

DéjàVu Project as a digital solution to Assist the appraisal of documents Focused on digital humanities

Bianca Hernandez

Postgraduate Program of Information Science,
Universidade Federal de Santa Catarina, Brazil.

Email: bianca.hernandez@ufsc.br

Orcid: 0000-0002-3742-8882

Luciane Paula Vital

Postgraduate Program of Information Science,
Universidade Federal de Santa Catarina, Brazil.

Email: luciane.vital@ufsc.br

Orcid: 0000-0003-2526-227X

Abstract

This article aims to analyze the DéjàVu Project as a complementary solution to the archival function of document appraisal in the digital humanities. It is exploratory and qualitative research with technical procedures related to the bibliographic study of digital humanities and documents. The digital solution refers to the combination of digital activities of data collection and cleaning, followed by organization and synchronization for later understanding and inference in various fields, including appraisal support. We concluded that using the solution developed in the DéjàVu Project has the potential to be used complementary to document appraisal. It provides a context substrate for a more accurate, pluralistic, and representative appraisal of social demands. It primarily facilitates the attribution of social values in the Macro-Appraisal process.

Keywords: document appraisal; digital humanities; DéjàVu Project

How to Cite: Hernandez, B., & Vital, L. P. (2023). DéjàVu Project as a digital solution to assist the appraisal of documents focused on digital humanities. In E.B. Alvarez (Ed.), *Digital Humanities: Visions and Applications. Advanced Notes in Information Science, volume 3* (pp. 22-46). Pro-Metrics: Tallinn, Estonia. DOI: 10.47909/anis.978-9916-9906-1-2.45.

Copyright: © 2023, The author(s). This is an open-access work distributed under the terms of the CC BY-NC 4.0 license, which permits copying and redistributing the material in any medium or format, adapting, transforming, and building upon the material as long as the license terms are followed.

1. INTRODUCTION

The Information Society was established in the 20th century with the development of computing-driven technologies (Capurro & Hjørland, 2007), which initiated a generalized process of redefining the world. In Archival Science, this was reflected in discussions on the theory, practice, and epistemology of science (Soares, Pinto & Silva, 2015) and also on the nature and purpose of the archive and the role of archivists and institutions in preserving documentary heritage (Eastwood & Macneil, 2016, p.7).

In Canada, Archivology found fertile ground for renewing its discourse, especially under the leadership of the theorist and archivist Terry Cook, within an approach that became known as Functional or Postmodern Archival Science and which also has names such as Brien Brothman, Rick Brown, Bernardine Dodge, Verne Harris, Eric Keteelar, Joan M. Schwartz, and Tom Nesmith, to name a few.

The Postmodern Archival Approach resides, according to Tognoli (2013), in Archivology, taking into account the sociocultural and ideological contexts of document production, which constitute assets for the construction of social memory. The same author argues that in this approach, the context is more important than the document itself since information registration has been endowed with symbolism, ideologies, and intersubjectivity since its creation, and these must be considered. As important as the context for understanding documents, postmodernists face two central questions, as Eastwood and Macneil (2016) and Cook (1997) point out: That archivists are agents in the formation of documentary heritage, and that archival theory must also be based on the context of production and use of information by users, instead of presenting itself as unique and static.

In this document, we propose as a background the Document Appraisal archival function, which refers to the detailed study of the uses of documents regarding the period in which they were produced to establish values and, based on these values, the periods of custody and the destination of the sets (disposal or permanent storage). The appraisal aims to maintain the administrative, tax, and legal capacity for the necessary time, vestigial probative and informative capacity, and the production of the documentary heritage of an organization or community. This ulterior purpose is directly related to humanistic interests since the trail of remaining documents allows society, based on the past, to act in the present and future, correct routes and social injustices, and exercise citizenship and social leadership. However, the archivist must know the document production and use context to carry out the assessment. In this context, digital technologies can streamline the process once done manually; or they can even generate patterns of social behavior that are auxiliary in understanding the context beyond that of the organizational production, that is, from external sources of information.

The DéjàVu project emerges as an initiative at Unicamp that uses content mining from different sources, on and off the web, to reconstruct contexts related to phenomena, objects, and specific places, followed by the organization of the material based on spatial and temporal coherence and, finally, opening space for different inferences regarding the content. Therefore, it is assumed that the survey of material related to the context through the digital solution can benefit the appraising of a set of documents. A problem that motivates this research is knowing which aspects of the DéjàVu Project can be used to appraise documents. Therefore, this article aims to analyze the DéjàVu Project

as a digital humanities solution that complements the archival function of appraising documents.

2. METHODOLOGY

Based on the purpose, the proposed research is mainly exploratory since it is a preliminary study to provide greater familiarity with the DéjàVu Project to aid the document appraisal process (Lakatos & Marconi, 2003). It also has a descriptive character, considering that the description of the conceptual aspects reflected upon their insertion in the dimensions of mediation is sought. The research can also be considered predominantly qualitative, given that it “deeps into the world of human action and relationship meanings” (Minayo, 1994, p. 22), a side best represented by the focus on subjectivity.

As for the technical procedures, this research is bibliographical, dealing with a “general overview of the main works already carried out, important, for being able to provide current and relevant data related to the theme” (Lakatos & Marconi, 2003, p.158). Therefore, it comprises a survey of information dealing with the Digital Humanities themes, especially from the perspective of the Digital Humanities Research Group of the University of São Paulo and the researchers Rodríguez-Yunta and Galina-Russel; and on Appraisal of Documents, from the understanding of Rousseau and Couture, Eastwood and Macneil, Trace, Schellenberg, and Cook.

Subsequently, the research was complemented with the technical procedure of a Case Study or Monographic, which starts by studying a representative case in depth and about its particular aspects (Lakatos & Marconi, 2003). In this article, the study refers to the Project “DéjàVu: Temporal, spatial, and characterization coherence of heterogeneous

data for integrity analysis and interpretation,” here called DéjàVu solution, based on information published on the project’s blog by the team of producers from Unicamp. Finally, from the reflections obtained by the thematic interlocutions and the software, an interpretative and reflective analysis was carried out, resulting in the present article.

3. DIGITAL HUMANITIES

Since the second half of the twentieth century, society has undergone a restructuring process driven by new information and communication technologies and the establishment of a dominant form of informational capitalism, which is reflected in various sectors of everyday life, such as the economy, the business world, work, culture, time, and space. Technologies mediated and facilitated the establishment of an Information Society, which, according to Castells (2016), acquired characteristics such as the use of information as an input, the penetration of new technologies, the connection of groups to achieve goals processed in networks and systems integrated and based on flexibility and dynamism.

Science has also used computing and digital technologies as a key to processing scientific information in general, even more so considering that first-hand sources began to be produced in digital media, and many of those found in analog media have already been digitized. Digital Humanities is an interdisciplinary field combining digital technologies with research and work in the humanities. The Digital Humanities “incorporate the methods, devices, and heuristic perspectives of the humanities and social sciences while mobilizing digital technology’s unique tools and perspectives” (Grupo De Pesquisa Humanidades Digitais, 2021, online).

Some elements indicating the consolidation of this field, as Galina-Russel (2011) points out, are the formation

of associations, organizations, and specialized centers, graduate programs, publications, projects, and events. Overall, the objectives of Digital Humanities are:

To create databases with digital resources relevant to the Humanities. This includes capturing, structuring, documenting, preserving, and disseminating data. 2. To develop methodologies for generating new elements derived from this data. 3. Generate research and knowledge to increase our understanding of the Humanities (Galina-Russel, 2011, p. 3).

The interest in the “Digital Humanities” label has been present in the social and cultural demand for humanistic sources; in the epistemological reflections of the humanities in terms of the evolution of these disciplines and the impact of new technologies; and because the word “humanist” denotes a defense of what is human, suggesting a humanization of technology. As Rodríguez-Yunta (2013) points out, the Digital Humanities arouse particular interest in the sciences that deal with information, especially regarding the creation of libraries, repositories, and digital archives for research in the humanities and the development of electronic resources for their access. Thus, the Digital Humanities are critical in safeguarding and preserving digital heritage.

Also in this context, the United Nations Educational, Scientific and Cultural Organization (UNESCO) brought in its “Charter on the Preservation of the Digital Heritage” the recognition that this heritage is composed of cultural, educational, administrative, or scientific resources, which contain technical, legal, medical, or other types of information originating from the knowledge or expression of human beings, created digitally or later converted to the digital format (UNESCO, 2003). In addition, this

document aimed at highlighting the awareness that access to digital heritage allows more excellent opportunities related to the creation, communication, and sharing of knowledge, including the risk associated with the disappearance of this heritage if its conservation is not treated as an urgent matter.

As the Digital Humanities Research Group (2021) pointed out, a common feature in Digital Humanities projects has been to provoke the relationship between traditional practices and new technologies. It is also worth noting that in the Information Society, the perspective of digital technologies combined with networks allows for an infinity of possible connections, which help provide information and enable the opening of new opportunities at a low cost to facilitate the optimization of traditionally carried out processes. Networked digital technology allows the segmentation of users, the maintenance of a multifaceted semantic context, and the phenomenon of mass self-communication (Castells, 2016).

Thus, information science, such as Archival Science, is special in understanding this contemporary social structure, digitally technological and structured in networks, and may represent a great engine for the Information Society and the Humanities in the 21st century. The Digital Humanities inserted in this context can provide instruments to enable greater agility in obtaining and applying informational inputs, ultimately favoring the achievement of humanistic objectives directly related to the records of human activities.

4. DOCUMENT APPRAISAL AS AN ARCHIVAL FUNCTION

Rousseau and Couture (1998) defined seven main archival functions (production, classification, appraisal, description, acquisition, diffusion, and preservation); however, we will

only bring the appraisal function to the center of the debate. The appraisal of documents emerges intimately linked to representation as it selects from the documentary background the representative part of this whole that will be kept for posterity and the part that will be discarded, thus being an essential task for the professional archivist. Carol Couture (2005, p.83) corroborates this understanding by stating that appraisal “is one of the most significant aspects of contemporary archival science: it is its “noble function,” its central core.

According to Eastwood and Macneil (2016, p. 8), recent literature has given special importance to the role of archives in ensuring organizational and collective personal accountability “[...] on the one hand [due to] the scandals that resulted from the incapacity, from public or private entities, to provide an appropriate account of their activities through their documents,” which denotes a concern that necessarily involves thinking about how activities are being represented and kept for reference; moreover, thinking about these activities long before the destination. However, before entering into the discussion on the appraisal of documents from the postmodern perspective, we must understand how this function appeared in the international literature.

According to Trace (2016), the first approaches to the subject date back to the end of the 19th century. Before the French Revolution, the appraisal favored the historical importance and the reduction of the documental volume, and it was not concerned with the administrative use of the documentation. In pre-World War I, England, there was an attempt to establish retention periods concerning the document’s age, use, and integrity. The British Hilary Jenkinson, an influential theorist of his time, limited the archivist role in the appraisal to separating identical

copies. At the same time, the destruction of the material that could be eliminated would be decided by the producers at the production site. In Germany, at the beginning of the 20th century, according to Trace (2016), theorists Karl Otto Mueller and Heinrich Otto Meisner also studied the value of documents. They defined it as the age and location of the producer in the administrative hierarchy.

In the United States, in the 1940s, State Department Archives employee Phillip C. Brooks started to state that documents were related to their source activities, and knowledge of these relationships was the key to identifying formal and content similarities and their destination. For him, producers could learn about the usefulness of documents for administrative functions, and archivists could learn about the effectiveness of documents for historical research. His successor G. Phillip Bauer, in turn, defended the sole and restricted custody based on the frequency of use: Even if the document had historical importance, it should be discarded if its use did not justify the cost of keeping it (Trace, 2016).

It was in the 1950s that the American Theodore Roosevelt Schellenberg, also an employee of the State Department Archives, presented an appraisal theory that defined the bases for the modern appraisal of documents. Schellenberg's document appraisal approach was based on documents' primary and secondary values. Aware of the growth in document production, driven by government activity and the use of reproduction technologies shortly after World War II, Schellenberg (1973) defended the careful appraisal of documents as a solution to the storage problem and a facilitating instrument for erudite research. According to the author, the appraisal process should be based on the analysis of the values inherent to the documents: Primary, related to the interest it represents for the producing body

itself, for administrative, fiscal, and legal purposes; and secondary, relating to the use by other entities and private users, for purposes other than those for which they were produced. Secondary values, in turn, could be divided into probative and informative.

For Schellenberg (1973), the determination of document values did not meet general standards, although the general principles presented by him could guide it; such regulations should be applied with moderation and common sense, based on a deep analysis of the documentation and not on intuition. The archival professional, according to Schellenberg, should look for specialists to help them in the assignment of value, not without first doing the work of analyzing and describing the information contained in the documents; thus, the archivist would act as the moderator with the specialists; that is, they would be responsible for showing them that a careful selection would not imply damage to the research or the maintenance of the institution's memory.

Schellenberg theorized aiming at solving the government document proliferation issue from the mid-twentieth century onwards, and this approach proved to be effective for a long time; however, for contemporary archival science, this motivation may seem somewhat reductive given current discussions involving the public interest, social memory, and the role of archivists, which Schellenberg did not contemplate. As shown by Trace (2016), other appraisal theories emerged in the 1970s and 1980s, among which she mentions those by Brichford, Frank Boles, Julia Young, F. Gerald Ham, and Hans Boom, with an emphasis on the last two who signaled an initiative to broaden the appraisal scope. Ham's contribution was to suggest that the archivist's work should be more collaborative, look beyond its destination to the larger context of information, the social;

and recognize through archives those who had no voice in society; assumptions that were also part of Hans Boom's appraisal theory, called "The Documentation Strategy," according to which the document should continually be appraised with the society at the time in which it was created, with the importance that society attributed to them.

But it was with the advent of the digital document and postmodern discussions that two new Canadian appraisal theories dictated the discussion in the last decade of the twentieth century and on which this research intends to focus: The Social Theory of Appraisal, by Terry Eastwood, and the Macro-Appraisal, proposed by Terry Cook. The Social Theory of Appraisal is based on the idea that appraisal in a democratic society must be found on the uses to which documents were submitted by the producing society during the document's existence, integrating its immediate use in the past and current use; and which should meet citizens' needs to know about their identity and place, as well as serve as accountability for democratic governments. For Terry Eastwood, the best appraisal decisions are demonstrably based on the standards of the breeding society. Therefore, archivists would better use their time by dedicating themselves to understanding the immediate and situated use of the documentation to allocate (Trace, 2016, Couture, 2005).

Terry Cook, in turn, combined theory, strategy, and methodology to develop his Macro-Appraisal. Between 1989 and 1990, Terry Cook began theorizing about the appraisal of archival information with a new perspective without giving it a specific name. In the publication *The Archival Appraisal of Records Containing Personal Information: A RAMP study with guidelines*, presented to the International Council on Archives, Cook suggested that archivists should shift the appraisal focus from explaining

the reason for the disposal to justifying the number of records to be kept. According to Cook, as documents result from transactions involving people and a context, they reflect an image of society. An aggregate of personal documents should be kept if it refuted the “official” image created by an institution so that it could provide evidence to review the “official” picture. A preservation scheme should be created once these evidence values have been verified. This perspective, which reviews man’s interaction with structures to capture dissonant voices, became the embryo of the Macro-Appraisal theory (Cook, 1991).

For Cook, a conventional appraisal was doomed to perform a microanalysis of the document or organizational structure. According to Trace (2016), Cook also rejected the idea that historical research producers and trends should be responsible for assigning some value to the document. With some influence from pioneers Gerald Ham and Hans Boom, Cook placed social values at the center of appraisal.

In a single summary sentence, the macro-appraisal assesses the social value of the structural-functional context and culture of the workplace in which records are created and used by their creator(s) and the interrelationship of citizens, groups, organizations – “the public” – with this structural-functional context. Suppose the appraisal assigns the long-term value of the record content – or series of records’ content – to their possible research values. In that case, the macro-appraisal will assess the significance of their creation and contemporary use context. A conventional appraisal is about records; the macro-appraisal is about its broader context (or “macro”) (Cook, 2005, p. 102).

In documents referring to the Macro-Appraisal of Documents in the Canadian national archive system, drafted by Terry Cook and approved by Library and Archives Canada, the Macro-Appraisal is described as based on the context and contemporary use provided by social values, those which, in turn, are only understood in light of the interaction between administrative structures, business functions, and citizens.

While no one can *objectively* know or state with complete assurance what the elements of societal value(s) are or have been within any given generation, archivists can develop appraisal strategies and methodologies that are most likely to provide a comprehensive documentary memory of what has transpired in society over time (Library And Archives Canada, 2001, p.4)

In summary, the idea is to research functional profiles and governmental structures and verify their impact on individuals so that it can document how the government operates and affects Canadian society. It is not intended to provide archival evidence as an end in itself (Library And Archives Canada, 2001).

According to Cook (1991, 2005), document appraisal is a co-creation of files, the exercise of power over memory; therefore, it requires an interventionist character from the archival professional, capable of identifying generic points of conflict and intersections between producers, socio-historical models, and clients or citizens in the structure. Given its importance, Macro-Appraisal should not be a routine procedure but a detailed work of archival, diplomatic, and historical analysis of functionalities, structures, values, systems, flows, and technologies and their presence over time. Therefore, macro-appraisal constitutes

an intellectual stage before micro-appraisal, which is the operationalization of the assessment itself (or traditional assessment, based on primary and secondary values).

In short, the main trends of Canadian postmodern approaches to appraisal are, according to Couture (2005, p. 100):

- Those proper to the National Archives of Canada: macro-appraisal, based on the appraisal of the institution (the context of creation) that precedes the appraisal of its documents (from the top down);
- Those that support the use of archives as a primary consideration for appraising documents;
- Those that support the importance of providing evidence of the whole of society; and
- Those proper to Québec insist on the essential role of a retention schedule closely and necessarily tied to a needs assessment.

In conclusion, new discussions on Document Appraisal have become part of the archival discourse and the international scenario from the postmodern perspective, involving accountability mechanisms that need to be incorporated into document management and those related to social memory. Such an arduous task, which requires knowledge of social dynamics, can benefit from using digital technologies such as the DéjàVu project, which we will see below.

5. THE DÉJÀVU PROJECT

As the scope of digital humanities and the Appraisal of Documents concept have been outlined, we can return to the purpose of this article, that is, analyzing how the DéjàVu solution can be a technological increment to the appraisal function to function as a digital support.

The project “DéjàVu: Temporal, spatial, and characterization coherence of heterogeneous data for integrity analysis and interpretation,” here called DéjàVu solution, was created by Professor Anderson Rocha, from the Instituto de Computação da Unicamp. Inspired by similar projects from international universities, the idea behind DéjàVu is to retell what has already happened, therefore reconstructing visual and textual records linked to an event to interpret it better. The project works by mining content from different sources – on and off the web. This information from textual and audiovisual pieces is captured, synchronized, and organized around an event, place, or object. Finally, the organization of this data can provide a denser understanding of the object or fact, the order of events, and its relationship with people and other objects, and offer subsidies to checking facts and other purposes, such as the appraisal of documents (Ferreira, 2018). Figure 1 provides an overview of the project.

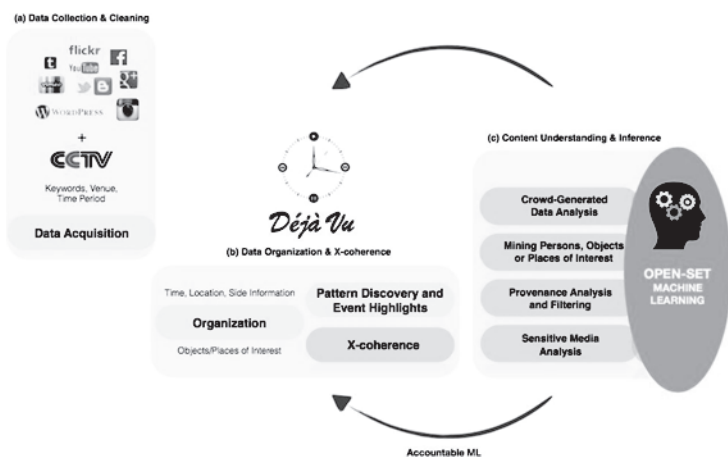


Figure 1. DéjàVu’s project overview. Source: Rocha; Pedrini & Carvalho (2021, online).

Therefore, the structure of the DéjàVu project is divided into three main and interrelated strands, namely: A) Data Collection and Cleaning; B) Data Organization and Coherence X (data synchronization); and C) Understanding and Inference regarding the content. The data collection and cleaning strand, according to Rocha, Pedrini, and Carvalho (2021), refers to the acquisition of data itself related to a specific location, object, or event, which has been posted in various places (social networks, video security cameras, information collections) or that have been seized in operation or grouped for any other purpose.

The strand of data organization and coherence X (synchronization of events) consists of putting together the pieces of information on the previously chosen topic based on their spatial (geographical) and temporal (position in time) coherence, called Coherence X (being X, the topic to be analyzed). This step includes analyzing how to connect different information related to the event in the physical and virtual environment; how to synchronize the pieces in the temporal aspect; how to organize the pieces about the event (before, during, and after)

With pieces consistent with topic X, we can: (i) look for specific clues (authorship, related topics); (ii) mine possible suspects, objects, or places of interest through biometrics and machine learning techniques; (iii) fake combat news and media reuse through provenance analysis methods; and (iv) evaluate sensitive content through various filters (e.g., violence), all based on open-set recognition techniques (Rocha, Pedrini & Carvalho, 2021, online).

This organized information and inferences will be used in strand C) Content understanding and inference. Therefore, in strand C) that we find scope for the Appraisal of Documents in an organization, as the correlations allowed by the software from a large quantity of data provide a substratum for evaluative decision-making.

Marino, Silveira, and Nicodemo (2020) highlight, for example, the use of the DéjàVu project together with the Centro de Humanidades Digitais IFCH-UNICAMP for the development of the Memória COVID-19 Brasil project, in which they corroborate to understand the tool affinity for Writing of Digital History around the global synchrony event of the COVID-19 Pandemic; mapping the circulation of publications, including false news or fake news; and the development of a relevance detection algorithm to help historians and archivists work with digital documents in large volumes (Big Data).

The aim is to establish phylogenetic relationships within a database of publications based on crossing both the metadata of these records, such as authorship and date of publication, and the semantic analysis of the publication content. Using an identification tool specific to social networks and NoSQL data processing (Non-Structured Query Language), the aim is to map the circulation of news and rumors about the pandemic, making it possible to know when and how certain contents circulated, what reactions they engendered among users, and which audiences were most affected. As stated above, this tool will make preserving what should not be remembered possible but is no less important (Marino, Silveira & Nicodemo, 2020, online).

We believe in the software's potential not for writing history based on archival selections but as a support for selecting documentary sets based on different criteria. The main hypothesis we advocate is that to appraise documents, an organization needs to be equipped with an analysis of events, objects, and places of interest; and that their understanding can be maximized by software, such as that developed in the DéjàVu Project. This is because from the mining of varied data, inside and outside the organization, produced organically by other entities and correlated with data collected via crowdsourcing, it will be possible to understand the social context (in the form of behavior patterns, trends, and possible correlations) to make inferences about values than (in the case of Macro-appraisal, about social or macro-appraisal values, and in the case of micro-appraisal or traditional appraisal, about primary values – administrative, juridical, and legal – and secondary values – probative and informative) about what documents this organization should keep and for how long. However, it is not defended that software will replace a thorough human analysis; the solution is complementary to the traditional appraisal function and should be used to facilitate the process. Nor is the interventionist character of the archival professional working with the solution excluded, who can identify the generic points of conflict and exceptional interest in the structure.

It should be noted that the context becomes an essential element for Archival Science, an importance that extends to the Appraisal of Documents. Context can be understood as

[...] the environment in which the action recorded in the document takes place. In analyzing the context of an archival document, the focus ceases to

be the document itself and begins to encompass the entire structure surrounding it, that is, its documental, legal-administrative, procedural, provenance, and technological context. (Conarq, 2020, p.19)

We would also add to the above concept the relational context of the documents with the citizens who use the service, which is of special attention when dealing with public archives, as well as the sociocultural context in which the document transactions occur, implying power, customs, and interest relations that form the social fabric. Thus, the advantage provided by the DéjàVu Project is mainly presented regarding the digital tabulation of information that helps to understand the context in which certain activities and events are documented; and that, if collected manually, would take much more time and resources to be carried out. Therefore, it is also inferred that the DéjàVu Project can be especially beneficial in social analysis for the attribution of social values, which is a typical function of Macro-appraisal.

This is because the DéjàVu Project collects/mines the greatest amount of data on an event, given the parameterized sources for each event, to reconstruct it, check it, and retell it from various media so that it can, among other things, provide a basis for the archivist appraisal. The appraisal with this data tends to be more precise and suppress gaps in understanding about the phenomenon, clarify ambiguities, and make the impacts of events on citizens' lives more noticeable.

Let us take as a model the example of Rocha (2018), about the fire that devastated the Grenfell Building, in London, on June 13, 2017, for two days, resulting in 71 deaths at the time. The results of the survey of security cameras,

police vehicles, records posted online by residents, and other forms of coverage mined by the Project would facilitate police work and forensic work in understanding how the fire spread in a timeline, how each death followed, and whether there would be a need to investigate suspects concurrently with the investigations. Additionally, all the documentation derived from the episode and accumulated by public security agencies later would need to be evaluated, and for the appraisal to take place, the set of documents related to the event would need to be contextualized. Therefore, the DéjàVu solution would come as revisiting the media that portrayed the event, providing an overview of the event, posterior and diachronic to the time it happened. This panorama of the event, before, during, and after it, answers the questions “how, where, when, and who” was involved; that is, it makes it possible to understand the situation, which is crucial to the appraisal. Additionally, the algorithms to be developed by the system also allow the creation of relevant criteria to aid in selecting materials.

Thus, from a primary value assessment perspective, the set of administrative, juridical, and legal documents related to the event and the functions of the security agency in the anecdotal example would be observed; from a secondary value perspective, the probative and informative potential of the accumulated documents would be observed; and in the view of a macro-appraisal, the social values related to the importance of the building, considering the social commotion and the impact of the event for Londoners, for the history of civil construction and public safety, etc. All these aspects could be aligned with the media repertoire accumulated by the DéjàVu Project so that the attribution of values takes place with adequate density and informational precision.

As advocated in this study, the DéjàVu Project can potentially consider the phenomenon's impact on the organization and the informational density necessary to record the analyzed phenomena that should (or should not) be kept for posterity. Therefore, it is technically and epistemologically aligned with the needs of a postmodern Appraisal of Documents. By extension, it is also inserted in the Digital Humanities perspective since the purpose of Appraisal of Documents is, ultimately, humanistic, being linked to the creation of heritage and social memory and the strengthening of community identity. There is also a connection with Digital Humanities based on the assumption of the Digital Humanities Research Group (2021) that it would be provoking the relationship between traditional practices (in this case, appraisal) and new digital technologies, the latter as facilitators of conventional procedures.

Additionally, the same organization could use the fact-checking and possible false news appraisal process, as well as the original analysis of the materials it wants to incorporate into its collection, which is also functionalities made possible by the DéjàVu solution, but which we will not delve into in this study.

6. FINAL CONSIDERATIONS

The appraisal of documents has become one of the archival functions that comprise an issue to be resignified by Postmodern Archival Science. Not only because it is the function that allows selecting what will constitute the informational asset of an institution aiming at the best use of financial resources for its management and maintenance (which, in itself, is already a great responsibility that falls to the archivist), but also because the social aspect of this function started to be studied with greater focus by some

scholars. In other words, the appraisal of documents has been analyzed as a function that defines, on the one hand, which evidence of groups, organizations, and structures will represent the informational heritage of an organization or nation for posterity, and, on the other hand, which evidence will be relegated to oblivion.

To dive deeper into these discussions, it is assumed that digital technologies with a humanistic background can facilitate the appraisal and determination process of macro and micro values based on a better understanding of the social context. Among these technologies, this research sought to study the DéjàVu Project case. The digital solution developed by the DéjàVu Project refers to the combination of digital data collection and cleaning activities, followed by their organization and synchronization for later understanding and inference in various areas. Within the scope of an organization that seeks to appraise its documents, Project data mining is carried out internally and externally to the organization, providing administrative and social data that allow apprehending the context related to the production, processing, and use of documents as well as the impact of the activities recorded in the society in which the organization is located. Analyzing a set-in relation to a producing event, for example, answers the questions “how, where, when, and who” was involved, allowing disambiguation and filling in gaps that may exist after simple textual analysis.

Therefore, it is assumed that the solution can make the assessment more precise by suppressing gaps in understanding about the phenomena that generate the documents and clarifying uncertainties, as well as making the impacts of events in the lives of citizens more noticeable, contributing, to a second moment, so that the remaining

documents from the appraisal form a documentary heritage worthy of documenting the social memory, of documenting how organizations operate, and representing the identity of either that organization or, in the case of documents from public bodies, of society itself.

Despite the DéjàVu Project facilitating the apprehension of the context for the attribution of values, its primary function overall is to help define social or macro-value values, those typical of Macro-Appraisal and which refer to social criteria by which specific sets of documents must be maintained. This is because the Project comprises data mining that goes beyond the activities of the organization itself produced by external entities in various media, with diverse and plural interpretations of the events/phenomena that originated documents, which would allow covering more extraordinary nuances of the community and the maintenance of a heterogeneous documentary heritage, which would serve the creation, communication, and sharing of knowledge in the future – to the detriment, for example, of protecting only “the producer’s version” or “an official version”. Social analysis carried out traditionally would involve more time and resources than the time and resources employed when using the software. Even so, it might not have the same success in collecting and tabulating heterogeneous social information.

Finally, it appears that the solution developed in the DéjàVu Project can be complementary to the appraisal of documents, providing a context substrate to allow a more precise, plural, and representative appraisal of social demands.

REFERENCES

- CAPURRO, R., & HJORLAND, B. (2007). O conceito de informação. *Perspectivas em Ciência da Informação*, 12(1), 148–207. <https://doi.org/10.1590/s1413-99362007000100012>
- CASTELLS, M. (2016). *A Sociedade em Rede* (17a ed.). São Paulo: Paz e Terra.
- CONARQ. (2020) *Glossário da Câmara Técnica de Documentos Eletrônicos*. Available at https://www.gov.br/conarq/pt-br/assuntos/camaras-tecnicas-setoriais-inativas/camara-tecnica-de-documentos-eletronicos-ctde/glosctde_2020_08_07.pdf
- COOK, T. (2005). Macroappraisal in Theory and Practice: Origins, Characteristics, and Implementation in Canada, 1950–2000. *Archival Science*, 5, 101–161.
- COOK, T. (1991). *The Archival Appraisal of Records Containing Personal Information: A RAMP Study with Guidelines*. Paris: Unesco General Information Programme and UNISIST.
- COOK, T. (1997). What is Past is Prologue: A History of Archival Ideas Since 1898, and the Future Paradigm Shift. *Archivaria*, 43, 17–63.
- COUTURE, C. (2005). Archival Appraisal: A Status Report. *Archivaria*, 59, 83–107.
- EASTWOOD, T. & MACNEIL, H. (2016). *Correntes atuais do pensamento arquivístico*. Belo Horizonte: Editora UFMG, 2016
- FERREIRA, A. (2018). *DéjàVu Post-doctoral Fellowship*. Available at <https://recodbr.wordpress.com/2018/01/08/dejavu-post-doctoral-fellowship/>
- GALINA-RUSSELL, I. (2011) ¿Qué son las Humanidades Digitales? *Revista Digital Universitaria*, 12 (7).
- GRUPO DE PESQUISA EM HUMANIDADES DIGITAIS. (2021). *O que são Humanidades Digitais*. Available at <https://humanidadesdigitais.org/o-que-sao-humanidades-digitais/>
- LAKATOS, E. M. & MARCONI, M. A. (2003). *Fundamentos de metodologia científica* (5ª ed.). São Paulo: Atlas, 2003.
- LIBRARY AND ARCHIVES CANADA. (2001). *Appraisal Methodology: MacroAppraisal and Functional Analysis Part A: Concepts and Theory*. Available at <https://www.bac-lac.gc.ca/eng/services/government-information-resources/disposition/Documents/MacroappraisalPartA.pdf>

- MARINO, I.K., SILVEIRA, P.T. & NICODEMO, T.L. (2020) Arquivo, Memória E Big Data: Uma Proposta A Partir Da Covid-19. *Cadernos do Tempo Presente*, 11(01), 90-103.
- MINAYO, M.C.S. (1994). *Pesquisa social: teoria, método e criatividade*. Petrópolis, RJ:Vozes.
- ROCHA, A., PEDRINI, H. & CARVALHO, A.C.P.L. (2021). *DéjàVu: Feature-Space-Time Coherence from Heterogeneous Data for Media Integrity and Interpretation of Events*. https://www.ic.unicamp.br/~dejavu/#data_aquisition
- ROCHA, A.R. (2018). *Projeto DéjàVu*. Available at <http://www.jornal.ecuesta.com.br/conteudo/tecnologia/projeto-dejavu.htm>
- RODRÍGUEZ-YUNTA, L. (2013). Humanidades digitales, ¿una mera etiqueta o un campo por el que deben apostar las ciencias de la documentación? *Anuario ThinkEPI*, 7, 37– 43.
- ROUSSEAU, J. & COUTURE, C. (198). *Os fundamentos da disciplina arquivística*. Lisboa: Publicações Dom Quixote.
- SCHELLENBERG, T. R. (1973). *Arquivos Modernos: princípios e técnicas*. Rio de Janeiro: Editora FGV.
- SOARES, A. P. A., PINTO, A. L. & SILVA, A. M. (2015). O paradigma pós-custodial na rquivística. *Páginas a&b*, 3(4), 22-39.
- TOGNOLI, N. B. (2013). A representação na arquivística contemporânea. *Revista Ibero- Americana de Ciência da Informação*, 5(2), 79-92.
- TRACE, C. B. (2016). Dentro ou fora do documento? Noções de valor arquivístico. In Eastwood, T. & Macneil, H. *Correntes atuais do pensamento arquivístico* (pp.77-106). Belo Horizonte: Editora UFMG.
- UNESCO. (2003). *Carta para la preservación del patrimonio digital*. Available at http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/charter_preservation_digital_heritage_es.pdf