

A JOURNAL ON THE WEB: WHAT WE ARE NOT, WHAT WE DO NOT WANT

Alfio Ferrara¹; Massimo Parodi².

¹ Dipartimento di Informatica e Comunicazione, Università degli Studi di Milano
Via Comelico 39, 20135 Milano, Italy
e-mail: ferrara@dico.unimi.it

² Dipartimento di Filosofia, Università degli Studi di Milano
Via Festa del Perdono 7, 20122 Milano, Italy
e-mail: massimo.parodi@unimi.it

*Codesto solo oggi possiamo dirti,
ciò che non siamo, ciò che non vogliamo.*
(Eugenio Montale, *Ossi di seppia*)

Abstract

In this paper we discuss the experience of publishing a journal in both a traditional and an electronic edition. In particular, we take into account our journal *Doctor Virtualis*, devoted to the history of medieval thought, trying to understand what the limitations and problems of electronic publishing are and where the paper edition is really different from the electronic one. Starting from our experience, we then try to better understand what electronic publishing actually is. We critically discuss the usual analogy between traditional publishing and electronic publishing, by proposing a new analogy between the web and the medieval cultural environment. This new analogy helps in understanding some complex processes on the web and in proposing new approaches to transform paper texts into electronic products. To this end, we show how rhetoric plays a crucial role in adapting the text to the medium and how new paradigms for text editing could help in finding a (preliminary) definition of electronic publishing.

Keywords: electronic publishing; text representation; rhetoric; web publishing platforms; electronic journals.

1. Introduction

A fundamental assumption behind the discussions about electronic publishing is that an analogy holds between the activity of producing and representing contents on the web and the traditional publishing activity. This analogy involves the content management workflow, the content representation, the usability of the final product, and the distribution of contents. Electronic publishing products are usually expected to be much easier to use, to reach a wider community of potential readers, and to be characterized by a lower level of meticulousness, which is instead typical of academic and/or scientific products like journals.

According to this assumption, we propose an approach for measuring the similarity between a digital copy of a text and its paper version, and for comparing different kinds of electronic publishing products and general-purpose web applications. The idea is that we can define some classes of properties featuring both the electronic version of a text and the web application that is used to access the text and distribute the content.

From the comparison, it is easy to see how the electronic edition of our online journal *Doctor Virtualis* [1] has the same content of its paper edition, but uses the same medium as many other different web applications. In the middle, between medium and content, there is a third level of analysis, which is relevant to our research. We call this level *rhetoric area*. Our thesis is that this is the field where the content is supposed to be adapted to the medium. Thus, this is also the field where contents can be transformed into electronic contents. In order to better understand this point, we studied the analogy between the classical rhetoric and a possible rhetoric for the web, in order to propose an example on how digital technologies can be used to transform the text in a new way.

Both the analysis of electronic and paper products and the discussion about the representation of the rhetorical features of the contents are based on the analogy between traditional and electronic publishing. But the difficulty in defining the notion of electronic publishing is probably due to the fact that the analogy is wrong. To this end, we propose a different analogy for understanding the relationship between the web and its contents. This new analogy holds between the web and the production and sharing of cultural contents in the Middle Ages. Traditionally, publishing is seen as mediation between authors and readers, but this role is disappearing from the web, where a self-publishing process is more typical. Authors provide their contents on the web directly to the readers. The contents selection depends on the usability of the publishing platforms and the tools used for content retrieval. These kinds of relationships between authors and readers is typical of the medieval period, before the introduction of printing. The lack of mediation between cultural production and its circulation is typical of the web. The web itself is

both the platform for authoring and editing texts and the distribution channel. The fact that there is not a physical dimension on the web makes the distinction between distribution and creation hard to be defined.

In other terms, the web as a whole can be seen as a huge publishing product. According to this approach, the problem is not to find methods for organizing new problems in old categories, or to reproduce the publishing tradition in new forms, but to find a way to deal with the world as it is. This process is related to the success of a paradigm that rules the transformation of knowledge into common sense; the web makes this process faster and causes the difficulties in the paradigm definition. The risk when such a paradigm is missing is well known: the authority, which is no longer a cultural authority, but, in a world of words, becomes inevitably a political authority.

The paper is organized as follows: in Section 2, we discuss the problem of evaluating the quality of electronic publications and we provide two example comparisons among publishing applications of the web and general-purpose web sites. In Section 3, we propose a reference architecture for a publishing platform capable of representing the rhetorical features of texts. In Section 4, we critically discuss the initial analogy between traditional and electronic publishing by proposing a new analogy between electronic publishing and the medieval cultural environment. In Section 5, we discuss the risks in the current electronic publishing scenario. Finally, in Section 6, we give our concluding remarks.

2. The problem of evaluating electronic publishing

One of the main problems in comparing publishing products and web contents is their heterogeneity. In this paper we are interested in studying the evolution of texts from the paper edition to the electronic edition by relying, in particular, on journal articles. At the same time, we also want to understand what the process of producing and editing an electronic text is and what makes a text on the web an *electronic* or *digital* text. This involves the more general problem of how textual contents are communicated on the web, not only when we look at publishing contents, but also when we focus on general-purpose web sites.

In order to compare these heterogeneous phenomena, we propose two different sets of comparison criteria:

- 1) **Analysis of contents transformation.** In this kind of comparison we analyze the transformation of textual contents into web sites by focusing on the final features of electronic texts and on the availability of editorial metadata in the final publishing platform. Moreover, we compare platforms providing paper

texts in some electronic format (e.g., journal web editions, digital libraries) against applications hosting textual contents specifically conceived for or created on the web (e.g., social networks, blogs).

- 2) **Analysis of the electronic product.** In this second comparison, we analyze web sites and electronic publishing platforms in order to analyze both the web site application infrastructure hosting textual contents and the final format of text content. Again, we compare web applications created to support the publishing and/or distribution process against web applications created to support community-based network interactions.

Both the comparisons have been executed by defining two sets of measures A and B. These measure categories (two for the first comparison and two for the second comparison) have then been used in order to evaluate web sites and publishing platforms. We have selected eight web sites hosting textual contents and we have assigned an evaluation in the range between 0 and 1 to each of the measures in the categories A and B for the first comparison, and an evaluation in the range between -1 and 1 for the second comparison [2]. Concerning the web sites to be compared, we selected the following heterogeneous examples:

Doctor Virtualis (<http://riviste.unimi.it/index.php/DoctorVirtualis>). *Doctor Virtualis*, our online journal about the history of medieval thought, is an example of how a traditional academic journal could be published on the web by exploiting the Open Journal System (OJS) platform [3]. This is an application created to support the publishing workflow online from authors' submission to final distribution.

Time Magazine (<http://www.time.com>). The online edition of *Time* magazine has been chosen as an example of web site hosting non-scientific articles for a large public. The platform is a typical content management system that allows the adaptation of contents to the web environment, but also the creation of specific contents on the web.

Blogger (<https://www.blogger.com>). The *blogger* application is a well-known web site for authoring and hosting blogs. It has been chosen in order to take into account a peculiar kind of text content collections, such as blogs, where a multitude of authors can submit or comment textual contents produced by other users. The application supports users in creating texts and easily manages the layout and the organization of their web site.

Wikipedia (<http://www.wikipedia.org>). *Wikipedia* has been chosen as an example application for the collection of user generated contents. Texts are created, edited and checked by the community of *Wikipedia* users. The application is conceived to support user interaction and the so-called collaborative environment.

Google Book (<http://books.google.com>). The Google book application, together with the Google Book Search service, is an interesting example of a potentially new approach to the problem of book search, retrieval and publishing. More specifically, Google Book is conceived to reproduce the distribution phase of the publishing workflow on the web, where users are provided with advanced search functionalities to retrieve books and with a preview of the books' content.

ACM Library (<http://portal.acm.org>). The ACM Library is an example of a digital library hosting scientific literature and provided by the Association for Computer Machinery. The application supports the distribution of PDF versions of the articles and the retrieval and citation of existing literature in the field of computer science.

SpringerLink (<http://www.springerlink.com>). SpringerLink is an interactive database for scientific journals, book series and books, produced by the Springer publishing house. This application has been chosen as another example in the same category of the ACM Library.

Facebook (<http://www.facebook.com>). Facebook is a well-known application for the so-called *social networking*. It is not a publishing platform, but has been chosen as an example of a new and interesting way of collecting user-generated content and the sharing of it among a wide community of users. An interesting feature of the Facebook application is the use of multimedia and multichannel technological solutions to represent contents.

In general, we preferred to propose a comparison among different kinds of publishing technology hosting different kinds of textual contents for a different target, rather than focus our comparison on homogeneous applications hosting the same kind of textual contents. The main reason for this is to stress the fact that heterogeneity of media and contents is a key feature on the web and every kind of platform; even an application like Facebook which is not conceived for publishing could be interesting in order to understand how the circulation of textual information is changing and what texts will be like in the future.

2.1. Analysis of contents transformation

About the transformation of contents, we are interested in understanding whether traditional paper contents are actually transformed into electronic contents or if they are just reproduced on a different *medium*. At the same time, we want to also measure the editorial quality of the final result by taking into account the editorial metadata (e.g., bibliographic references, unique identification) that are actually transformed for the web. These criteria are described by two sets of measures that

are defined as follows:

- A. Content transformation.** The idea behind this first set of measures is that some features of electronic texts can be used as good indicators of the fact that the text itself has been transformed into an example of real web content. In general, low values in the category A mean that a text content has been just repurposed for the web, while high values denote texts either transformed into web contents or created for the web. In the following, we discuss the five measures proposed in this category.
- a. *Number and kind of hyperlinks.* A text on the web is supposed to be a hyper-text [4]. Thus, the number and quality of links connecting a text with other texts are considered as distinguishing features of electronic contents. The higher the number of good links, the higher the level of text transformation.
 - b. *Availability of tags and semantic annotations.* Modern web applications are moving towards the semantic web [5], where contents are annotated with information about the semantics of data. The availability of this kind of annotations (spanning from simple tag or keywords to RDF metadata) is considered a distinguishing feature of electronic texts for the web.
 - c. *Structuring of textual contents.* From a computer science point of view, texts are collections of data. In modern information systems, data are stored in a structured form, by using some kind of database technology (e.g., relational DBMS, XML). Thus, the transformation of a text for the web also means that the text itself is stored in a structured form.
 - d. *Multimedia.* A distinguishing feature of electronic texts with respect to the paper edition is the availability of images, audio and video contents. With this measure, we evaluate the presence of multimedia in the electronic edition of a text.
 - e. *Text format.* From a typographical point of view, paper texts and electronic texts are different. The electronic edition should be re-structured by taking into account the text usability on the web [6][7].
- B. Editorial quality.** The idea behind these measures is that a text on the web should be seen as something that could be searched, retrieved and referenced by other contents or web sites. Then, we try to evaluate the availability online of text metadata together with the quality of the text. Low values in this category denote the fact that the text metadata are missing and/or the text is not produced through a professional editing workflow. High values denote texts professionally edited and/or easy to retrieve and reference. The measures defined are the following.
- a. *Availability of editorial metadata.* Text metadata, such as the contributors,

the title, the publishing house, should be stored and shown online, in order to make them available for searching and citing.

- b. *Content identification.* Any electronic product should be identified uniquely on the web in a persistent way. Web URI [8] is often not enough, since a web resource could be moved to a different location. Recently, some standards for persistent identification of publishing products, such as the DOI [9], have been proposed. The usage of a persistent identification system is then considered as a distinguishing feature for electronic texts.
- c. *Availability of references.* Important information associated with texts is provided by the references. These should be available and reachable on the web by means of links or other retrieval facilities.
- d. *Availability of multiple formats.* If we look at the web as a publishing platform, we should also take into account that a web content should be available in a number of different formats, also according to different devices that could be used to access the content itself.
- e. *Validation and editing of contents.* This measure is introduced to evaluate if a given application provides checked and professional editing of contents. Low values denote self-authored contents and/or a non-professional editing workflow. High values denote applications supporting peer reviews or any other kind of professional workflow for checking and editing the published contents.

The results of our evaluation with respect to the analyzed web sites and according to the measures reported above are shown in Table 1.

A comparative analysis of collected values is shown in Figure 1, where we report the average evaluation collected for each web site in categories A and B, together with the F measure, that is calculated as follows:

$$\text{F measure} = 2 \cdot (A \cdot B) / A + B$$

The F measure, which is calculated as the harmonic mean between values of category A and values of category B gives higher values of web sites with a balanced evaluation for A and B.

The results of Figure 1 show how the analyzed web sites could be divided into three main categories: the first category, including the ACM Library, the Springer Link service, and also Doctor Virtualis, contains applications which mainly provide traditional contents (usually PDF files) in the context of more or less advanced systems for metadata management. In this category the quality of available texts is usually high, but contents are more or less digital copies of the paper original products. The focus of these web applications is on the distribution of textual material rather than on the production of new kinds of products specifically tai-

Web site	A.a	A.b	A.c	A.d	A.e	AVG
Doctor Virtualis	0.2	0.0	0.2	0.0	0.0	0.08
Time Magazine	0.5	0.2	0.4	0.8	0.2	0.42
Blogger	0.8	0.4	0.4	0.6	0.6	0.56
Wikipedia	1.0	0.4	1.0	0.6	0.8	0.76
Google Book	0.5	0.5	0.4	0.2	0.4	0.40
ACM Library	0.5	0.4	0.4	0.0	0.0	0.26
Springer Link	0.2	0.4	0.4	0.0	0.0	0.20
Facebook	0.8	0.6	0.4	1.0	0.6	0.68

Web site	B.a	B.b	B.c	B.d	B.e	AVG
Doctor Virtualis	0.8	0.2	0.0	0.3	1.0	0.46
Time Magazine	0.0	0.0	0.4	0.8	1.0	0.44
Blogger	0.0	0.8	0.6	0.3	0.0	0.34
Wikipedia	0.2	0.8	0.9	0.3	0.2	0.48
Google Book	0.6	0.2	0.4	0.4	0.8	0.48
ACM Library	1.0	1.0	1.0	0.2	1.0	0.84
Springer Link	1.0	1.0	0.7	0.2	1.0	0.78
Facebook	0.0	0.0	0.0	0.0	0.0	0.0

Table 1: Analysis of content transformation.

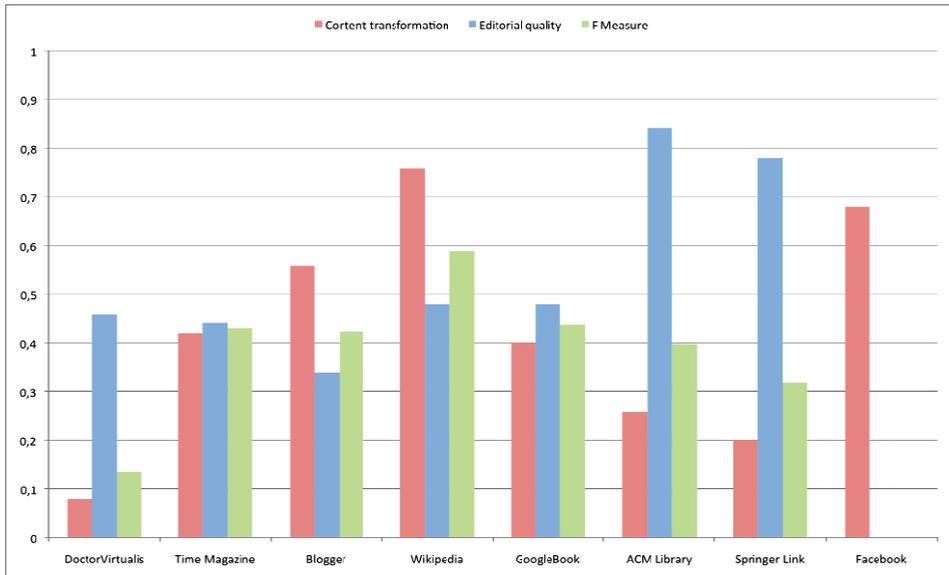


Figure 1: First comparison among web sites hosting textual contents.

lored for the web. The second category, including Time Magazine and Google Book, contains advanced web applications that provide few or less quality metadata but where textual contents are effectively transformed from traditional paper products into real digital contents. In the case of Time Magazine contents are often produced for the web or provided as HTML text. In the case of Google Book, even if the focus is on retrieval and distribution, texts are proposed in a format that can be directly accessed from the web browser. The third category, including Wikipedia, Blogger, and Facebook, represents web applications where the quality of the editorial process is not the priority. Facebook is focused on user interaction rather than on contents. Wikipedia and Blogger represent a relevant phenomenon of content production and diffusion which is however different, and maybe alternative, to the traditional publishing activity.

2.2 Analysis of the electronic products

In the second proposed comparison, we focus on the evaluation of the chosen web sites by taking into account both the characteristics of textual contents and the features of the web application. We adopted the same measures both for category A and B by assigning a value ranging from -1 to +1 to each of them. When applied to A, our measures are evaluated by taking into account only the format of text contents, while when applied to B, the proposed measures are evaluated by taking into account only the functionalities of the web application hosting the textual materials. Measures proposed for the second comparison are the following:

A. & B. Text and application.

- a. *Number and kind of hyperlinks.* As in the first comparison, also in this second one we consider the number and quality of links as a distinguishing feature for the evaluation of both texts and web applications.
- b. *Availability of tags and semantic annotations.* When applied to texts (category A), this measure is based on the number and quality of semantic annotation in the text. When applied to the web application (category B), this measure evaluates if the web infrastructure supports users in providing this kind of annotation for their contents.
- c. *Structuring of textual contents.* This measure evaluates the availability of a database for persistent storage of data (category A) and metadata (category B).
- d. *Availability of multilingual contents.* The web is a world wide application. Publishing content on the web means the opportunity to reach millions of users speaking different languages. This measure is introduced to evaluate if users

speaking different languages can access the web application and the texts. In particular, we take into account the availability of translations for the application commands (e.g., title, menus) and text contents (categories B and A, respectively).

- e. *Web usability*. With this measure, we evaluate the usability of the web site as a whole (category B) and of the text materials in it (category A).

The results obtained for this comparison on the considered web sites are shown in Table 2.

Web site	A.a	A.b	A.c	A.d	A.e	AVG
Doctor Virtualis	-1.0	-1.0	-1.0	-1.0	-0.2	-0.84
Time Magazine	-0.2	-1.0	-0.2	-1.0	0.6	-0.36
Blogger	0.4	-0.2	-0.2	-1.0	0.6	-0.08
Wikipedia	1.0	0.2	1.0	1.0	0.6	0.76
Google Book	-0.6	-0.2	0.6	-1.0	0.0	-0.24
ACM Library	-1.0	-1.0	-1.0	-1.0	-0.2	-0.84
Springer Link	-1.0	-1.0	-1.0	-1.0	-0.2	-0.84
Facebook	0.4	0.6	-0.2	-1.0	1.0	0.16

Web site	B.a	B.b	B.c	B.d	B.e	AVG
Doctor Virtualis	0.2	-0.6	0.2	-1.0	-0.2	-0.28
Time Magazine	0.6	-0.2	0.6	-1.0	0.6	0.12
Blogger	0.6	0.6	0.6	1.0	0.6	0.68
Wikipedia	1.0	0.6	1.0	1.0	0.6	0.80
Google Book	-0.6	-0.2	0.6	-1.0	0.0	0.64
ACM Library	0.2	-0.2	0.6	-1.0	0.2	-0.04
Springer Link	0.2	-0.2	0.6	1.0	0.2	0.36
Facebook	1.0	0.6	0.8	1.0	1.0	0.88

Table 2: Analysis of electronic products.

In order to comment on the obtained results, it can be useful to visualize them by using categories A and B as the axes of a bi-dimensional plot, where we can put markers to denote the web sites, as shown in Figure 2.

Our hypothesis is that category A is a good indicator of the fact that a text has been actually transformed into electronic web content. Thus, low values on the horizontal axis denote traditional paper-style textual contents, while high values denote electronic text editions or materials. Similarly, category B is an indicator of how advanced and usable is the web application hosting texts. Thus, low values

on the vertical axis denote traditional and less usable web applications, while high values denote advanced, usable web applications. According to this interpretation, we can see four areas in the plot. On the top-right corner, we have advanced web application hosting electronic texts that are texts that have been actually transformed or created for the web. On the top-left corner, we have advanced web applications hosting traditional texts that are texts available in digital formats but not really transformed for the web environment. On the bottom of the plot we have traditional or less usable web applications; among them, on the left we have web sites hosting traditional texts, while on the right we have texts transformed for the web.

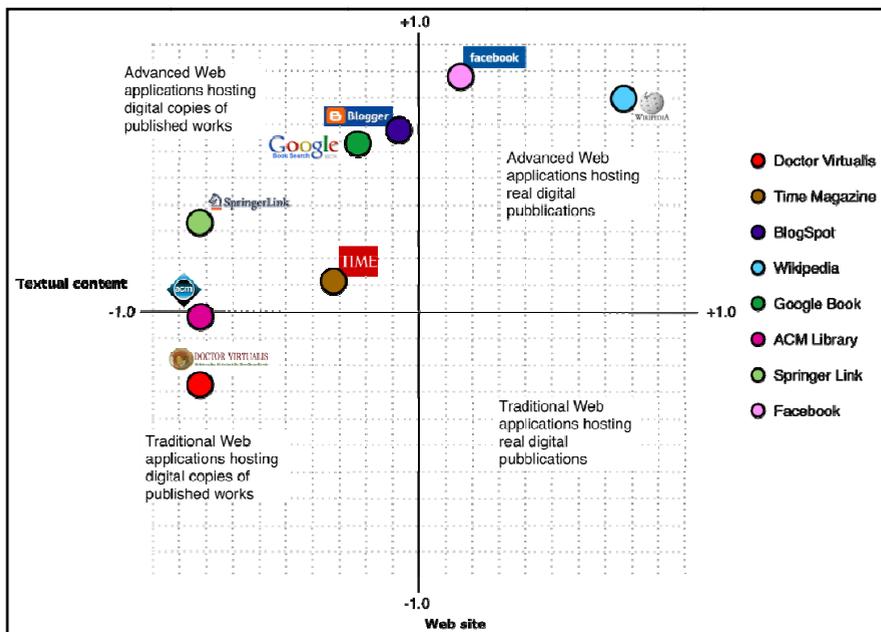


Figure 2: Second comparison among web sites hosting textual contents.

As expected, we do not have examples of traditional web sites hosting well-transformed texts. Doctor Virtualis is collocated in the category of traditional web applications hosting paper-style texts. The reason is mainly that text contents are not transformed but simply provided as PDF files. The OJS web application then is a traditional content management system and it is configured in a way that does not provide semantic annotations for contents. An interesting result is that the main part of analyzed web applications seems to be advanced if we consider the application itself, but not really new if we consider the kind of textual contents available. For example, Google Book does provide advanced search functionalities

and navigation facilities, but texts are only partially transformed for the web, in that they look like PDF-based previews of the paper edition of the text. Again, as in the previous comparison, Wikipedia seems to be something different from electronic publishing platforms like OJS or digital libraries like ACM Library or Springer Link, in that not only the web site is quite advanced but also the texts hosted are conceived for the web and produced on the web. Moreover, they are accessible in many different languages and they are structured and searchable easily.

2.3. A preliminary comment

The goal of our comparisons is not only to propose a methodology for evaluating the transformation of texts when published on the web, but also to suggest new problems in the discussion about electronic publishing. According to our results, the success of some web applications, such as Wikipedia or the several platforms for blogging, seems to be due to the capability of providing advanced facilities for the retrieval and circulation of texts, together with a new way of conceiving editorial activity and text transformation. Tagging and semantic annotations, availability of multiple formats, community-based editing and authoring, seem to be the keys of the success on the web.

If our hypothesis is true, the web should not only be seen as a new and powerful way of distributing contents, but should also be seen as a phenomenon that changes the nature itself of textual communication. Publishing a text on the web is more than producing a digital format of the text and making it available in a nice and advanced web site; the text itself needs to be changed. Technology provides tools to represent and visualize the different layers of meaning in the text; authors and readers can interact in a new way; the text becomes a resource that users need to access, share and reuse. But, is Wikipedia or Facebook really the model? These two applications have a huge success, more or less ten times the number of visitors of a successful service like Google Book, but do they provide editorial contents? Facebook is a social networking platform. It is an example of how communication on the web can change, but it is not of course a publishing platform. Wikipedia is a collection of texts produced by a community of users where the process of creation, edition and distribution of contents has nothing to do with the traditional publishing workflow. At the same time, we should remember that publishing is not only business; historically, the publishing industry has been a powerful device for producing and distributing contents to wide communities of readers. Today millions of users share ideas and contents using Facebook, Wikipedia or blogs. If this is not the same thing as reading a publishing product, it is at least a phenomenon that cannot be ignored by the publishing actors.

In this context, our hypothesis is that this is due not only to the business model of these kinds of applications, but also to the kind of products and technology used and also to a different view of what communication and distribution of content is. The rest of the paper tries to introduce a discussion about these transformations and the opportunities and the risks they imply.

3. The rhetoric area

The comparison among web products is usually based on two main criteria: on one side we have the evaluation of the contents and on the other side, the contents as the publishing process transforms them, including the final web application used for content distribution. Looking at these criteria in the case of Doctor Virtualis, it is easy to see how the digital content is almost identical to the paper edition of the journal, while the web application is more or less similar to other web applications like Wikipedia.

In this section we try to show that it is possible to define an intermediate area between the contents and their final distribution format, namely the *rhetoric area*. The definition of the rhetorical features of a text is seen as a fundamental phase of the process of transformation of a text into its final form. In fact, rhetoric is strictly connected to the problem of adapting a text to a medium, which in our case is the transformation of the text into its electronic edition.

In order to better understand this point, we studied the similarity between the classical rhetoric and a possible rhetoric for the web. The western rhetorical tradition is focused on three main features, called *dispositio*, *elocutio*, and *memoria*, which refer respectively to: a) the order given to argumentation and contents; b) the way the different argumentation steps are used for presenting with respect to the goals of the argumentation; c) the set of possible contents from which the ones that are expressed are chosen. Starting from these elements we define a model for a publishing application capable of highlighting the rhetorical shape of contents on the web in order to suggest an effective use of the rhetorical features in electronic publishing. The model is required to support the authors in defining the rhetorical structure of their contents and to make the rhetoric effective on the web. This means, for example, that if a metaphor is a distinguishing rhetorical feature of a text, the correct representation of the same metaphor in a web edition of the text may be realized by means of a link or some other kind of device, which makes such a relation active and usable on the web. But such a link should not be realized according to the hyper textual approach, in which we have arbitrary and free associations among contents, but should be defined according to the definition of metaphor;

that is, in the rhetorical tradition calling something with the name of something else. When the use of analogies or metaphors is a strategy of argumentation, the electronic devices should be used to reconstruct the heuristic function of the rhetorical features by explicitly showing the internal mechanism of the text to the reader. As an example, we consider the following analogy:

Water is to liquid as ice is to solid.

When this text is transformed into an electronic object, the tagging mechanism, usually used just to give the text a layout format, should instead be also used to make the internal relations explicit. Thus, we expect to tag terms according to their function in the sentence and to have a database storing the internal text relations according to an ontology of rhetorical features [10]. An example could be:

```
<S id="S1">
  <S id="S2">
    <T id="T1">Water</T>
    <R id="R1">is to</R>
    <T id="T2">liquid</T>
  </S>
  <R id="R2">as</R>
  <S id="S3">
    <T id="T3">ice</T>
    <R id="R1">is to</R>
    <T id="T4">solid</T>
  </S>
</S>
```

where the database contains:

```
analogy(S2, S3)
corresponds(T1, T3)
corresponds(T2, T4).
```

In our trivial example, we just show how tagging the text components and storing the internal rhetorical shape of the text can enrich the information provided to the final reader. In such a way a user could access analogies in the text by using a search function and the internal structure of the analogy is made explicit.

A reference architecture of an electronic publishing platform for rhetoric is shown in Figure 3.

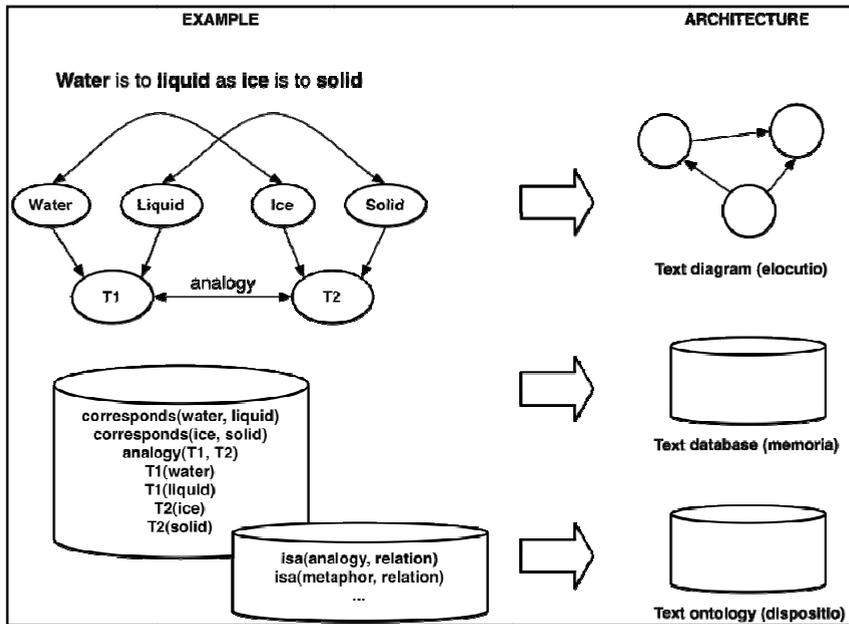


Figure 3: Reference architecture for rhetoric representation

According to Figure 3, the *dispositio* is represented as an ontology of the published text, where contents are connected by specific rhetorical relations such as dependency, semantic translation, analogy. This solution enables new search functionalities and provides a sort of electronic language for the text.

The *elocutio* is represented as a diagram depicting the structure of the argumentation, according to the classification proposed by Perelman of the different types of roles played by contents [11].

Finally, the *memoria* is seen as a database, representing the alternative argumentations that are possible starting from a point of the text. The main goal of this approach is to show how the web can be used in order to publish the contents in a specific and new way, instead of reproducing the traditional paper format in the electronic edition.

4. The analogy revised and the medieval cultural environment

Both the discussion about the comparison of electronic publishing product and the discussion about the rhetorical features of texts, especially for texts in the humanities, are based on the analogy between electronic publishing and traditional publishing. According to a classical definition of analogy, seen as a similarity of rela-

tions, such as in the case of mathematical proportions, the structure of the traditional workflow and activity organization of paper publishing is intended to be similar to the structure and organization of electronic publishing.

At the same time we realized how difficult it could be to give a precise definition of electronic publishing by focusing on its peculiar characteristics and by avoiding the well-known discussions about the problem of the technical supports and the accessibility of contents that seem to be only superficial characteristics of electronic publishing. Looking at this problem, we claim that the difficulties in understanding what electronic publishing is, are due to the fact that the analogy between electronic and traditional publishing does not work; it seems to be that the similarity between the two activities could not help in understanding electronic publishing because the analogy between the two is wrong and misleading.

Thus, we try to change our approach by introducing a different analogy to explain the relationships between the web and the processes of cultural content production. To this end, we will use the characteristics of these processes in the Middle Ages, which is the historical period that is studied and discussed in our journal. The medieval period was an age in which the vitality of the culture, the heated debates – sometimes too heated – about philosophy and theology, the great production of written works as well as the success of Universities were not supported neither by the printing nor by the publishing industry.

A first reason why the analogy between Middle Ages and the web seems to be interesting is that the role of mediation between authors and readers played by the publishing industry is typical of the “Printing Age”. Only after the invention of printing did the publishers acquire the power of selecting contents and of defining the modalities of their distribution, which is one of the most important and crucial phases of the publishing process, both for authors and for publishers. This typical role of mediation is no longer crucial on the web, where the different phases of the publishing process are often reduced to a unique phase, the self-publishing one. As it was before the printing invention, the author provides his/her contents directly to the reader by also skipping that phase typical of the Middle Ages, represented by the work of copyists. Also the selection of contents can be reduced to the work of organizing texts according to the technological platform requirements and to the tools available for content and information retrieval.

The relationship between reader and author is then similar to the relationship between reader and author in the Middle Ages; the same can be said also about the notion of author. Ideas and contents on the web seem to be much more important than the name of their author, which is relevant only when it is far enough in time to be considered as part of the so-called tradition. The same was true for the medieval notion of *auctoritas*: “Someone says that...”, but Augustine would answer that...”; “Someone claims that this thesis is based on the ideas of Aristotle”. In

these examples, the names of contemporary authors vanish in the fog of the present moment, while the great names emerge from the light of the past.

The idea of texts as something which is closed, neither usable nor modifiable, and protected by some kind of copyright is vanishing as well. But also the idea of copyright is something that has been adopted in a very limited period of the thousand-year-old history of human culture and was completely unknown in the Middle Ages.

Sometimes those publishers who desperately try to find a place in the world of electronic publishing look like a medieval copyist who continued to copy pages and pages of the Bible while the Gutenberg machine produced an infinite number of copies of the same text.

On one side, the analogy with the medieval world can help in understanding how the relation among publishers, authors, and readers is changing; on the other side, it can also help in understanding why the analogy between traditional publishing and web publishing is wrong. With respect to the problems we are dealing with, it is more useful to look at the web as a huge editorial product. Thus, the analogy should be found not between traditional and electronic publishing, but more bravely, between the electronic publishing and the world itself. Again, we can use the Middle Ages: in that period many authors described the world as the word of God, a world where the Creator talks to the creatures and God's thinking becomes a speech: God speaks and by speaking He creates the reality in a sort of fantastic early uploading process.

Out of these considerations the crucial role of rhetoric is confirmed with its components as they were presented in the previous sections. The similarity between the memory of the orators is evident from Quintilian onwards and the huge database of contents provided by the users which constitutes the memory of the web; or at least the similarity between the memory and those limited databases that constitute the contents of applications like Wikipedia, where each author can use, cite, elaborate every content, such as orators with argumentations taken from the tradition. This happens because in such a world what is important are the ideas, the words that become reality, rather than the author or the publisher who provided them.

5. Rules for a world of words

As soon as we accept the new approach, by changing the analogy used to discuss electronic publishing the problem itself changes: we do not have to find a way for organizing and repurposing old problems by adopting new categories, not to rec-

reate the publishing tradition with new devices or technologies. The problem now is to deal with the world as it is and not the world as we think it should be. If we do not change the paradigm (or the analogy) that we use to understand the new phenomena on the web, we will not be able to see what the world actually is.

If we do not accept the world, in our case the web, as it is, we have to face two kinds of risks, which are the usual risks that have to be faced by those who want to understand the reality using old categories. On one side, the risk is to create new mechanisms for controlling or censoring culture – we can also call them “peer review” or “impact factor” –, new ways for simplifying the world and for trying to throw contents out of the web. On the other side, the risk is to leave the production of contents on the web completely out of control, which is the ideal situation in which the rich – in terms of technological means, knowledge, and/or money – prevail on the poor.

In our opinion the solution, which means a long process of research and trials, is to accept the new and actual processes of production and circulation of culture in order to provide theories and tools to govern them. This means that we should start to accept the fact that lower-level quality of contents could be functional to a wider access to cultural products. However, at the same time, this also means investing research efforts, time, and money to study new ways of representing contents in the digital environment in order to distribute contents on the web effectively, instead of repurposing old traditions and methodologies.

Looking at human history we see how even the revolutions were powerful enough to accelerate the definition of new cultural paradigms which transform culture into common sense. But nowadays, it is the specific function of the web that makes the transformation process faster. The consequence of this is that the definition of new paradigms (and not only new methods or tools) becomes more and more difficult. In this context, working with wrong analogies is risky: without a clear view of the new communication processes, the request for rules and regulations increases. In the world of words that we have to face, less ambitious but more effective, authorities based on strength can easily replace an ineffective cultural authority.

6. Concluding remarks

In this paper, starting from the experience of Doctor Virtualis, we have tried to show how the state of electronic publishing initiatives is often far away from what the web is or is going to be. The gap between new forms of the circulation of ideas, such as Wikipedia or blogs, and traditional or new initiatives in the publishing

field is still relevant. Doctor Virtualis is much more a traditional paper journal on the web than a real electronic product. This is the reason for our reference to Montale's famous verse: the electronic journal is what we are not. But we do not want to transform Doctor Virtualis into a thematic version of Wikipedia or into a personal blog, because we do not think that this is the way ideas should be communicated and discussed. At the same time, we do not think that some new forms of publishing, either commercial Google-style ones or open OJS-style ones, will be the answer (or at least, the answer for us). The reason is that we do not think that peer-reviews, impact factor(s) and in general regulations, are the solution for achieving high-quality electronic publications. These kinds of solutions should be redefined for at least two reasons: they will lose the competition against cooperative platforms for content production because these platforms provide fast and free access to contents. But it is not just a matter of business: the second reason is that we think that the web requires new paradigms for content production and distribution. The focus on contents more than on authors; the availability of contents in a multi-format and multi-channel environment; the need for reusing and sharing contents: these and others are the characteristics of the web itself. Wikipedia is not a successful application on the web; it is the web that is a successful application and Wikipedia is only an example of how we can exploit the web for re-thinking publishing rather than just repurposing it.

Acknowledgments

Some of the main ideas in this paper come from the research activity of Dr. Chiara Selogna, who worked on the theme "Rethoric in humanistic contents communication on the web". The work of Dr. Selogna was partially funded by the project "Dote Ricercatori" of the Lombardia Region. The project work from June to September 2000 was performed in collaboration with Università degli Studi di Milano (tutor Prof. Massimo Parodi) and with the LED publishing house of Milan.

Notes and References

- [1] Doctor Virtualis is available online at <http://riviste.unimi.it/index.php/DoctorVirtualis> (April 2009).
- [2] The goal of our evaluation is only to provide an example of how the proposed evaluation methodology can be applied to real web sites. A real and extensive evaluation should be executed by adopting statistical methods of organizing data and collecting results.
- [3] KOPAK, R. Open Access and the Open Journal Systems: Making Sense All Over. *School Libraries Worldwide*, 14(2), 2008, p.45-54.

- [4] LANDOW, GP. *Hypertext 3.0: Critical Theory and New Media in an Era of Globalization*. Baltimore: The Johns Hopkins University Press, 2006.
- [5] BERNERS-LEE, T. et al. The Semantic Web. *Scientific American*, May 2001.
- [6] KRUG, S. *Don't Make Me Think: A Common Sense Approach to Web Usability*. Berkeley: New Riders Press, 2000.
- [7] NIELSEN, J. and LORANGER, H. *Web Usability*. München: Addison-Wesley, 2006.
- [8] BERNERS-LEE, T. Universal Resource Identifiers in WWW. CERN Report, June 1994. Available at <http://tools.ietf.org/html/rfc1630> (April 2009).
- [9] PASKIN, N. Digital Object Identifiers for scientific data. In BECKER, E; et al. *Digital Rights Management - Technological, Economic, Legal and Political Aspects*. Lecture Notes in Computer Science 2770 Springer, 2003, p. 12-20.
- [10] CAMARA PEREIRA, F. *Creativity and artificial intelligence : a conceptual blending approach*. Berlin ; New York : Mouton de Gruyter, 2007.
- [11] PERELMAN, C. and OLBRECHTS-TYTECA, L. *The new rhetoric: A treatise on argumentation*. Notre Dame: University of Notre Dame Press, 1969.

June 2009
Printed on demand
by "*Nuova Cultura*"
www.nuovacultura.it

Book orders: ordini@nuovacultura.it