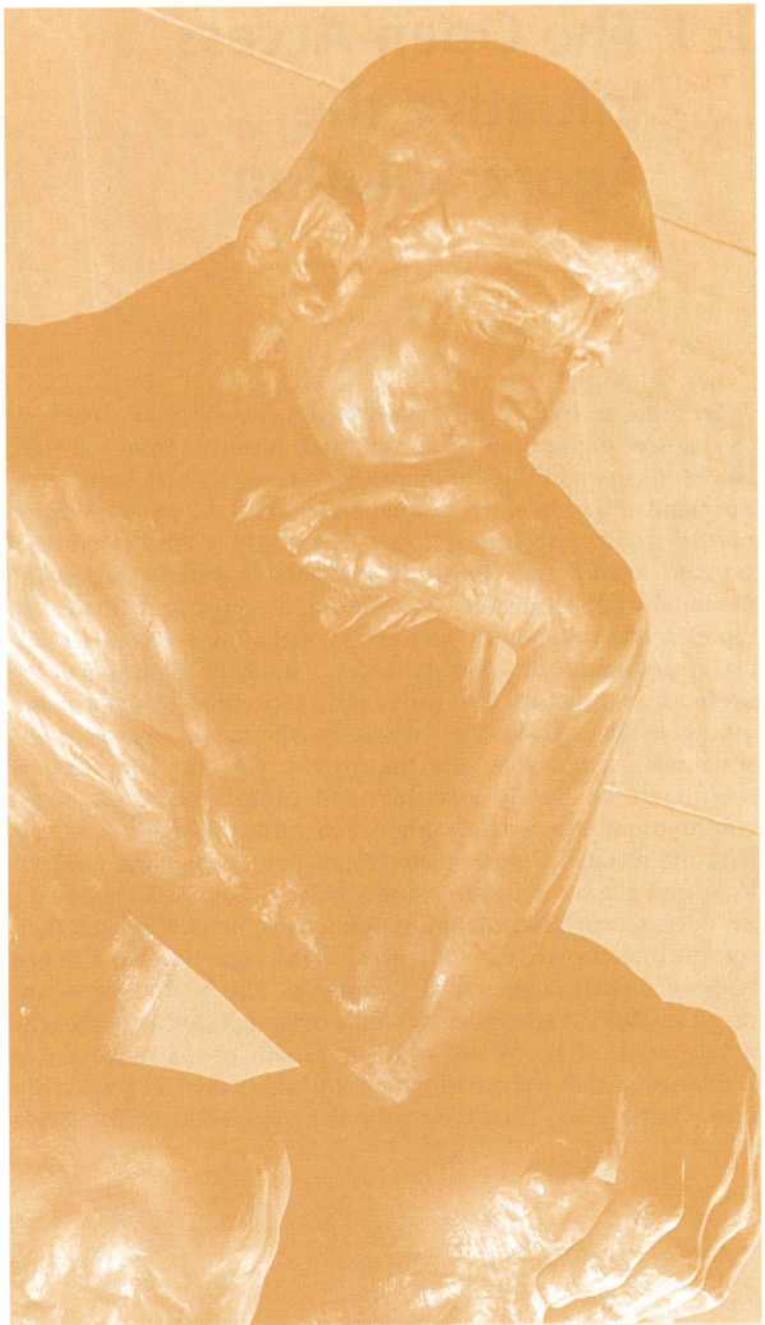


Chapter 2

The Vision of a Web of Culture and Science

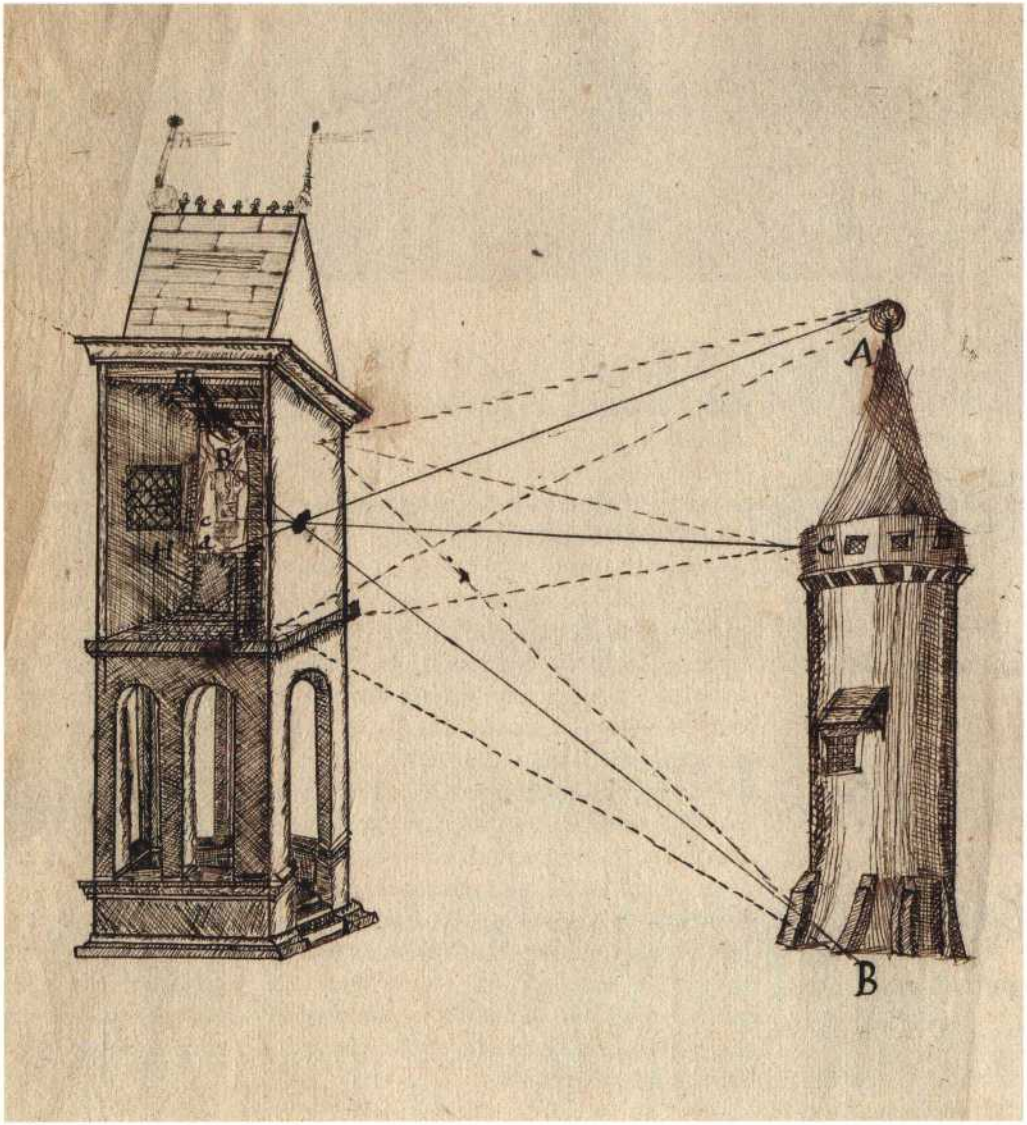
2.1 The Open-Access Initiative the Agora Solution

In order to initiate the far-reaching upheaval that a comprehensive digitization of our cultural, historical, and scientific heritage would amount to, neither missionary zeal, brute force, nor standardization and coordination efforts alone will do. What is needed is rather an infrastructure that makes scientific contributions as rapidly and effectively available as possible, using the potential of the Internet to constitute a global and interactive representation of human knowledge, including cultural heritage and the guarantee of worldwide access. This is the main goal of the open-access initiative. In order to realize its vision of a global and accessible representation of knowledge, the future Web has to be sustainable, interactive, and transparent. Content and software tools must be openly accessible and compatible. Open access contributions must include original scientific research results, raw data and metadata, source materials, digital representations of pictorial and graphical materials and scholarly multimedia material.
An open-access infrastructure should enable users to pursue their specific interests while contributing, at the same time, to a shared body of digitally represented knowledge. This is the key of the Agora solution. It aims at launching a dynamics that combines the development of the whole with the benefit of the individual, a combination that has actually been the hallmark of all great civilisatory enterprises, beginning with the foundation of the Greek

polis which achieved such a synthesis of interests in its agora. A self-accelerating dynamics leading to an ever-more comprehensive electronic representation of culture and science heritage can only emerge if certain minimal conditions are fulfilled. Among them are the requirements of open access, interoperability, modularity, and interactivity. Only if digital sources are made freely available on the Web, only if the same tools can be applied because they share compatible structures, only if diverse digital collections can be integrated to yield an interconnected whole, and only if it is possible to combine the power of computing with the power of the human mind in the analysis of sources, will a set of data turn into a meaningful representation of human knowledge.

It has turned out that even the most convincing standards, models, or tools will remain island solutions as long as those still lacking expertise in electronic information management or access to appropriate equipment are unable to join in. It would be an error to consider the implementation of the agora solution simply as a matter of technological developments which, once completed, have to trickle down from the initiated to the laymen. It makes just as little sense to develop standards without the tools to implement them as it does to develop tools without understanding the questions they should help to answer. The real challenge of the agora solution is thus to achieve an integration of intellectual and technical work and to promote technological developments which are driven by content. Its realization therefore presupposes an environment in which not only technology but also the knowledge about its innovative application to the pressing problems of culture and science is spreading.

In summary, the aim of the agora solution is to establish an open-source culture of the public and scholarly exploitation of cultural and scientific heritage on the Internet, comprising the promotion of content-driven technology in information management. The resulting infrastructure should allow every scientific institution, archive, library, museum, or educational institution to make their resources available online with little effort and in a way that guarantees their interoperability with other representations of human knowledge. In order to make participation in the agora attractive, every potential contributor should gain a surplus value when entering the agora by making contents or tools available on the Web. In particular, all possible meaningful links between a newly available corpus of materials and the already existing ones should be enabled; tools developed for particular aspects of culture or science should be transformed into modules of a universal working environment applicable to all pertinent domains of human knowledge.

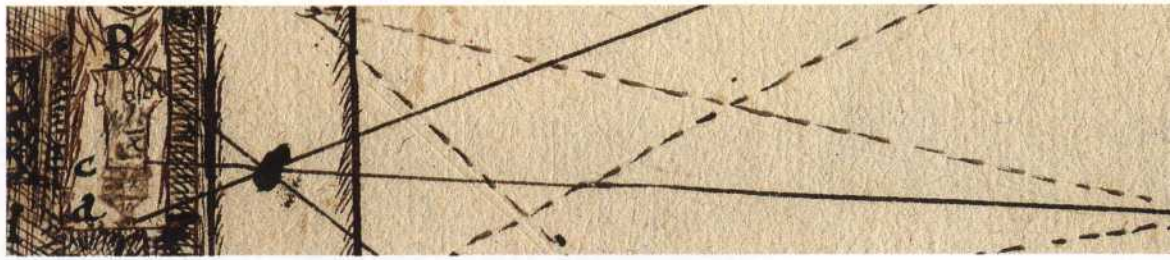


2.2 A Web of Culture and Science

The establishment of an open-access infrastructure must go along with a further transformation of the Web. The Web represents a powerful achievement in the connectivity of human knowledge that until recently seemed inconceivable. The revolution it has caused has rightly been compared with those of the invention of writing and of printing technology. But the rapid development of the Web itself is about to surpass the basic development of the technologies it is based on. In its present form, the Web makes more promises than it can actually keep – at least as long as it is restricted to the specific paradigm that has originally given rise to it.

As was the case when the Internet was created by turning a network of computers into a medium representing a universal hypertext, also its future will rather depend on requirements and possibilities that are revealed only in the context of innovative usage scenarios. Such usage scenarios will emerge when the Internet is used as a virtual public think-tank, a Web of culture and science serving as a medium of reflection on current global challenges of human civilization such as the destruction of ecological equilibria, social impacts of epidemics and drug addictions, or terrorism and other devastating consequences of oppression and increasing mass impoverishment.

If both the natural sciences and the humanities should even under



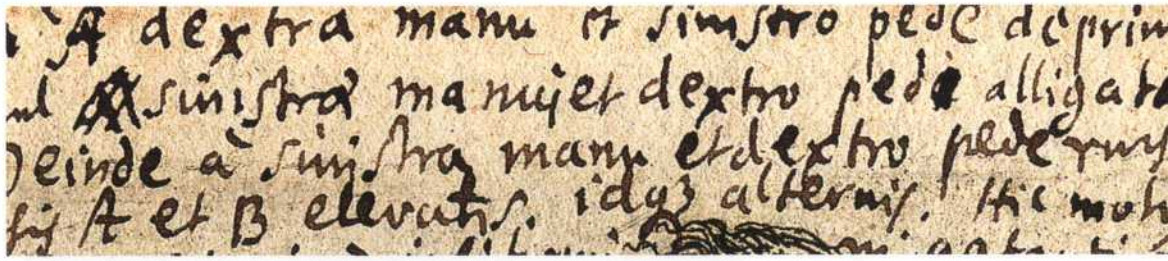
these conditions of global challenges be capable of providing the knowledge crucial for solving the problems of the human species, then this knowledge must also be represented, integrated, and made available in a form allowing for global orientation and action. It is time to take up the opportunity offered by the Internet to create such a medium of global human reflection.

Due to its origin in the idea of hypertext, the World Wide Web is centred on textual data enriched by illustrative insertions of audio-visual materials. The status quo paradigm of the Web is a client-server interaction, that is, a fundamentally asymmetric relationship between providers inserting content into the Web hypertext (server) and users who essentially read texts or provide answers to questions by filling out forms (clients). The hyperlinks of the Web represent structures of meaning that transcend the meaning represented by individual texts, but, at present, these "webized" structures of meaning, lacking any longevity, can only be blindly used e.g. by search engines which at best optimize navigation by taking into account the statistical behaviour of Web users. However, these meaning structures themselves can so far hardly be made the object of interventions by the Web community. There is at present no way to construct complex networks of meaningful relations between Web contents. In fact, the providers have no influence on the links to the contents provided by them and the users have no impact on the available access structures to the content, except by becoming content providers themselves.

This asymmetric client-server relation largely determines the functionalities of the existing web-software. Web servers are not the standard tools of users, while the web-browsers used by them are restricted to accessing existing information in a standard form with only limited possibilities of further processing that information (such as e.g. changing fonts and background colours). As a consequence, the present Web offers no possibility for (radically) different views of the same underlying content, depriving users from the creative potential inherent in the dynamics of the ever-changing Web hypertext.

The Web of the future will thus continue to be essentially based on the representation of meaning by text. However, contrary to the existing web, its emerging paradigm is no longer constituted by the client-server asymmetry but by informed peer-to-peer interactions, that is, by a cooperation of equally competent partners who jointly act as providers and servers at the same time. Future users will work on shared knowledge by constructing new meaning while accessing the existing body of knowledge represented in the Web through meaningful links to texts and text corpora. An important framework for creating such meaningful links can be provided by what is presently discussed as the semantic web, that is, the


A
C
D
B
E
F

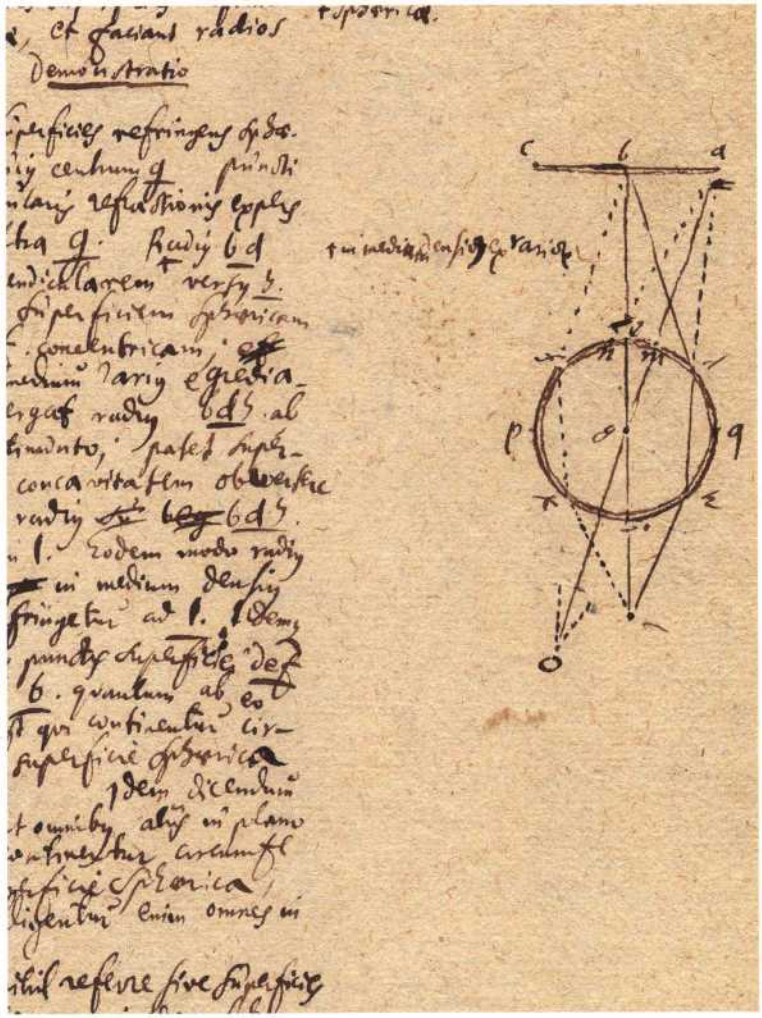


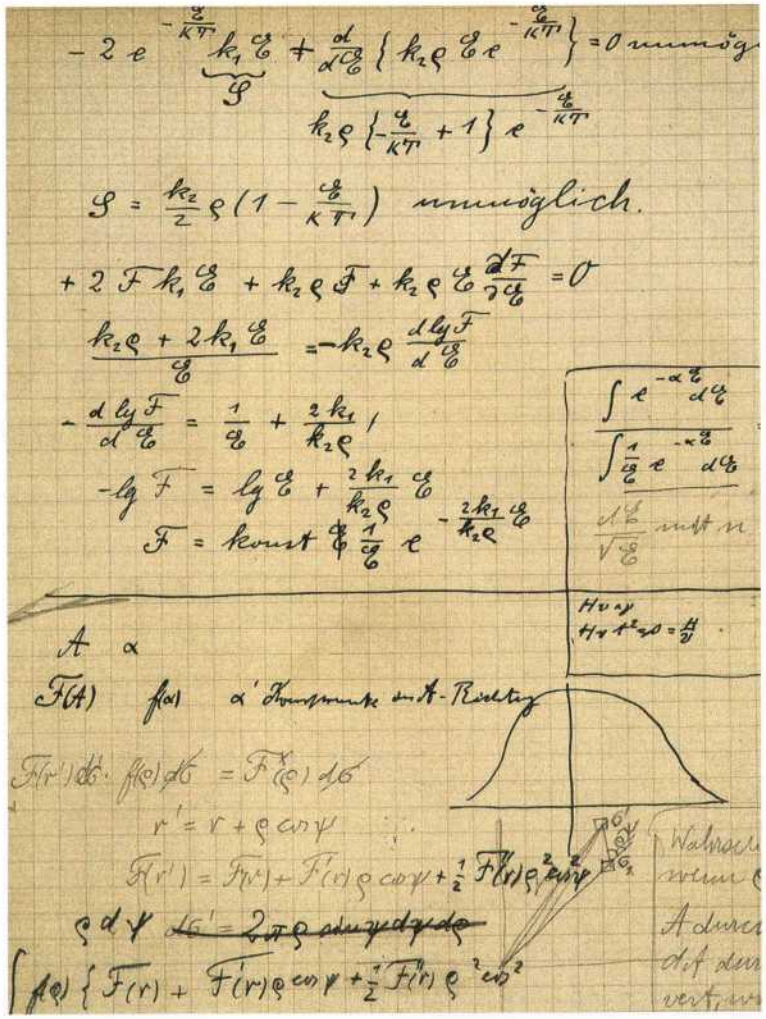
automated creation of links between machine-understandable meta-data. In a further perspective, however, such semantic linking will not be restricted to the use of specifically prepared metadata sets but will exploit the meaning structure of the Web itself in order to provide a content-based semantic access to information.

A basic faculty of human thinking is in fact the ability to reflect on existing knowledge and to produce metadata in a much more general sense than is presently current. The outcome of such reflections typically constitutes a network of meaningful relations. Although the representation of such networks of reflection within the Web does not raise unsolvable technical problems, the very mechanism of such a network of reflection can only to a very limited extent be mapped onto the structures of the Web due to the incompatibility of its traditional paradigm with the creation of meaningful relations between contents. The generation of the specific kind of metadata produced by the sciences and the humanities through analysing, annotating, and reformulating the content of texts is hardly supported by the present infrastructure of the Web. For this reason interactive working environments have to be developed that support the creation of such metadata, and which can then also be used for an improved navigation through scientific contents.

The new infrastructure of the Web which could emerge from a realization of the explanatory power of metadata in a more general sense provides a solution for a serious problem of the current as well as of the future Web, that is to produce order in the ever-growing complexity of the Web by content-sensitive linking. Even the most sophisticated search engines will reach their limits as long as the search criteria can at best exploit the statistics of human decisions about the quality of data. If, however, navigation can be based on content-specific metadata resulting in dynamically changing ontologies, the situation will change. Future "knowledge weaving Web environments," will have powerful link editing functionalities and thus engender a self-organizing mechanism of the Web, improving its hypertext structure.

„If both the natural sciences and the humanities should even under these conditions of global challenges be capable of providing the knowledge crucial for solving the problems of the human species, then this knowledge must also be represented, integrated, and made available in a form allowing for global orientation and action."





„It is time to take up the opportunity offered by the Internet to create such a medium of global human reflection."

„ ... science is above all, communication."

Josiah Williard Gibbs , physicist