The Digital Humanities Role and the Liberal Disciplines

Dino Buzzetti
formerly
Dept. of Philosophy
University of Bologna
Bologna, Italy
dino.buzzetti@gmail.com

Claudio Gnoli
Science and Technology Library
University of Pavia
via Ferrata 1
Pavia, Italy
claudio.gnoli@unipv.it

Paola Moscati
CNR – Istituto di Studi sul
Mediterraneo Antico
Area della Ricerca di Roma 1
Via Salaria km 29,300
Monterotondo Scalo (RM)
Roma, Italy
paola.moscati@isma.cnr.it

ABSTRACT

The subject proposed for discussion during this panel consisted in evaluating the role of digital humanities from the point of view of specific disciplinary research practices.

The first analyzed discipline, involved in the digital humanities practices and theories, is Knowledge Organization (section 2); the second is Archaeological Computing (section 3).

Categories and Subject Descriptors

A.0 [**General**]: Conference Proceedings; J.5 [**Arts and Humanities**]: Liberal Disciplines – *methodology, tools and applications*; H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing – *indexing methods, thesauruses*; J.5 [**Computer Applications**]: Arts and Humanities.

Keywords

Digital ecosystem; digital humanities; archaeology; archaeological computing; history of applications; digital information; knowledge organization systems (KOS); controlled vocabularies; taxonomy; classification; ontologies.

1. INTRODUCTION[1]

Humanities in the digital ecosystem was the main theme of this year's conference. As it is apparent from the outline of the programme, the conference comprised a keynote address on its main theme, followed by sessions of invited papers and a panel discussion, to end up with the presentation of posters and accepted papers.

The invited papers were organized by direct comparisons of traditional and computational approaches to a specific research problem, whereas the panel discussion was meant as an overview of the positioning of digital humanities from the point of view of the research practices in some significant disciplines. To provide a plausible middle ground for possible interaction a speaker on knowledge organization was invited to participate in the exchange.

AIUCD'14, September 18-19, 2014, Bologna, Italy
ACM 978-1-4503-3295-8/14/09.
http://dx.doi.org/10.1145/2802612.2802629

[1] Section by Dino Buzzetti.

An archeologist, an historian and a philologist were invited, but unfortunately only two contributions were presented for final publication, an overview of the state of the art on archeology and computing, and an outline of knowledge organization principles and techniques.

To present the attitude of influential scholars, active in specific humanities domains, towards the increasing presence of digital humanities projects and applications, was deemed of great importance to understand their research requirements and to uncover the possible answers and relevant contributions that humanities computing could usefully provide, and on the other hand, to make the still doubtful scholars and research communities better informed about the opportunities that a computational approach could now offer to their discipline's methods and practice.

Computing procedures and their employment do indeed affect and transform the customary methods of well-established disciplines and it would be pointless to insist on the necessity of their adoption, unless the evidence of actual added valued could be produced as based on firm grounds and positive results.

A dialogue between the disciplines and a newly established humanities-specific information science seems to be absolutely necessary to bridge the divide that still exists between the two communities. A crucial requirement seems to reside in the joint capacity to find suitable *data structures* for a specific discipline, i.e. "the methods needed to represent the information within a specific domain of knowledge, in such a way that this information can be processed by computational systems," combined with the capacity to find suitable *algorithms* applicable to that discipline, i.e. "the methods needed to formulate the research questions and specific procedures of a given domain of knowledge, in such a way as to benefit from the application of computational processing" [1].

Besides the intrinsic value of the two contributions presented here, the limited success of this panel, hampered by significant desertions, testifies the need to encourage a productive dialogue aimed at overcoming the difficulties still extant for a full recognition and acceptance of humanities computing methods and for a confident employment of digital tools in many a discipline of humanities studies.

2. INFORMATION FRAGMENTATION AND KNOWLEDGE ORGANIZATION[2]

The explosion of digital information is leading to fragmentation of knowledge into many different sources, media, formats, channels, and document types. While this enhances the potential for recombination and reuse of information, it also has problematic consequences on the cognitive side, as it makes deep and complex arguments as developed in traditional monographic works more difficult to be followed. This situation makes implementation of the principles and techniques of knowledge organization (KO) even more necessary. The origins of KO as a domain evolved across library and information science, computer science, philosophy, and linguistics, and the main types of KO systems, are reviewed

2.1 Fragmentation of sources

The digital era has produced an explosive multiplication of information contents. Together with the obvious benefits of being able to access vast amounts of them through the Internet, this is also bringing an increasing fragmentation of information. Information users now have to look in many different places, and to continually switch from one source to another, in order to assemble the knowledge they are looking for.

As we still rely on paper for much relevant information, we have to switch between printed documents or handwritten notes and digital devices. The latter in turn exist in various forms – personal computers, laptops, tablets, smartphones, etc. – and store information in various kinds of memories – hard disks, compact disks, USB bars, online drives... Classical texts are supplemented by audio and video documents, each with a variety of possible formats, not all compatible with each other: just to mention textual information, it can be saved as TXT, as RTF, as DOC, as PDF, as TEI, as XML, etc.

Even in the shared online environment, every author or institution can express its views and knowledge by a plurality of channels: e-mail and mailing lists for personal and collective discussion are now facing the competition of web forums, of blogs, and of social networks, both general, such as Facebook, and specialized on particular topics or media, such as Anobii. Recently, social networks have also reached the academic communities, but again several of them are available, so that while some researchers keep their Academia.edu profile updated, others do the same in Research Gate, or in disciplinary networks such as PhilPapers. Clearly, this situation makes it complicated for authors to provide access and visibility to the contents of their work in a consistent way, and for readers to find all the relevant channels where an author of interest can be followed.

An especially critical aspect of social networks is the technical interoperability of their information contents. As emphasized in his TED conference by the very inventor of the Web, Tim Berners-Lee, many social networks connect people only within their own system, that has to be accessed through a distinct password, while making connection towards and from contents in other networks problematic. The interconnected architecture of the Web is based on such open standards as Hypertext Markup Language (HTML) and uniform resource locators (URL); but often the URL of a specific resource within a social network cannot be displayed in explicit ways nor integrated easily in an external website, suggesting that the actual purpose of these tools, rather than connecting people, is keeping them bound to their own contents and advertisements.

The semantic Web would offer simple technologies to implement open connections between any information content. For example, an author can connect herself to others by specifying their names and URLs in a FOAF (Friend Of A Friend) file, and this information could be read automatically and leveraged to develop open social networks across the Internet. Unfortunately, however, most people finds it easier to use bounded networks with their friendly, appealing interfaces, than to publish data in appropriately open formats.

When checking a specific information source online, a further factor of fragmentation is the default presentation of its contents in reversed chronological order, common in blogs and social networks. While this produces a sense of being always updated, it only suits news-like contents, but makes it difficult to follow the thread of more thorough discourses.

2.2 Relevance of long arguments

All this has profound implications on the cognitive side. Almost one century ago, an important philosopher already complained that modern life made contemplation and deepening of reflection very difficult [2]. While we are submerged with information items, we have few occasions to connect them to each other in less shallow ways in order to produce *knowledge*, which ultimately consists in a graph of relationships between information units; as well as to select from the graph the items most relevant to our actual life, in order to reach *wisdom* [3].

Relationships also are a key feature in knowledge representation and indexing. An information resource only indexed with a set of isolated tags, such as `#bibliography`, `#teaching`, `#literature` and `#history`, can mean many different things depending on the relationships actually holding between the corresponding concepts: are we dealing with a bibliography *about* teaching *of* the history *of* literature, or with teaching *of* literature *about* the history *of* bibliography, or what else?... It has been noted that the evolution of bibliographic description standards, such as RDA, towards aggregates of separate metadata may imply the loss of relevant information contained in the very sequence of elements as presented in a traditional cataloging card: the new isolating approach seems not capable e.g. of distinguishing between the place of the second publisher and a second place of the first publisher [4].

Learning resources are also undergoing an evolution from traditional handbooks bound in a single volume to a plurality of multimedia materials, often organized as modules that can be used independently from each other through online learning platforms. While they can make learning more rich and interesting, they also imply a risk of focusing the learner's attention on technical sides (and technical problems, as mentioned above) rather than on the very articulation of complex contents as it can be expressed in discourse form.

In centuries, printed books have developed ordered usage formats, such as the monograph, gathering and organizing textual and image contents in a single, easily usable object. Treatises, encyclopedias, dictionaries, journals have organized their subject matter in consistent ways. While converting to digital formats, we should be careful not to lose the information contained in their very organization and presentation.

Admittedly, the digitalization of texts brings many advantages: digital words can be quickly retrieved by search engines

[2] Section by Claudio Gnoli.

independently from their position within a text or their location in any website; and can be connected to any other digital content by hyperlinks. Such free navigation across the whole Web is a very powerful knowledge tool. At the same time, it makes it more demanding to build a synthetic view of a topic from the variety of resources scattered in the Internet.

The information unit, once consisting in a whole printed volume or issue, has become in Web "1.0" a shorter document carried in a file, e.g. an article published as a web page. With Web 2, it has further been fragmented into shorter posts, tweets, etc. Modularized pieces of information, such as an image or a video or an RSS text feed, can be embedded into pages obtained by a *mash-up* of several sources. Web 3, that is the semantic Web of linked data, means the ultimate fragmentation of information into individual *data*, recorded as RDF triples of the form Subject-Relationship-Object. The essential role of relationships is well acknowledged by this structure, making any data freely connectable and combinable with any other, including data recorded in different sources (*hosts*). However, building of consistent knowledge graphs out of these atomic elements is largely left to remote users.

2.3 KOS types

A tool for better connecting and organizing atoms of information, which is familiar to social network users, are hashtags, that is keywords identified as such by prefixing a # symbol. They have been introduced in Twitter (reportedly by Chris Messina onAugust 23, 2007) to allow grouping of tweets by different users or tweeted at different times. In order to be effective they require authors to use them consistently, which is not always the case. They have rather become a fashionable style of expression, conveying additional connotations besides just indexing, like in the famous case of `#Enricostaisereno` ("keep tranquil, Enrico"), a hashtag tweeted by Italian politician Matteo Renzi to his rival in the same party Enrico Letta, then reused by people of different political views for months afterwards.

Hashtags are but a simple example of a *knowledge organization system* (KOS), that is, any technique used to index the conceptual contents of knowledge on a given topic and make them findable by a subject search. KOSs are alternative and complementary to descriptive indexing elements, such as titles, formats or author names, which are rather aimed at identifying an information resource by its formal features [5].

There are many KOS kinds, ranging from simplest and easier ones to most advanced and demanding ones. Hashtags belong to a simple KOS family, that of *tags* or *keywords*, basically being words selected by authors or other users as especially representative of a resource content, without any fixed rule. Other examples of them include keywords of academic papers and "categories" of blog posts. By selecting a tag, users can be redirected to more resources on similar topics.

As mentioned, this only works if tags are assigned in consistent ways. The establishment of rules to do it means moving to the next type of KOSs, *controlled vocabularies*. In this case, an information resource cannot be assigned any word, but only selected ones from some list. Such selected words or phrases are then called *terms*. Terms in a controlled vocabulary, such as a subject heading list or a thesaurus, are also normalized as to their spelling, grammatical form (plural vs. singular, noun vs. verb), and lexicon (selection of a preferred term over its synonyms).

Unlike subject heading lists, thesauri also include standard hierarchical relationships between terms, specifying that a given term is narrower or broader than another. The same kind of relationship is usually found in *taxonomies*, that is hierarchies of categories used to present the classes and subclasses of some phenomena, like families of plants or of languages, or the items of a menu, like in a recipe-book or a website.

Further sophistication can be introduced in the KOS by fixing a meaningful order between classes listed in the same array. This can be done by associating to each class a notational code specifying its position within the sequence. A *classification* is then produced.

We will finally mention a KOS type having recently become popular among professionals of information representation and automatic processing: *ontologies*, which take their name from philosophy and provide, besides all the other features described above, a formal specification of the relationships between classes and of logical restrictions on them.

2.4 KO as a field

Some KOS forms have been used for millennia by philosophers and scientists trying to organize the variety of available information into consistent presentations. This was the case with classifications of traditional disciplines as taught by Andronicus of Rhodes, Martianus Capella, Francis Bacon, Jean-Baptiste d'Alembert, Melvil Dewey, Henry Evelyn Bliss, S.R. Ranganathan, and many others. The last three authors applied classification to the organization of books in library shelves.

Of course, systems of knowledge always evolve, so that there is no ultimate KOS. On the other hand, some KOS is always better than none, as knowledge needs to be presented, learned and processed in systematic ways, in order to stimulate further research. Indeed, quoting Herbert Spencer, "science is organized knowledge".

In 1974, Ingetraut Dahlberg introduced a more systematic approach to the principles of what she called *Wissensordnung*, "knowledge ordering". This field, synthesizing principles from library and information science, computer science, linguistics, and philosophy, has then become known by its English name *knowledge organization* (KO) [6], and is now well structured with regular conferences and a journal promoted by the International Society for Knowledge Organization (ISKO). Still many authors throughout different knowledge fields are actually doing knowledge organization every day without being aware of its theorization. Indeed, the field is a very general and basic one, that is much needed in all knowledge activities.

Nowadays, knowledge to be organized is usually stored in documents, such as books or Web pages or linked data, although oral knowledge can also be organized by formulas, narrations, or mnemonic techniques [7; Dousa, pers. comm.]. In that it addresses the conceptual contents of knowledge, the field of knowledge organization should be distinguished from the complementary ones of knowledge representation (concerning the recording of knowledge in standard formats), knowledge engineering (concerning the processing of represented information), and knowledge management (concerning the flows of information within private or public organizations).

All these specialties, and knowledge organization in particular, are much necessary in order to face the contemporary fragmentation of information sources, and to leverage them in more systematic and fruitful ways.

3. TOWARDS 2020. THE LEGACY AND CHALLENGES OF ARCHAEOLOGICAL COMPUTING[3]

When the organizers of the 2014 AIUCD Bologna Conference kindly invited me to present a paper illustrating the state-of-the-art of digital archaeology, I focused on the legacy and challenges of archaeological computing, a discipline that originated in the 1950s. Changing the order of priority, I would rather start with the most recent challenges that find in technology the necessary complement to an effective implementation of strategies for data acquisition, processing and representation [8]. I will come back to the issue of the legacy in the concluding remarks, by addressing some theoretical aspects underlying the application development process, which is still permeated – as a timeless hallmark – by the pioneers of computerization in the field of archaeology.

To narrow down the vast field of investigation, I will examine some of the latest and most innovative trends, which characterize three basic areas of archaeological practice: fieldwork, laboratory analyses, and cultural resource management. In this discussion, the Journal *Archeologia e Calcolatori*, now in its twenty-fifth year of publication, has provided a vantage point for illustrating the current research scenario (http://soi.cnr.it/archcalc/).

3.1 Field Archaeology

3.1.1 Sites and landscapes

In archaeological fieldwork, spatial data acquire relevant importance and become an integral part of the investigation and a source of knowledge for ancient sites and territory, when analyzed within their own geographical context. In line with landscape archaeology purposes, the computer-based approach benefits from the emergence of Geographical Information Systems (GIS) which, since the 1990s, have represented one of the strategic cornerstones of archaeological computing [9].

As a result, digital mapping has found its way into a methodological revitalization process. Spatial data analysis – applied since the 1970s in the Anglo-Saxon countries in the wake of quantitative geography and ecology studies – changes its perspective, while remaining anchored to its rationale: the logic of archaeological data distribution originates from policy choices that, never by chance, have led to aggregation and concentration processes. With the emergence of GIS, spatial analysis in general and the concept of spatial data in particular expand from a precise attribute intended to produce distribution maps, to an investigation tool focused on the reconstruction of complex structures, in which land morphology and human action can help to understand the dynamics of settlement patterns.

The methodological debate, therefore, is hinged upon the procedures required to record and formally represent spatial data. In archaeological field-walking, data acquisition criteria take into account the way in which information is collected; in excavations, the analysis of topological relationships and the study of the functional use of space nicely complement the precise location of material remains. In addition to conventional topographic and photogrammetric surveys, data can now be acquired by means of specific sampling methodologies based on satellite and airborne remote sensing [10]. For example, over the past decade the LiDAR remote sensing technique has been widely used for landscape and urban analysis, thanks to data acquisition accuracy and the ability to penetrate dense vegetation canopy.

[3] Section by Paola Moscati.

In geophysical prospecting, state-of-the-art GPR technology has provided 3D high-resolution images of the subsurface, leading to major achievements in investigating the stratigraphy of archaeological features across the upper layers of the subsoil and in locating natural and artificial cavities. Underwater archaeology, generally supported by highly sophisticated sonar technology, is further enhanced by the use of robotics for data acquisition purposes. Robotics is also used in complex investigations, e.g. when almost inaccessible environments (hypogea, caves, underground passages) are to be explored or remote visits are to be planned.

In landscape archaeology, spatial statistics are currently based on Digital Terrain Models data within a GIS environment and are generally aimed at investigating space distribution phenomena, also in order to predict the location of archaeological sites. In landscape modeling, distance calculus takes advantage of geo-referenced data to quantify relevant factors, like the distance from watercourses, between sites and routes, and between major and minor settlements. Other forms of analysis facilitate the study of the impact of past natural and artificial barriers vis-à-vis spatial movement patterns. In addition, visibility analysis focuses upon the role played by human visual perception in settlement, urban and architectural choices.

The transition from a deterministic approach to a socio-historical dynamical configuration of land organization structures has resulted in a new archaeological research field: archaeogeography. Under the leadership of French Medievalists, archaeogeography investigates the memory, transmission and transformation over time of specific rural and urban planning structures, such as the Roman centuriation or the Medieval “encastellation”, thanks to a chronologically transversal approach and an integrated investigation procedure, which is based on a dialogue between archaeologists, historians, geographers and paleoecologists.

3.1.2 2D and 3D digital archaeological drawings

Computer graphics is, without any doubt, another field at the cutting edge of technological innovation so much so that someone has even talked once again of the “death” of archaeological drawing in the same way as occurred with the advent of photography. Whilst in the early 21th century the archaeological scenario was still essentially characterized by 2D digital documentation techniques, the use of 3D models has now prevailed and is supplemented with new solutions aiming at documenting and visualizing data relevant to sites, monuments and objects, with a significant impact on reconstruction and restoration activities [11].

In topographic field surveying, innovation is connected with the advancement of tools geared to carry out extensive and high-speed prospection campaigns and with the growing integration of multiple techniques, which call for a multi-resolution approach intended to produce metric reconstructions consistent with different geometries. As a result, robotic total stations and differential global positioning systems (DGPS), aerial photogrammetry and airborne and satellite remote sensing are being increasingly used.

To assist in field surveying operations, hand-held and tablet devices have proved particularly useful for transferring information to an archaeological laboratory with suitable computer equipment, by using a wireless connection. A new approach is therefore emerging, referred to by some as tele-archaeology, a kind of “remote” archaeology whereby the time elapsing between data acquisition and post-processing is cut down and fieldwork methods renewed. Webmapping, webGIS and GIS Cloud techniques can be used in a virtual environment in which multiple investigation procedures that so far had been kept separate, such as field operations and lab data processing, can now be synchronized.

Innovation is also part of the archaeology of architecture (in short "archaeotecture"), specifically aimed at "reading the walls". Although it derives the bulk of its instruments from archaeological excavation techniques, data recording aims at documenting built heritage characteristics, from wall stratigraphy to construction materials and building techniques. When documenting plane surfaces, such as the facades of historic buildings, the combined use of monoscopic photogrammetry and 3D laser scanner is replacing the use of direct survey methods and instruments. Photomodeling is also gaining ground as a user-friendly technique that makes use of digital cameras to capture 2D images and return 3D metric models manageable in a GIS environment.

3.2 Laboratory analyses

3.2.1 Artefact classification

In artefact classification, the statistical approach is still a well established methodological reference model, although applications are decreasing. The most widespread trend is in favor of a specialization addressed to specific issues such as the reliability and representativeness of the sample population. Data quantification tends to overcome the traditional encoding procedure based on presence/absence criteria in search of more nuanced investigation solutions, which are increasingly flexible but not neutral, such as those provided by multi-criteria analysis techniques to support the decision making process during the systematic organization and synthesis of complex information.

Building on the success of the *Analyse des Données* approach promoted in France since the 1970s, multivariate statistical techniques, in particular Correspondence Analysis, are still opted for in exploratory data analysis, since they apply well to investigation fields where it is necessary to recover as much synthesized information as possible. These methods identify the variables that most significantly contribute to data classification and as such they offer guidelines and suggest possible interpretation of data structures by highlighting systematic relationships that can be used for predictive purposes as well.

Another investigation field that is being revitalized is the one applying "New Artificial Intelligence" tools [12]. Computer simulation based on formal modeling provides an ideal method to experimentally investigate ancient societies and to understand past human behaviors in relation to environmental, social and cultural variables. In addition to agent-based models, capable of detecting agents (individuals, families and settlements) and processes (food production, adaptation to the environment, social organization) that prove useful for modeling complex systems, Artificial Neural Networks, whose architecture mimics the brain's learning system, are employed in archaeology especially for simulating settlement processes but also for classification purposes.

3.2.2 From databases to digital repositories

In the process of cataloging and recording cultural heritage resources, archaeology has undergone a natural evolution from databases to multimedia systems. Major research projects reveal the intent to shift from computer-based item and collection cataloging to integrated classification systems in which the web turns into an environment available for consultation and knowledge sharing, especially in cases where government departments, universities, research institutes, regional and local authorities are called upon to make a joint effort to promote and coordinate cultural heritage conservation and planning policies.

A case in point is provided by large archaeological *corpora* that by tradition represent the solid foundation upon which any comparative study should be based. In this context, the Internet has played a primary role in innovating archaeological communication. As is customary in any digital archive, the focus is placed on the creation, description and preservation of digital contents as well as on the repository management and data sharing through the use of exchange protocols [13].

International projects, such as the Beazley Archive Computer Project, or the CVA and LIMC associated initiatives, improve data harmonization and standardization, with a revival of interest intext encoding, metadata collection and indexing operations. As archaeological data exchange via the web intensifies, representation standards and data transmission protocols, such as the OAI-PMH (Open Archives Initiative-Protocol for Metadata Harvesting), gain momentum from a supranational viewpoint and e-infrastructures and services are planned to facilitate interoperability between digital repositories, as in the very recent case of the Ariadne European project.

In research work oriented to the identification of ontological structures for data classification, in view of a more targeted retrieval, the CIDOC CRM (Conceptual Reference Model), promoted by the International Council of Museums can be hailed as a supranational semantic approach for extracting domain taxonomies in the cultural heritage sector. In 2011 CLAROS (Classical Art Research Online Services) international portal went live. Today, it includes several archives that can be simultaneously searched through Semantic Web technologies, and ensures interoperability and virtual access to major Greek and Roman collections. Each content provider keeps its own data format and website unchanged, and users can formulate their queries as if coming from a single source.

In this context, reference to the open access and open archives movement can hardly be omitted. The Journal *Archeologia e Calcolatori* was pioneering in this perspective. Not only did it join the Open Archives Initiative as early as in 2005, but it has also promoted the debate on Open Science. For this reason, the ArcheoFOSS Scientific Committee selected the Journal to publish the proceedings of some of its most recent annual conferences. Furthermore, the SITAR working group started a fruitful cooperation with the Journal to publicize and disseminate the communication features of the Archaeological Geographical Information System of Rome, whilst specific contributions from expert scholars addressed the theme of open digital archives and open access archaeological publishing.

3.3 Cultural Resource Management: preserving and enhancing the archaeological heritage

3.3.1 Preventive archaeology

Protection, conservation and enhancement of archaeological heritage is an area that more than any other has witnessed the involvement of archaeologists in the economy of each country. By developing methods and priorities of rescue or preventive archaeology, and reconciling infrastructure works in urban and rural environments with the integrated conservation of cultural heritage, archaeologists are now part of the process of safeguarding the knowledge of the history of civilizations.

In the European arena, the pioneering experience of the French INRAP (Institut National de Recherches Archéologiques Préventives) was critical for ensuring coordination between archaeological protection and urban planning requirements, based on the polluter-pays principle, and for developing understanding and awareness of archaeology in the general public. In Italy, the expression preventive archaeology was first introduced in association with the development of some legislative and regulatory issues. However, the practice of preventive archaeology was the outcome of a well-established activity – e.g., the test bench of the metropolitan train networks and the high-speed train projects – which is based on traditional archaeological methods not only for knowledge acquisition and dissemination purposes, but also for adequate expert evaluation of potential risks for the archaeological heritage by development and building of infrastructures.

From a technical viewpoint, GIS can once again be regarded as an ideal IT platform that is capable of generating – according to well-established monitoring procedures – risk assessment maps that help estimate the level of aggressiveness of physical, chemical, biological and environmental factors and the vulnerability index. In terms of built heritage, both in ancient urban centers such as Pompeii, and large architectural complexes such as Hadrian's Villa in Tivoli, results coming from the monitoring of archaeological risk are recorded as GIS attribute values.

For archaeological prediction and risk management, statistical analyses are widely used: the combined reading of different sets of data allows researchers to evaluate the levels of probability of encountering archaeological deposits. Predictive modeling can guide urban planning and environmental policies to mitigate, through preventive measures, the incidence of risks or at least to reduce the damage [14]. Therefore, the term "archaeological risk maps" has been recently replaced by "archaeological potential maps", which indicates in the historical presence in any given region an added value for an efficient and shared urban and infrastructural planning policy.

3.3.2 Virtual musealization

Programs and facilities for the exploitation and enjoyment of archaeological sites, monuments and objects have greatly increased due to Virtual Reality techniques that have affected a number of sectors, generating new exploration paths: the landscape turns into a dynamic place in which man-environment interaction is driven by past and present practices; monumental complexes regain their initial appearance, with relevant implications for the diagnosis of the state of deterioration and for restoration work, favoring the education industry.

Musealization greatly benefits from 3D modeling and Virtual Reality techniques to design exhibition set-up and display of finds and collections, also in relation to new museum concepts, such as narrative museum or diffused museum system [15]. Museology meanings and purposes are no longer confined to collecting, preserving, and displaying items of cultural and artistic significance, but special efforts are increasingly devoted to a capillary knowledge dissemination activity. The focus is therefore shifted towards the community of visitors, through a dialogue that goes well beyond the boundary of a virtual tour and engages visitors in a more complex information system.

Innovative communication systems are tested giving rise to scientifically relevant achievements and producing an important impact on knowledge and education: the requalification of specific monumental complexes, in particular those inaccessible for the general public; the development of systems to automate traditional research and analysis procedures, such as the computer-aided recomposition of fragments or the virtual integration of missing architectural components; the development of systems designed to explore 3D digital models and experience direct haptic contact especially designed for impaired people.

Archaeologists, whose primary goal is to preserve the contents that are virtually represented and animated, test innovative research methods, among others the development of museum networks connecting various heritage resources stored in different geographical locations, by virtually reunifying what is physically set apart. As a result, scholars diversify the channels through which their knowledge is made available, with the support of experts in cognitive science and communication. The outmost reaches of the virtual world are explored, with solutions provided by telepresence systems and augmented reality. Visitors/users immerse themselves in a museum scenario, by interacting with the exhibited items, customizing their itinerary and engaging in a dialogue with an avatar that is almost indistinguishable from a human.

3.4 Theoretical legacy

The theoretical legacy of archaeological computing, despite today's technological innovations, is still directly connected with the early achievements of the pioneers of computer applications [16]. The collection of materials required to set up the virtual museum of archaeological computing – a joint project between CNR-ISMA and the Centro Linceo Interdisciplicare "Beniamino Segre" of the Accademia Nazionale dei Lincei – is undeniable proof of this trend.

An extended community of protagonists, each following his own research route, has succeeded in allowing archaeology to reach its own current *status*, on a par with that of other exact sciences. It is not possible to follow in the footsteps of all the pioneers, especially those scholars who, in the 1960s, gave birth to the New Archaeology movement. However, a tribute should be paid to the emblematic figure of Jean-Claude Gardin, the unquestionable protagonist of the birth of archaeological computing, who died in 2013 [17]. The Fund "Equipe Archéologie de l'Asie Centrale et Jean-Claude Gardin", kept in Nanterre in the Archives of the Maison Archéologie & Ethnologie, René-Ginouvès, has been instrumental in drawing a detailed picture of his unremitting efforts to promote innovative research routes and to establish new laboratories where the formalization of some aspects of archaeological research was forged, like data representation and classification processes and the construction of scientific knowledge.

The modernity of Gardin's thought is unparalleled: after establishing the *Centre mécanographique de documentation archéologique* in 1957, later known as *Centre d'analyse documentaire pour l'archéologie*, Gardin actively participated in the cultural movement that revolved around the new methods of automatic processing of scientific information, promoting a number of initiatives supported by UNESCO and EURATOM and working in cooperation with important French scholars such as Henri Seyrig, Fernand Braudel, Jean Leclant, André Leroi-Gourhan and Claude Lévi-Strauss.

Some of Gardin's most famous statements, taken from his writings and from archival documents, bear witness to the innovative nature of his approach. First, the need to draw attention to formal methods and their theoretical implications when interpreting archaeologist's reasoning, instead of giving priority to the tools that technology was generating.

Second, the central role of data encoding that, according to Gardin, accounts for the essential preliminary phase, which is followed by data input in the analytical documentary process, either automatic or manual. Data encoding, thus conceived, tends to break down and describe individual components as well as their interrelationships, and therefore can be applied to an extensive class of archaeological materials and at different descriptive levels with varying degrees of complexity. So, in the cataloging of archaeological items for classification purposes, Gardin recognizes the potential to give rise to a collection of *catalogues virtuels* or *classifications potentielles* that could be multiplied at will, thus introducing the concept of virtual representation much ahead of time.

The typical instruments used to carry out documentary research are combined together by Gardin under the same heading: meta-languages, which should include a specific lexicon of terms for indexing purposes, a semantic organization of these terms, and a grammar governing their relationships. By quoting Gardin again, we can find some consonance with the words of modern partisans of data standardization and knowledge sharing according to the principles promoted by the Semantic Web: «The merits of such designations tend to be threefold: objective, i.e., conform to fixed standards of description, irrespective of personal appreciations; international, i.e., independent of national differences in the process of naming identical entities; analytical, i.e., capable of being broken down into several terms, which makes for a more compact storage, and a more flexible retrieval of information».

This is what inspired the theoretical approach of Gardin and of all those pioneers who accepted the challenges launched by the advent of computers: research should follow a spiral-shaped cycle in which data analysis and interpretation lead to a new data formalization, the outcomes of which must be integrated into further analysis. The theoretical impact of this spiral should represent, in the debate on the relationship between archaeology and science, a caveat for those who believe that digital archaeology is only the outcome of technology.

4. REFERENCES

[1] Thaller, M. 1999. Defining humanities computing methodology. In *Computing in Humanities Education: A European Perspective*, K. de Smedt et al., Eds. University of Bergen, Bergen, Ch. 2.3. URL=http://www.hd.uib.no/AcoHum/book/fm-chapter-final.html.

[2] Hartmann, N. 1926. *Ethik*. De Gruyter, Berlin.

[3] Marshall, B. 2013. Information, knowledge, wisdom. In *Twitter*. http://twitter.com/flowchainsensei/status/408167162344648704.

[4] Petrucciani, A. 2008. Le nuove RICA nel contesto dell'evoluzione delle normative catalografiche. *Bibliotime* 11, 1 (Mar. 2008). http://www.aib.it/aib/sezioni/emr/bibtime/num-xi-1/petrucci.htm.

[5] Gnoli, C. 2014. Knowledge organization systems. In *Subject access to information: An interdisciplinary approach*, K. Golub, Ed. ABC-CLIO, Santa Barbara.

[6] Dahlberg, I. 1993. Knowledge organization: Its scope and possibilities. *Knowl. Org.* 20 (1993), 4, 211-222.

[7] Ong, W. 1982. *Orality and literacy: The technologizing of the world*. Methuen, New York.

[8] Moscati, P. 2013. Archaeology and Computers: Perspectives of the 21st Century. In *Mediterranean Archaeology: A GID-EMAN Training Course* (Rome, October 2012). Atti dei Convegni Lincei 275, Roma, 91-108; Earl, G. et al., Eds. 2013. *Archaeology in the Digital Era. Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology* (Southampton, March 26-29 2012). Amsterdam University Press, Amsterdam.

[9] Bodenhamer, D. J., Corrigan, J., and Harris, T. M., Eds. 2010. *The Spatial Humanities. GIS and the Future of Humanities Scholarship*. Indiana University Press, Bloomington.

[10] Lasaponara, R. and Masini, N., Eds. 2012. *Satellite Remote Sensing A New Tool for Archaeology*. Springer-Verlag, Dordrecht-Heidelberg-London-New York.

[11] Curci, A. and Fiorini, A., Eds. 2012. Documentare l'archeologia 2.0. Atti del Workshop (Bologna, 19 aprile 2012). *Archeologia e Calcolatori* 23, 209-325; Curci, A. and Fiorini, A., Eds. 2013, *Documentare l'archeologia 3.0. Atti del Workshop* (Bologna, 23 aprile 2013). *Archeologia e Calcolatori* 24, 325-394.

[12] Ramazzotti, M., Ed. 2014. ARCHEOSEMA. Artificial Adaptive Systems for the Analysis of Complex Phenomena. Collected Papers in Honour of David Leonard Clarke. *Archeologia e Calcolatori* Suppl. 6.

[13] Kansa, E.C., Kansa, S.W., and Watrall, E., Eds. 2011. Archaeology 2.0: New Approaches to Communication and Collaboration. *Digital Archaeology Series*, 1. https://escholarship.org/uc/item/1r6137tb.

[14] Kamermans, H., van Leusen, M., and Verhagen, P., Eds. 2009. *Archaeological Predictionand Risk Management. Alternatives to Current Practice*. Leiden University Press, Leiden.

[15] Moscati, P., Ed. 2007. Virtual Museums and Archaeology. The Contribution of the Italian National Research Council. *Archeologia e Calcolatori*, Suppl. 1.

[16] Moscati, P., Ed. 2009. La nascita dell'informatica archeologica. Atti del Convegno internazionale(Roma, 24 ottobre 2008). *Archeologia e Calcolatori* 20.

[17] Moscati, P. 2013. Jean-Claude Gardin (Parigi 1925-2013). Dalla meccanografia all'informatica archeologica. *Archeologia e Calcolatori* 24, 7-2.

ASSOCIAZIONE PER
L'INFORMATICA UMANISTICA
E LA CULTURA DIGITALE

in collaboration with

ALMA MATER STUDIORUM
UNIVERSITY OF BOLOGNA

SCHOOL OF ARTS, HUMANITIES, AND CULTURAL HERITAGE
DEPARTMENT OF CLASSICAL PHILOLOGY AND ITALIAN STUDIES
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Humanities and Their Methods in the Digital Ecosystem

Proceedings of **AIUCD 2014**

Selected papers

aiucd
2014
La metodologia
della ricerca umanistica
nell'ecosistema digitale
III convegno annuale
18-19 settembre, Bologna
Biblioteca del Dipartimento
di Filologia Classica e Italianistica (BFI)
via Zamboni 32 - Piano I

Ed. by Francesca Tomasi, Roberto Rosselli Del Turco, and Anna Maria Tammaro

Published by ACM

The Association for Computing Machinery
2 Penn Plaza, Suite 701
New York New York 10121-0701

ACM ISBN: 978-1-4503-3295-8

CONFERENCE COMMITTEE

General Chair

Dino Buzzetti, University of Bologna, Italy – AIUCD President

Local Chair

Francesca Tomasi, University of Bologna, Italy - Department of Classical Philology and Italian Studies

Program Committee

Maristella Agosti, University of Padova, Italy;
Fabio Ciotti, University of Roma Tor Vergata;
Maurizio Lana, University of Piemonte Orientale;
Federico Meschini, University of Tuscia;
Nicola Palazzolo, University of Perugia;
Roberto Rosselli Del Turco, University of Torino;
Anna Maria Tammaro, University of Parma.

Local Program Committee (University of Bologna)

Luisa Avellini - Department of Classical Philology and Italian Studies
Federica Rossi - Leader Library of the Department of Classical Philology and Italian Studies
Marialaura Vignocchi - Leader AlmaDL and CRR-MM
Fabio Vitali - Department of Computer Science and Engineering

Secretary (University of Bologna)

Marilena Daquino - Research assistant
Federico Nanni - PhD

PREFACE

The conference call and topics

AIUCD 2014,[1] the third AIUCD (Associazione per l'Informatica Umanistica e la Cultura Digitale[2]) Annual Conference, was devoted to discussing the role of Digital Humanities in the current research practices of the traditional humanities disciplines.

The introduction of computational methods prompted a new characterization of the methodology and the theoretical foundations of the human sciences and a new conceptual understanding of the traditional disciplines. Art, archeology, philology, philosophy, linguistics, bibliography, history, archival sciences and diplomatics, as well as social and communication sciences, avail themselves of computational methods to formalize their research questions, and to innovate their practices and procedures. A profound reorganization of disciplinary canons is therefore implied.

New emerging notions such as Semantic web, Linked Open Data, digital libraries, digital archives, digital museum collections, information architecture, information visualization have turned into key issues of humanities research. A close comparison of research procedures in the traditional disciplines and in the digital humanities becomes inevitable to detect concurrencies and to renew their tools and methods from a new interdisciplinary and multidisciplinary perspective.

We invited therefore submissions related, but not confined to the following topics:

— Digital humanities and the traditional disciplines
— Computational concepts and methods in the humanities research practices
— Computational methods and their impact on the traditional methodologies
— The emergence of new disciplinary paradigms

Submission, review and selection process

The call showed a large interest on these topics from both the computer science and the humanistic communities: 29 abstracts were submitted for a first evaluation; of these, 10 were accepted as main conference speeches and 6 as posters. Final papers were subsequently submitted for a first peer-review. The members of the Program Committee as well as additional reviewers were given the task to evaluate them. A second peer review was necessary in order to verify the authors' acceptance of the reviewers' suggestions.

Eventually 13 submissions were accepted for the present proceedings, and organized in full papers and poster papers. Also the invited talk papers were evaluated by the program committee.

Acknowledgments

Our thanks go to: the School of Arts, Humanities, and Cultural Heritage, the Department of Classical Philology and Italian Studies and the Department of Computer Science and Engineering for the sponsorship of the event.

We also thank all the attendants of the conference: more than 100 people physically attended the meeting in Bologna.

The Program Committee thanks our external reviewers for the effort in the revision process.

Finally, we are grateful to the Library of the Department of Classical Philology and Italian Studies for having accepted and hosted the workshop.

[1] Conference web site: http://aiucd2014.unibo.it/.

[2] FB: http://www.facebook.com/groups/aiucd; WIKI: http://linclass.classics.unibo.it/udwiki/; WEB SITE: http://www.umanisticadigitale.it/; LIST: aiucd-l@humnet.unipi.it; BLOG: http://infouma.hypotheses.org/.

TABLE OF CONTENTS