

## A DISCIPLINE OF REFERENCE IN PRESENT DAY RESEARCH: DIGITAL HUMANITIES

**Corina Moldovan, Assist. Prof., PhD, and Gabriela Rus, PhD, Researcher, "Babeş-Bolyai" University of Cluj-Napoca**

*Abstract: The impact of the digital era on the humanities is evident. Beginning with the electronic text and the online libraries, with the search platforms and the creation of databases increasingly performant, the two fields, the „digital” and the „humanist” have merged and created a new interdisciplinary field, called the Digital Humanities. In this paper we will analyse the main aspects of the history of the relationship between the computer and the text, the main applications of the new discipline, the textual and the visual reading, the digital cartography or the 3D reconstruction. We take into question the pedagogy of the Digital Humanities as more and more DH Centers are created around the world. Finally we will investigate the epistemologic value of the new science and its capacity to create new research patterns and new instruments of knowledge of the modern world.*

*Keywords: Digital Humanities, digital era, the post-print, modern humanist research.*

“An incredible number of dice, always rolling, dominate and determine each individual existence: uncertainty, then, in the realm of individual history; but in that of collective history . . . simplicity and consistency. History is indeed ‘a poor little conjectural science’ when it selects individuals as its objects . . . but much more rational in its procedures and results, when it examines groups and repetitions.” (Fernand Braudel)

The complexity of the research method and discipline, represented by *Digital Humanities* can be very good explain through the Fernand Braudel’s quote mentioned above, related to the way of understand history. When we are dealing with a large field, the best method to deal with it isn’t that of stitching together different parts of knowledge about different cases and situations, because it doesn’t represent a sum of those cases, so, in order to see it in perspective we have to create a collective system, that should be apprehended as a whole. This pattern must be applied when we want to better understand the problem of *Digital Humanities*.

A qualitative rather than quantitative approach of the *Humanities* disciplines focused on the literature, music, arts, dance, theater, architecture, philosophy etc. can be made using computing like a modeling tool, like a practice of representation, like a way of thinking, and like a way of ontological responsibility. *Digital Humanities* describe the work at the intersection between digital technology and humanities disciplines. Here is a more elaborate definition:

“The digital humanities, also known as humanities computing, are a field of study, research, teaching, and invention concerned with the intersection of computing and the disciplines of the humanities. It is methodological by nature and interdisciplinary in scope. It involves investigation, analysis, synthesis and presentation of information in electronic form. It studies how these media affect the disciplines in which they are used, and what these

disciplines have to contribute to our knowledge of computing.” (The European Association of Digital Humanities)

### **Elements of the *Digital Humanities* history**

The milestone of the interdisciplinary experiment between humanities and the computer was set in the early years after the second world war, when an Italian Jesuit priest, Roberto Busa, started a monumental task that consisted in creating an *index verborum* of all the words in St. Thomas d’Aquino writings, in total some eleven million words in Latin. Busa got an unexpected but precious help from an IBM director, in the USA, and so a very laborious, both physical and intellectual work with punched cards and huge computers, following a concordance program written by Busa for the project began. The originality of Busa’s enterprise was the fact that he refused a purely mechanized program. Organizing the entries into “lemmas”, he succeeded in putting them into an order that made sense, like in a dictionary. His project lasted about 30 years and eventually produced in the 1970s the 56 printed volumes of the *Index Thomisticus*, currently to be found on the internet, where the search of one entry, or lemma, provides concordances (the occurrences of the given terms or expressions in their context), and also offers statistics information on them.

The advent of the computer made possible the record of the frequency of a word appearance in a text and many such exercises were made, by pure curiosity or even for legal purpose. The period also saw the first attempt in bringing together humanists and those involved in computer issues; in 1964 IBM organized a first conference entitled “Literary Data Processing Conference Proceedings”, discussing the difficulties that the text encoding procedures had at that time. This was the first of a series of conferences that preceded the organization of the *Association for Literary and Linguistic Computing*, born at Cambridge in 1970. In the same time, a journal was created, the *Computers and the Humanities*, (1966) and also some centers were created, dedicated to the use of computer in the humanities were established, mostly in the UK but also in Italy or Germany and of course in the USA.

If the first period of what was not yet a new field is to be described like technologic, the second phase is that of the knowledge dissemination. Use of technologies was more and more and rapidly growing, knowledge was spread. Conferences on the subject were organized every year, with now an emphasis on the linguistics, but most of all it provided a forum for researchers in a field that now had a name, the corpus linguists.

Academic computing centers were organized and the first platform where a researcher had the opportunity to register his paper was the “Oxford Text Archive”, in 1976, that maintained an author’s electronic text, provided its copyright permission: it was the first digital library in the world. In the same time a group of scholars began to create archives of texts, such as *Thesaurus Linguae Graecae*, in California, or *The Old English Corpus for the Dictionary of Old English*, in England.

Meantime, the period saw the introduction of academic courses on various aspects of humanities computing, emphasizing on stylistic analyses or linguistic applications, in the humanities field, or on the mechanics of using specific software programs, in the technical universities. This second period lead to the recognition of the need for standard tools for text encoding, one the initiative being the COCOA concordance program in Britain, the predecessor of the OTA (Oxford Text Archive).

In the third phase, this is mid-80s to early 90s, significant developments were registered. As personal computers became a necessity of scholarly life, the access to the three existent DOS-based text analysis programs became more facile, and so the creation of the hypertext (the link between file cards) was rapid which came along with the invention of the networking, or at least the electronic mail, by the mid 80s. The first electronic seminar was launched by Professor Willard McCarty, in the name of a group of people working in the field and who thought that they had to get in touch with each other on regular basis. The first message was sent out on May 7 1987 and since the first electronic or discussion list or group became a model with excellent standards of editing and with high level of discussion that is still functioning today, making a significant contribution to the development of the digital humanities community. (*The Humanist Discussion Group*).

This period also saw the publication in print of large-scale attempt of bibliography, in the *Humanities Computing Yearbook*. Also, in 2001, as a result of a massive effort and million dollars support in North America, a new mark up scheme for text encoding was created, the TEI (Text Encoding Initiative), and it was the first systematic attempt to categorize and define all the features, in the so called tags, in the humanities. The TEI adventure began with an international conference of the ALLC where scholars examined the possibility of creating a standard encoding scheme for electronic texts. In all, some 400 encoding tags were specified and then published in the TEI guide-line. This was the beginning of the digital libraries in the western world so common nowadays.

The digital humanities activities developed rapidly as the Internet era appeared. The first graphical browser called Mosaic was launched in 1993. Academic groups already involved in linking computer and humanities rapidly saw the great opportunity of the possibility of sharing knowledge and publishing on the web. Electronic resources were gathered in “archives”, such as the William Black archive created by the Institute of Advanced Technology of the University of Virginia, archive meaning a collection of material where the reader will follow the road of his own search and editing meaning the publishing of an e-text enriched with the comments of a group of researchers. One of the first digital humanities experiment was, and still is, the Woman Writers Project, a web site created at the Brown University that offers to the scholar the possibility to read texts written by pre-Victorian women writers, project supported by the National Endowment for Humanities.

A revolutionary dimension was added to the humanist electronic resources with the rise of the multimedia information. Suddenly a new language emerged, an audio-video expandable language that changed the classic or normal text into what it is called today a “post-text” which means transmedia crisscrossing and the creation of a new and very interesting concept of a kind of sensorium of humanist knowledge focused on transferable and interactive research platforms. Many examples can be given from the *Valley of the Shadow* from the Virginia University to *The Republic of Letters*, from the Stanford University or the biggest European platform that gathers million of pieces of information of the European heritage called *Europeana.eu*.

In 1990s was a great deal of concern about the death of the book, and due to this fact, scientists have been written and published a lot of articles and books on this problem. The digital age could kill off the print resources, but few years after it was demonstrated a

complex relationship between print works and electronic resources, and print resources continued to co-exist with digital ones.

The humanities scholar had as main activity reading and analyzing texts and computational techniques simplify those processes by creating working tools.

### **Applications of *Digital Humanities*:**

#### **Reading texts**

The “reading” refers at processing content and is a form of data mining that allows a text to be processed and analyzed. Promoters of the method argued that this kind of text processing present aspects of literary and historical texts at a large scale that is not possible for human readers. Those results are very useful tools in starting new researches.

If we refer to the stylistic analysis studies, the Computational Stylistics is focused on searching patterns in language related to the writing and reading processes that could not be made without computational methods. For example, is very easy to analyze the patterns of similarities and differences in Shakespeare plays in terms of spoken dialogue. Here is a grouping in these plays, of genre and chronological sections. Computational stylistic work can never be a fix one because of the variability of the semantic operation of language that means the same textual form in different context, has different meaning.

Authorship studies determine ‘yes or no’ aspirations to the situations that are formed. Authorship attribution is based on Stylistics that are been accepted in a specific contexts and areas. Humanities computing deal with attribution based on stylistic features specifically based on internal verification and comparing the results.

A fundamental tool for any kind of analysis and research on language is represented by a corpus that in digital era has been taken an electronic form. Corpora of literary works were assembled to facilitate stylistic analyses and authorship studies and it represent the general language. Since 1990 began the ‘golden era’ of linguistic corpora; a different corpora are compiled most of them by government-funded projects in Europe, USA and Japan as: monolingual corpora and multilingual corpora that covered multiple languages. Those numerous corpora are available in the USA through the Linguistic Data Consortium (LDC) and in Europe through the European Language Resources Association (ELRA).

The most used representation format for linguistic corpora is XML. Some of corpora are tagged using the EAGLES XML Corpus Encoding Standard (XCES), and a Text Encoding Initiative (TEI). There are two types of information that may be encoded with XCES:

1. *Gross structure*: universal text elements down to the level of the paragraph as volume, chapter, footnotes, titles, headings, tables, figures, etc., features of typography and layout, and non-textual information like graphics.
2. *Segmental structure*: elements appearing at the sub-paragraph level as quotations, orthographic sentences, orthographic words, abbreviations, names, dates and highlighted words.

As regarding the annotations, are linked with data is using XML conventions. Corpus annotation is grouped in some major section type: morpho-syntactic annotation, parallel alignment, syntactic annotation, semantic annotation, discourse-level annotation and annotation of speech and spoken data.

Reading texts in digital humanities language go from the pure text approach such as stylistic analysis and authorship studies, analysis of linguistic corpora, the electronic editing, textual analysis, thematic research collection, print scholarship and digital resources and also interdisciplinary applications such as cognitive stylistics, multivariant narratives or robotic poetics.

### **Examples of text reading:**

#### Electronic editing

Our century challenges the editing process with new demands. New „editing” genres were born, such as post-print artefacts, products where content and structure are disputed between strategists and information designers or graphic artists. Our post-print era lies on the vision of assuring effective and sensitive communication to a diverse, sociologically and culturally, audience. An interesting transformation involves the „old” procedure of marking up of a text, that meant the conventional symbolic printer’s instructions in the margins of a text. This job was done by the „markup men” and applied by the editors, publishers and authors and it was indeed the genesis of today’s markup languages, such as XML or HTML, the language that enables the display of text and graphics on the web or the CSS (Cascade Style Sheet) that is the computer language that expresses the presentation of structured documents whose content can be reused and presented in many ways (color, font, layout). One of the successes of the electronic editing is the collaborative writing, with the help of word processor documents (Google docs, Zoho), that are supporting sharing the docs and also providing processing functions such as columns, inserting images, comments, formulas, etc.,. With the electronic editing the editor will be in charge of organizing the content but will let the collective reader to choose what they want to know.

#### Multivariant narratives

The way we tell and read stories is related to their material mediums. Emergence of digital media coincides with a crisis in literature around truth and authority. Digital media present new opportunities for the development of multivariant narratives. Those are stories that go through different cycles or loop, meaning that a story can have many different endings depending on how the person decides to play the stories out. This type of narrative is common when we try to construct truths from oral tradition. Over time truth becomes uncertain and our minds construct different endings. For Marie-Laure Ryan, the most important theoretician of the multivariant narratives, classic narrative is a „a medium-free, semantically based definition, that lives in texts but can emerge from stimuli that are not narrative; narrative is not only linear but vectorial, it must be followed from the beginning to end”. Digital texts are durable in that they are recorded but subject to unstable platforms, also they can be kaleidoscopic, producing a vast number of versions from a limited number of elements. One of the characteristics of the multivariant narrative is the fact that it implies a variable point of view, the reader having access to various characters’s perspectives, the most common example being that of the videogames (like a „god game”, an artificial life game that casts the player in the position of controlling the game as an entity with divine powers: Populous, Dungeon Keeper, Black@White, etc.).

### Thematic research collections

One of the easiest and most rewarding ways to begin engaging with the digital research is to build what Carol Palmer calls a „thematic research collection”. Palmer describes it as „digital aggregations of primary sources and related materials that support research on a theme”. Such a collection is distinct of an archive because it does more than simply collect materials. A thematic research collection also includes the research commentary on those materials, the researcher composing the text that will guide visitors through the collection, sort of museum exhibits contextualized. One excellent tool to create a thematic research is Omeka, a web-publishing platform that allows anyone to collect various items, documents, photographs, sound files and videos and organize them into an exhibit.

### **Reading visual**

The “reading” it can be use as well in processing images, and is so called cultural analytics as the author Lev Manovich described it. He use screen displays and digital tools to analyze, organize and sort a large number of images. Through the new process way based on values, color, degrees and difference he founded one of the core research areas of cultural analytics.

Digital Humanities use a lot of activity and visual format like timelines, diagrams, chart, tables and maps. Maps as a visual representation are not accompanied by instructions about the way of reading their encoding. Using GIS (Geo-Spatial Information Systems) as operational system in maps can be encoding all kind of theory. In Digital Humanities, maps are a very important tool because using it can transform artistic, literary or historic descriptions into a visual representation. From the earliest time human being was concerned in mapping different places and routs.

Mapping includes:

Spatial viewpoint

Temporal (out of date)

Conceptual (experiential vs. literal).

Mapping is just a record of experience not of things. Maps contain hypothesis that encapsulate cultural values in various fields, at a specific historical moments. All the materials are process in the present and fix them into a single geographical representation system, or are use elements from past presentations into map formats.

In his critical postmodern analysis Edward Soja, set a distinction between is concept of “space” as a physical territory and “place” as an experiential environment. According to his theory space is a construct that comes into life through the activities of experience. These concepts produce some challenges for the visual instruments of mapping that we use at present.

Digital cartography is the process by through a collection of data is changes to a virtual image. The first role of those maps is to reproduce a particular area; since early civilization people have created maps but only recently we have had the digital tools like Google Earth or Maps to create sophisticated and interactive maps. Geographic information systems (GIS) are computer applications concerned with the manipulation of geographic information and today such software are capable of represent, analyze and make visual any kind of information about the location of divers phenomena around the earth. Remarkable

experiences have been made in different part of the world, most of them being the result of the work and study of the “geocriticism” group, headed by Robert Tally and Bertrand Westphal.

Using digital technologies in humanities research it gave birth to many changes in different disciplines. For example using digital images for research and teaching in art history it requires the ability to manipulate digital images, and to use software.

“Digital teaching ... has stimulated the development and application of tools to simulate and enhance the experience of viewing art and architecture in new ways.... These tools make it possible to unfurl scrolls, move through buildings, zoom-in on details, overlay different states of an etching, track the build-up of a painting, animate structural forces, navigate 3-D reconstruction of ruins, model an unbuilt design, and map archaeological sites. ...These tools are yielding new perspectives on the objects of study...” (Ballon Hilary, Mariët Westermann)

### **Teaching the Digital Humanities**

The rise of the Digital Humanities as a coherent field was accompanied by the creation of many Digital Humanities Centers, as a result of the evolution of the Humanities Computer Center as an institution. Laboratories such as the one in the Oxford or King’s College London are very important for the application of information technology to the humanities, they create and develop the Digital Humanities as a field with a solid theory, and they are infrastructure nodes and models in the organizing of other centers all over the world. Also those Centers are inspirational regarding collaborative and interdisciplinary work; they bring students to learn how to work on project-base, hands-on and to learn in progress, notions that are a key to a modern university education and research.

From the late 80s through the mid 90s, pedagogy held place of pride in the Digital Humanities as the 2001 conference on “The Humanities Computing Curriculum” in British Columbia demonstrated. Since then a number of workshops dedicated to the digital humanities pedagogy were held in different occasions, the most recent one being the “Innovative teaching methods and practices in the digital humanities” to be held in Lausanne at the DH International Congress, July 2014.

As for now, relatively few centers in DH offer undergraduate and graduate training, the most important being the program offered by the Department of Digital Humanities at King’s College London, under the supervision of Professor Willard McCarty. The undergraduate offers a degree analyzing digital culture, comprehending the consequences and meaning of digitization in social media, gaming, digital memory, etc. At the University of Amsterdam a project called Coding the humanities is running, that’s aimed at collaborative tool building and reflecting on the tools used in the humanities. Other programs run in many Universities of Britain, USA, but also in Luxembourg, Graz, Milan, Madrid, Belgrade or Budapest. An interesting map of the DH Centers around the world is available on the CenterNet site.

Most colleges and universities in Europe have extensive resources for research and study. Contemporary *Digital Humanities* utilize these assets by taking as partners, libraries, museums, archival collections. Extra-mural partnerships, with professional societies,

historical associations, corporations, public entities can extend the possibility to reach more and more people.

*Digital Humanities* pedagogy emphasizes teamwork as well as attention to a larger set of skills beyond classical text-based critical thinking. It emphasizes the ability to think critically with digital methods, to understand, analyze and use data, and most of all, to work collaboratively, across disciplines and methodologies.

### **Advocacy and controversy**

If there still are some voices from the conventional humanities scholars that dismiss digital humanities, we consider that there are some crucial points on which we can advocate for DH as a field: the contribution to cultural record, through the database creation or the thematic collections, the virtual exhibits, by thinking beyond the ideologies and embracing the diversity, the original, the innovation.

In conclusion digital resources improve the humanities research area with a new way to see, and represent humanistic fields and with new tools that facilitate the process of interpreting humanities materials. A different point of view on analyzing texts can lead to a new way of reading and understanding print resources.

As we see in our presentation all applications of digital humanities are based on the fundamental interpretative acts like: parameterized (what can be measure), counted, sorted, and displayed. The research results of projects in humanities domain have to compulsorily be read in relation to those issues, about the corpus under investigation.

“Digital Humanities shouldn't only be about the production of knowledge. It's about challenging the ways that knowledge is represented and shared.” (Mark Sample)

### **Bibliography**

Alvarado, C. Rafael, *The Digital Humanities Situation*

<http://dhdebates.gc.cuny.edu/debates/text/50> . Accessed at 18.05.2014.

Ballon, Hilary, Westermann, Mariët, *Art History and Its Publications in the Electronic Age*, Houston: Rice University Press, 2006.

Braudel, Fernand, *L'histoire, mesure du monde*, in *Les Ecrits de Fernand Braudel*, vol. ii, Paris, 1997.

Drucker, Johanna, *Humanistic Theory and Digital Scholarship*

<http://dhdebates.gc.cuny.edu/debates/text/34> . Accessed at 19.05.2014.

Fitzpatrick, Kathleen, *The Humanities done Digitally*.

<http://dhdebates.gc.cuny.edu/debates/text/30> . Accessed at 19. 05. 2014.

Gold, K. Matthew (ed.) *Debates in the Digital Humanities*, Minneapolis, 2012.

Kirschenbaum, Matthew, *What Is Digital Humanities?* in *ADE Bulletin* 150 (2010): 1–7.

<http://mkirschenbaum.wordpress.com/2011/01/22/what-is-digital-humanities/>. Accessed at 19.05.2014.

Levine, George, et al., *Speaking for the Humanities*. American Council of Learned Societies, 1988.

<http://www.lib.muohio.edu/multifacet/record/mu3ugb1492819>. Accessed at 19.05.2014.

McCarthy, Williard, *Humanities Computing*, London & New York, 2005.



- Moretti, Franco, *Graphs, Maps, Trees, Abstract Models for Literary History*, London & New York, 2005.
- Idem, *Distant Reading*, London & New York, 2013.
- Mullen, Lincoln. *Digital Humanities Is a Spectrum; or, We're All Digital Humanists Now. Backward Glance*. <http://lincolnmullen.com/2010/04/29/digital-humanities-is-a-spectrum-or-were-all-digital-humanists-now/>. Accessed at 18.05.2014.
- Pannapacker, William, *Big Tent Digital Humanities, A View from the Edge* <http://chronicle.com/article/Big-Tent-Digital-Humanities/128434/> Accessed at 19.05.2014
- Ramsay, Stephen, *Who's In and Who's Out*. <http://lenz.unl.edu/papers/2011/01/08/whos-in-and-whos-out.html>. Accessed at 18.05.2014.
- Reid, Alex. *The Digital Humanities Divide*. <http://www.alex-reid.net/2011/02/the-digital-humanities-divide.html>. Accessed at 17.05.2014.
- Sample, Mark, *The Digital Humanities is Not About Building, It's About Sharing*. <http://www.samplereality.com/2011/05/25/the-digital-humanities-is-not-about-building-its-about-sharing/> Accessed at 17.05.2014.
- Schreibman, Susan, Siemens, Ray, Unsworth, John, (eds.), *Companion to Digital Humanities* Oxford, Blackwell, 2004.
- Selfe, Cynthia, *Computers in English Departments: The Rhetoric of technopower* in *ADE Bulletin* 90/1988.
- Spiro, Lisa, “*This Is Why We Fight*”: *Defining the Values of the Digital Humanities*. <http://dhdebates.gc.cuny.edu/debates/text/13>. Accessed at 18.05.2014.
- Svensson, Patrik, *Beyond the Big Tent*, <http://dhdebates.gc.cuny.edu/debates/text/22> . Accessed at 18.05.2014.
- Unsworth, John, *What is humanities computing and what is not?* <http://computerphilologie.uni-muenchen.de/jg02/unsworth.html>. Accessed at 17.05.2014.
- Zorich, M. Diane, *Transitioning to a Digital World: Art History, its Research Centers, and Digital Scholarship* in *Journal of Digital Humanities*, Vol. 1, No. 2 Spring 2012. <http://journalofdigitalhumanities.org/1-2/transitioning-to-a-digital-world-by-diane-zorich/> . Accessed at 17.05.2014.