New relationships between the digital and the humanities

Science

1. Introduction: the Digital and the Humanities

Computers, digital tools and the Internet have been radically changing the way scholars work, collaborate and publish their research and supported the creation, the storage, the analysis and the dissemination of data and information.

While many areas of study within the natural, medical, and social sciences have a long tradition with these technologies, most of the humanities disciplines have been more reluctant and have found it more difficult or inappropriate to integrate computational tools that are generally built to perform quantitative analysis.

Even thought in the last years new activities and new research opportunities have emerged from the intersection between the humanities and the world of digital technologies, what we call today digital humanities, represents an undefined and heterogenous set of studies and practices that aims at understanding the implications and the opportunities that digital technologies can provide as media, tools, or objects of study in the humanities [1, 2].

What is sure is that the enormous quantity of information coming from digital collections and in a more general sense the " digital world", is offering different opportunities to rethink the traditional research activities and tasks in the humanities (Moretti, Manovic…)

These new relationships between the digital and the humanities are rapidly demanding new modes of observation, exploration, and interpretation and in

this perspective information visualization and interfaces are becoming essential tools to explore and make sense out of the increasing quantity of available data.

Due to the fact that most of the methods and technologies adopted by the digital humanities come from other disciplines, today the biggest challenge in reshaping the relation between the digital and the humanities seems to imagine and design new research tools able to embrace and emphasize humanities endeavor (Drucker, 2009).

In this context the role of communication and interface design emerges as a central issue in the current research initiatives.

In addition to this, the nature of humanities inquiry and the work scholars engage with data and representations, provide an great opportunity for design and design thinking to experiment and successfully apply their interpretative and use-centered approaches.

However, despite the increasing attention that interfaces and visualizations are receiving among scholars and researchers, HCI and design communities, show little interest or involvement in the discussion. As Burdick points out: "[…] within the digital humanities, most scholars continue to design their own projects or use research assistants or IT staff to do so. But the most troubling development is that these same scholars validate their approach by theorizing visual communication modes such as visual layout, the image, interaction, and multimedia with scant reference to the existing discourse within film, design, or the arts" [5].

2. The cultural data

The technologies and tools conceived in the last years to collect, encode and store information in digital formats, has made possible for science, social and humanities scholars to access and explore an incredible amount of data coming from analog materials as well as born digital data generated from the use of computers.

The computer-readable forms of the data and the possibility to share and access this material remotely, has deeply transformed scholar activities that once were local and tacit.

The key role of digital data is so fundamental in many scholar activities that often it become the end product of the research itself or the input to scholarly publications, and inputs subsequent research and learning.

Even though the role of digital data is becoming fundamental in the humanities, the term data has a short history if compared with the sciences and their long tradition of quantitative and qualitative approaches.

Trying to understand what humanistic data are is an important problem and the definition will determine what data are produced, how they are captured, and how they are curated for reuse (Borgman, 2009) and it can influence the design process underlying the production of tools to visualize and explore humanistic data.

If in other disciplines, where the term data has a longer history, the definition is often associated with archival information systems(nota), in the humanities is difficult to find a common point of view.

For this reason the humanities scholars, the designers and the computer scientists involved in this kind of research, need to be particularly aware about the nature and assumptions about their data, sources and epistemology. Quoting Borgman " we are only beginning to understand what constitute data in the humanities, let alone how data differ from scholar to scholar and from author to reader" (Borgman, 2009).

A good starting point to discuss and better understand humanistic data can be through the analysis of which are the data sources in humanistic research, especially in history and literature, trying to spot the main differences between them and the sources of evidence of disciplines and practices that we usually considered data driven.

2. 1 Data Sources

The sciences and humanities are really different when we talk about data sources and the degree of control over those data (Borgman, 2007).

For scientists the creation and collection of data is usually conceived for specific scientific purposes and every step of the process in most of the cases follow formal research protocols and research questions. Even if the data sources in sciences can be very heterogeneous depending on the discipline, the control over the data is crucial and plays a key role in the research activities.

On a continuum of data source and control, the social sciences occupy the middle position between the sciences and humanities: " hose at the scientific end of the scale gather their own observations, whether opinion polls,

surveys, interviews, or field studies; build models of human behavior; and conduct experiments in the laboratory or field" (Borgman, 2009).

Other social scientist rely on born digital data coming from the Internet, not only to study the online culture, but in a more general sense as a unique data source to analyze society and culture (Rogers, 2009).

The situation in the artes and traditional humanities, especially literature and history studies, is quite different. They usually don't create their own data but they rely on records, whether newspapers, photographs, letters, diaries, books, articles; records of birth, death, marriage; records found in churches, courts, schools, and colleges; or maps. Basically, " any record of human experience can be a data source to a humanities scholar" (Borgman, 2009).

Cultural records and materials are usually stored in libraries, archives, museums or other public and private agencies under a complex system of access rules and, if already digitized, through different online platforms build with different technologies.