



# Social justice in the web archives: considerations to (re)think the documental appraisal

**Jonas Ferrigolo Melo**

PhD Candidate in Information and Communication in Digital Platforms, Universidade do Porto (UP), Porto, Porto, Portugal.

ORCID: 0000-0002-7312-3509

<http://lattes.cnpq.br/0594057350445795>

[jonasferrigolo@gmail.com](mailto:jonasferrigolo@gmail.com)



**Moisés Rockembach**

PhD in Information and Communication, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil.

ORCID: 0000-0001-9057-0602

<http://lattes.cnpq.br/1304688580274983>

[moises.rockembach@gmail.com](mailto:moises.rockembach@gmail.com)

Submitted on: 26/05/2023. Approved on: 28/02/2024. Published on: dd/mm/yyyy.

## **ABSTRACT:**

This paper examines the archival appraisal process in web archives through the lens of social inclusion, aiming to understand how archival theory can contribute to building web archives that promote social justice. By analyzing archival literature and empirical cases of ten web archiving initiatives, this research identifies elements that suggest ways to construct web archives aligned with social justice principles. The study emphasizes the importance of reconsidering web content evaluation processes in light of diverse cultural contexts and systemic inequalities that can lead to the permanent loss of historical records. It argues that web archiving requires a thoughtful consideration of the cultural dimensions influencing preservation decisions. This study contributes to the understanding of evaluation in web archiving and serves as a foundation for future research in the field.

**Keywords:** web archiving; web archive; archival appraisal; social justice.

## INTRODUCTION

Access to and use of digital resources on the internet often prove ineffective due to rapid technological evolution, creators' inability to maintain their websites for extended periods, and in different country's due to legal challenges and restrains. As *Tim Berners-Lee, the inventor of the web*, stated, "não há nenhuma razão em teoria para as pessoas mudarem de URLs, mas milhões de razões na prática" (Berners-Lee, 2013)<sup>1</sup>. Consequently, the need arises to study digital preservation strategies, such as web archiving, through various approaches and activities. These include selecting, preserving, archiving, and accessing websites deemed to have academic, cultural, and institutional value.

Web archiving has been examined as a primary source (Brügger, 2011; 2012; Milligan, 2018; Rodrigues; Rockembach, 2021; Rogers; Brügger; Milligan, 2018) because its content can be considered virtual cultural heritage or an information resource for research. Capturing positive experiences and ways of living has the potential to become part of the social and cultural memory to be preserved and recognized in the future. The United Nations Educational, Scientific and Cultural Organization (UNESCO) acknowledges that websites containing cultural heritage values are part of tangible cultural heritage (UNESCO, 2021).

Web archives are widely regarded as essential resources for preserving access to websites, social media, and other forms of web communication for future use and as fundamental sources for research in Digital Humanities (Brügger, 2018; Fritz et al., 2021; Milligan, 2019; Rockembach, 2019; Weber, 2020). Additionally, there are ongoing discussions exploring methodological approaches to web archiving (Bragg *et al.*, 2013; Khan; Rahman, 2019; Németh; Drótos, 2019), as well as its ethical aspects (Glanville, 2010; Mackinnon, 2021; Rockembach, 2017).

Web archiving has spread globally and is carried out by various organizations, including archives, libraries, universities, and companies<sup>2</sup>. The Internet Archive, founded by Brewster Kahle in 1996, is considered the world's first web archive and operates as a nonprofit organization with the mission of enabling "universal access to all knowledge" (Internet Archive, 2023, online). In Brazil, web preservation began to be intensely debated in 2015 with the introduction of a bill in the National Congress. By 2017, it had become part of research group discussions, and in 2018, the topic gained further momentum (Ferreira; Martins; Rockembach, 2018; Boeres; Saad, 2023; Rockembach, 2018a; 2018b; 2018c; Rockembach; Ferreira, 2018; Rockembach; Pavão, 2018). In 2020, the focus extended beyond academia, and in 2022 the National Council on Archives (CONARQ on the Portuguese acronym) approved the Digital Preservation Policy for Websites and Social Media (Melo; Oliveira; Rockembach, 2023).

The introduction of a web preservation policy opens up the possibility for Brazil to foster a culture of systematic preservation of these sources of information and social memory. Through this, it aims to, "[...] alcançar um senso de comunidade, identidade nacional e enraizamento

---

1 Translation: "there is no reason in theory why people should change URLs, but millions of reasons in practice" (Berners-Lee, 2013, editorial translation)

2 Website: [https://en.wikipedia.org/wiki/List\\_of\\_Web\\_archiving\\_initiatives](https://en.wikipedia.org/wiki/List_of_Web_archiving_initiatives).

entre os cidadãos brasileiros, no sentido em que se estará preservando informações que, de certa forma, moldam a identidade nacional [...]” (Melo, 2020, p. 21).<sup>3</sup> However, legislation alone will not suffice without accompanying research into practical solutions for the effective implementation of web archives (Melo; Oliveira; Rockembach, 2023), as well as debates on theoretical issues such as the selection process.

Although issues involving the power and politics of document selection and appraisal have historically concerned the archivist profession for decades, these observations point to a community generally resistant to discussions about “selectivity” (Ogden, 2022). The work of selection requires many difficult decisions for professionals, highlighting the need for clear and objective standards and techniques. Appraisal and selection in archives are among their main activities, as collections have the capacity to encompass intersections of all identities, including race, economic status, gender and sexuality, religion, and politics. Consequently, archivists can work to increase the representation of previously unrepresented or underrepresented people in the historiographic record. Such pluralistic approaches to archival practice have a long history, dating back to early efforts to create more inclusive archives (Booms; Joldersma; Klumpenhauer, 1987; Cook, 1992; Eastwood, 2002; Flinn; Stevens, 2009; Flinn; Stevens; Shepherd, 2009; Harris, 2007; Zinn, 1977).

In this context, social considerations are essential when forming collections and documentary archives. This includes addressing the privacy concerns of vulnerable individuals or those who have experienced trauma and may not want their stories archived. Additionally, it is crucial to respect diverse cultural beliefs and traditions regarding access, handling, and custodianship of materials. Another significant factor pertains to members of the LGBTQIAP+ community, who may wish to be represented by their chosen names and define their own pronouns. Hence, archives and other heritage institutions must be sensitive to the social issues inherent in the composition of these memory preservation instruments.

The concept of social inclusion extends beyond the actors previously mentioned, encompassing complex ideas associated with modern political thought and interconnected with poverty, deprivation, and inequality. ‘Social inclusion’ means individuals can fulfill their social roles (Levitas, 2004) and involves actions that integrate marginalized groups, such as homosexual and transgender people, indigenous, native and Black individuals, people with disabilities, or low-income populations, into the social fabric. These groups have historically experienced exclusion from socialization processes, which continues to this day. In this context, actions promoting inclusion and representation aim to provide equal opportunities regardless of socially imposed conditions. This social exclusion resonates within formal records that fulfill institutional missions and ultimately constitute historical and permanent archives.

Archivists must consciously incorporate diversity and inclusion in all facets of their work. This includes actively considering how materials are organized, described, and preserved. Given their unique responsibilities to society, archivists and archival professionals must

---

3 Translation: “[...] achieve a sense of community, national identity, and rootedness among Brazilian citizens, as it will be preserving information that, in some way, shapes national identity[...].” (Melo, 2020, p. 21, editorial translation).

recognize that their purpose, strategies, and technologies impact both the selection of materials for archiving and their subsequent accessibility. Keeping these aspects in mind is crucial when planning and executing archival activities.

This paper examines web archiving through the lens of the subjective nature inherent in the concept of social inclusion. Drawing inspiration from Harris's (2002) work on the relationship between document appraisal, power, and memory, we argue that web archiving necessitates an awareness among archivists of the cultural dimensions that shape decisions regarding how the web is preserved. In this context, we pose the question: **How can the literature on document appraisal contribute to the creation of web archives that promote social justice?** This paper aims to identify, within archival literature and empirical case studies, elements that suggest pathways for composing web archives grounded in the principles of social justice.

## METHODOLOGY

To address the need for answers regarding website selection criteria for web archives, this article initiates a dialogue to expand the discussion on potential document evaluation criteria in this context. To this end, it introduces methodological studies on evaluation in archives from the perspective of social inclusion, despite these theories sometimes being overlooked due to the corporate/administrative bias inherent in established document management practices.

We posit that when debates center on a limited technique, there is a risk of this technique eclipsing the preservation of contextualized information. This narrow focus can lead to archives inadvertently promoting the erasure of populations that are structurally marginalized or invisible in society. Therefore, the general objective of this article is to identify, within archival literature and empirical case studies, elements that suggest pathways for composing web archives grounded in the principles of social justice.

To achieve this general objective, the research methodology comprised two phases. The first phase employed a bibliographic research approach, utilizing the Scopus, Web of Science, and Dimensions databases to search archival literature for evaluation methods considering social issues as criteria. The articles were organized to facilitate the extraction of relevant information, theories, authors, and ideas aligned with this study's aims. Research topics included document appraisal with a focus on social justice and web archives encompassing socially excluded populations.

In the second phase, a secondary research approach (also known as documentary research) was employed. This involved summarizing, grouping, and synthesizing documentation related to existing web archiving projects with a social justice bias. Several web archiving initiatives were selected and analyzed. The selection criteria for our sample of web archiving initiatives included: (i) established web archiving initiatives; (ii) geographical distribution; and (iii) different biases within the scope of social inclusion. This approach was taken to study a representative sample of web archiving initiatives. Ten initiatives were selected.

The first study analyzed the DocNow project, an initiative focused on archiving web content related to social events and movements. The second study focused on the Community Webs platform, which aims to archive and preserve web content from local communities. The third case examined was the Web Archiving at Stanford platform, dedicated to preserving and making available the web content of academic institutions and related to politics and electoral campaigns. The fourth study examined MFigure, a Malaysian web archiving repository focused on social and cultural heritage. The fifth study observed the Huellas Incómodas platform, which encompasses issues related to feminism and gender equality. The sixth case explored the Archiving the Black Web (ATBW) project, which seeks to preserve and make accessible online records of the Black experience. The seventh case is the #SchomburgSyllabus, a web archive composed of educational resources related to Black studies, social movements, and experiences. The eighth study was dedicated to archiving LGBTQIAP+ web content in the Russian and Eastern European contexts. The ninth study was the University of British Columbia (UBC) archive for the study of Indigenous peoples, titled First Nations and Indigenous Community Websites. Finally, the tenth case is the collections of the largest government web archive in the world, the UK Web Archive (UKWA), which offers dozens of collections with a social justice promotion bias.

These case studies provided valuable insights into the role of web archiving in safeguarding cultural diversity and promoting social inclusion for groups historically marginalized in archives. By examining the intersections between archives, document appraisal, and social justice, this article underscores the importance of web archiving as a vital tool for documenting the voices and experiences of diverse communities, as well as for fostering more equitable access to the knowledge and history of these groups.

## WEB ARCHIVING

Between the 1980s and 1990s, the first movements emerged with the goal of developing strategies aligned with the rise of digital technologies, specifically aimed at enhancing the capacity of institutions devoted to memory conservation to capture and preserve digital artifacts as records of social phenomena (Schneider & Foot, 2008). However, the formal practice of web archiving dates back to the late 1990s, with pioneering initiatives launched in 1996 by the Internet Archive and the National Libraries of Australia and Sweden (Schroeder; Brügger, 2017). France also played a pivotal role, with the National Library of France (BnF) conducting its first web archiving experiments in 1999 (BNF, 2017).

As web archiving is an emerging field, precise and clear definitions can be elusive. Consequently, the term “web archiving” is often used interchangeably with “web preservation.” However, there is a general consensus that web archives are innovative systems designed to acquire, store, and preserve information published on the internet (Brown, 2006). Akin to traditional memory institutions, web archives also serve as valuable sources for research and materialize spaces for safeguarding information.

Rockembach (2018a, p. 9)<sup>4</sup> defines web archiving “[...] como um processo que compreende coletar, armazenar e disponibilizar a informação retrospectiva da *World Wide Web* para futuros pesquisadores”. The International Internet Preservation Consortium (IIPC), defines it as “[...] “[...] o processo de coletar porções da *World Wide Web*, preservando as coleções em um formato de arquivo e, em seguida, servindo os arquivos para acesso e uso” (IIPC, 2017, *online*)<sup>5</sup>.

Given this long-term perspective, web archiving requires a strategic approach regarding technologies, systems, policies, procedures, and resources to make it more than just the act of collecting and storing online content. In this sense, legal issues have significant implications for web archiving, as they influence policies on selection, preservation, and user access to archived content. Therefore, debates on the composition of a web archive must consider aspects related to copyright legislation, personal data protection, and the issue of illegal content that violates public policy.

As the web has evolved into a medium of communication, it has come to represent a vast collection of primary sources about our past. This wealth of diverse information provides the necessary conditions for the emergence of web archiving as a truly interdisciplinary field, bringing together professionals and academics from various areas: humanities, social sciences, computer and information sciences, libraries, archives, etc. (Ogden, Halford; Carr, 2017). Although in Brazil, actions related to web archiving are more centered in the field of information science, with some insertions in communication, technology, and cultural heritage (Melo, Oliveira; Rockembach, 2023), other fields will eventually join the debate, forming an interdisciplinary network at the national level, as occurs in other countries.

Given the importance of legislative, technical, and political elements in the creation of a web archive, it is crucial to provide adequate information and documentation about this context to users, effectively opening the “black box.” A notable example is the Portuguese web archive (arquivo.pt), which promotes transparency through videos that illuminate the archive’s internal workings. This is essential from the user perspective, as providing such elements aids in searching for and potentially archiving web pages.

Therefore, recognizing the web as a rich source of information, its archiving has become an interdisciplinary field that brings together specialists from diverse areas. In Brazil, it is hoped that debates will expand beyond Archival Science and Information Science, incorporating web archives into the broader discussions within these fields.

## DOCUMENT APPRAISAL

Document appraisal is commonly understood as the process by which archivists identify materials possessing significant enduring value, thus justifying their preservation

---

4 Translation: “[...] as a process that involves collecting, storing, and providing retrospective information from the World Wide Web for future researchers.” (Rockembach, 2018a, p. 9, editorial translation).

5 Translation: “[...] collecting portions of the World Wide Web, preserving the collections in an archival format, and then serving the archives for access and use” (IIPC, 2017, *online*, editorial translation).

within an archive (Pearce-Moses, 2005). While the values ascribed can vary due to the inherent subjectivity of this process, appraisal is typically guided by techniques grounded in research, documents, and legislation. This is a core activity in archival work, as the outcomes of appraisal decisions shape the historical record and, consequently, our understanding of the past and our collective memory (Jacobsen, Punzalan, Hedstrom, 2013).

Eric Ketelaar (2001) introduced the term “archivalization” to describe the process of determining what materials merit preservation. Appraisal choices, he argues, reflect the prevailing organizational logic and the perspectives of those assessing the documents. The outcomes of this process can shift depending on the relationship between the appraiser and the community, subject, or class being evaluated. Ketelaar suggests that archivalization precedes the actual act of archiving, and as such, it serves as a mirror reflecting the world prior to the archiving process.

Similarly, Verne Harris (2002) likens the appraisal process to a “sliver of a window.” He contends that many archivists perceive their work as constructing a faithful representation of “reality” from the evaluated documents, much like assembling a puzzle. However, few archivists actively participate in shaping the “reality” that the archival record supposedly “reflects.” Consequently, through the identification of records with archival value, appraisers—operating from an external perspective – create the value of the archive. This renders the archival record a fragmented glimpse through a window.

Over the past century, the archival community has formulated diverse appraisal approaches that guide the selection process. These approaches have evolved partly in response to the contingencies of their respective eras, including the advent of writing, printing, photography, and digital technologies, which have exponentially increased content generation (Cox, 2009). Various approaches considered peripheral to traditional archival science have emerged over time. For example, sampling techniques offer a systematic method to select records from large, inaccessible sets (Cook, 1991; Nye, 2009; Kolish, 1994). Archival institutions have also adopted collection development policies, which communicate the types of materials of interest to the institution. This technique is seldom used in Brazil, as it is considered an artificial strategy for composing an archive, countering the principle of organicity and provenance<sup>6</sup>. In contrast, it is widely used in the United States, although many archives lack a collection development policy (Sauer, 2001).

While appraisal discourses in archives are sometimes oversimplified due to misinterpretations, it is crucial to acknowledge that these marginal perspectives offer fresh opportunities and challenges for professionals. They are grounded in decades of post-custodial thought within the field, which has challenged conventional models of archival ownership (Bearman, 1991; Ham, 1981; Ribeiro, 2002; Silva, 2006; Silva; Ribeiro, 2011, 2012, 2020). The development of novel appraisal ideas, such as documentation strategy (Samuels, 1986),

---

6 In this regard, Emily Maemura *et al.* (2018) explore the provenance of web archives from the perspective of document appraisal, advocating for an approach that ensures researchers can understand the conditions under which a web archive was created.

macro-appraisal and functional analysis (Cook, 2004), and the information flow appraisal model (Ribeiro, 2005), was spurred by the need to adequately document society and its interactions with state institutions.

Documentation strategy is a collaborative methodology that guides the appraisal process by identifying archives to be preserved based on subject matter, document type, and geographical location (Nesmith, 2011; Samuels, 1998). Macro-appraisal and functional analysis provide frameworks for assigning value to documents based on their function and social impact. Conversely, the information flow appraisal model emphasizes measuring relevance, focusing on the informational acts and essential objectives of the document, its organic structure, and its competencies/functions.

The emergence of technology, particularly the opportunities afforded by Web 2.0, led some authors to propose more participatory archive models (Shilton; Srinivasan, 2007; Yakel, 2007), exploring modes of interaction between creators, users, and holders of archives. This social and cultural phenomenon of co-creation necessitates a departure from traditional paradigms of document organicity and archival practice.

Although appraisal is a complex topic within archival studies, it is a fundamental process that can be approached from various angles. While well-established appraisal practices exist in archives, the field is still in the early stages of investigating how archivists and web archiving tools can effectively conduct appraisal.

## **DOCUMENT EVALUATION AS A DRIVER FOR SOCIAL INCLUSION IN WEB ARCHIVES**

Before exploring document appraisal from this social inclusion perspective, it's crucial to define the term. Social inclusion gained prominence in the United Kingdom, and it wasn't until the early 2000s, with the adoption of the European Social Policy Agenda by heads of state and government, that comparable strategies to combat poverty and foster social inclusion were implemented. This agenda led to the concept's widespread adoption throughout Europe. It identified populations deemed at risk of social exclusion (the antithesis of social inclusion), including individuals with disabilities, refugees and asylum seekers, ethnic minorities, the elderly, single-parent families, women, the unemployed, and people with mental health issues.

Baumgartner and Burns (2014) offer recommendations for developing a tool to measure social inclusion. These recommendations also contribute to a more inclusive understanding of who might face social exclusion, thereby defining pathways to social inclusion. The authors summarize social inclusion as encompassing a sense of belonging to the community, the degree of participation in community and civic life, the autonomy to choose whether to participate, and the opportunities and barriers that affect participation. Without mechanisms supporting these tenets, social groups risk marginalization from democratic processes or feelings of isolation and lack of choice.

Another pertinent perspective on social inclusion frames it as an informational issue. While existing theories acknowledge the marginalization of certain groups due to economic,



social, and cultural factors, they haven't fully considered the role of information access. Those without adequate information are socially excluded, and the socially excluded may lack access to essential information sources or appropriate social capital (Caidi; Allard, 2005).

In 2013, the annual meeting of the *Society of American Archivists* and the pages of the *American Archivist* became the forum for a debate on the political nature of social justice work. Mark Greene (2013) initiated the discussion by arguing that archivists best serve their profession by maintaining neutrality. Jimerson (2013) countered, asserting that social justice is a personal, rather than professional, responsibility. These positions sparked a response from Michelle Caswell (2013), who contended that the very notion of "neutrality" is subjective and that such perspectives are constrained by the field's lack of diversity.

Debates surrounding neutrality in archives predate 2013. In the 1970s and 1980s, archivist Hans Booms and historian Howard Zinn raised concerns about the social ramifications of archival practices. Booms, Joldersma, and Klumpenhauer (1987) advocated for greater representation of society in public archives by including collections that reflected a wider range of experiences. Zinn, in his 1970 address to the Society of American Archivists, challenged the concept of archival neutrality. He questioned how archival practices could perpetuate the existing social order, favoring the powerful and wealthy while neglecting the "powerless and obscure" (Zinn, 1977). He urged archivists to enhance access to government documents and prioritize collecting records that documented the lives of ordinary people, not just the elite. The relationship between social inclusion and archives has been a subject of exploration among archivists and scholars for over fifty years. More recent studies have used the term "social justice" to frame similar investigations (Duff et al., 2013; Harris, 2002; Jimerson, 2007; 2009; Wallace, 2010).

The challenge of representing marginalized communities prompted shifts in archival collecting practices and inspired a thorough reassessment of foundational archival principles, which often prioritize state or corporate entities, along with elite and dominant cultures (Punzalan & Caswell, 2016). Since the 1990s, critical theory, especially post-structuralism and post-colonialism, has spurred efforts to challenge prevailing archival concepts. This critical lens aimed to broaden archival understanding and practice, leading to the recognition that a reconceptualization of the intellectual foundations of archival work was necessary to better serve historically underrepresented and marginalized communities.

An example of this reconceptualization is the ongoing discussion and expansion of the principle of provenance to liberate archival work from narrow interpretations of creation and ownership (Punzalan; Caswell, 2016). Tom Nesmith (1999, p. 146) advocates for a dynamic understanding of provenance, encompassing "the social and technical processes of inscription, transmission, contextualization, and interpretation of records, which account for their existence, characteristics, and ongoing history." He proposes that this concept consider not only the creators of documents but also the subjects represented within the records. Other authors have similarly addressed provenance by incorporating broader, community-based configurations (Bastian, 2003, 2006; Douglas, 2010; Hurley, 2005; Millar, 2002; Wurl, 2005).

In recent years, archival literature has amassed a significant body of work advocating for a participatory model of provenance definition. This model, informed by a plurality of voices, acknowledges social inclusion as fundamental to a democratic society (Cook; Harris, 2007; Jimerson, 2009; Schwartz, 2002). Through these and other contributions, archivists have begun to critically reflect on the social implications of their work. Archival activism has gained traction, influencing the development of alternative appraisal and selection models.

Aligned with this perspective, another branch of archival thought has developed theories about the role of archives in shaping and disseminating collective memory (Bastian, 2003; 2009; Caswell, 2010; Cook, 2013; Hedstrom, 2002; 2010; Jacobsen, Punzalan & Hedstrom, 2013; Josias, 2011; Punzalan, 2009). Recognizing that official narratives often overlook community perspectives, while official sources predominantly represent an elite viewpoint, various perspectives have emerged on recovering silenced voices and sustaining counter-narratives.

From this perspective arises the discourse of community archives, which aim to foster societal participation in archival endeavors. Community archives have formed around various identities, including ethnic, racial, or religious (Caswell, 2014; Daniel, 2010; Kaplan, 2000); gender and sexual orientation (Barriault, 2009); economic status (Flinn & Stevens, 2009); and geographic location (Flinn & Stevens, 2009). Some authors have observed a significant increase in community archive organizations since the early 2000s (Bastian & Alexander, 2009; Cook, 2013; Daniel, 2010; Flinn & Stevens, 2009; Flinn, Stevens & Shepherd, 2009; Mander, 2009).

Intertwined with these movements is the resurgence of debates advocating for more “activism” from the archivists (Jackman, 2012; Yaco; Hardy, 2012; 2013). These debates underscore the growing interest within the archival community in social justice (Caswell *et al.*, 2012; Cushing, 2014; Nelson Mandela Foundation, 2005; Greene, 2007; Harris, 2007; Jimerson, 2009; Jimerson, Harris; Wallace, 2008; Society of American Archivists, 2012a; 2012b; Wallace, 2010).

These trends and debates within the archival community mirror developments in various academic disciplines and professions responding to the pressing challenges of contemporary social justice struggles, such as globalization, the defense of democracy, militarism and fascism, poverty and inequality, the influence of large corporations, surveillance, environmental degradation, indigenous survival, female empowerment, LGBTphobia, and the survival of Black bodies, among others. While the phenomenon of social justice extends far beyond the realm of archives, it is imperative for the discipline to engage with this issue in a systematic and coherent manner.

Archives wield the power to create and perpetuate both justice and injustice through the technical decisions made within the field. Recognizing systemic inequalities and utilizing intellectual and physical resources to challenge and transform these structures of exclusion, marginalization, and domination should therefore guide actions in the archival realm. It is crucial to understand that this framework precedes the production and maintenance of

records, necessitating a proactive commitment to social justice. Social justice, in this context, requires archives and archivists to acknowledge these issues while demonstrating concrete connections between social justice, archives, and archival practice.

Duff *et al.* (2013) surveyed archival literature to identify areas where social justice is most evident: the inclusion of underrepresented and marginalized sectors of society; the reinterpretation and expansion of archival concepts; the development of community archives; archival education and training; and efforts to document human rights violations. Some tangible markers of social justice in archives include facilitating participation and access, recognizing the role of archives in promoting restorative justice through record preservation and protection, addressing exclusions, absences, and silences in archives as gaps to be filled, and actively resisting the exclusions and marginalizations perpetuated by archival institutions.

While much remains to be explored in the realm of classical archiving, the same holds true for web archiving. Despite recent studies on the technical infrastructures of web archiving, little is understood about how archivists interact with these systems. Technical interactions, such as the content selection process and the influence of capture tools on this process, remain significant questions in web archiving discussions within Information Science and Archival Studies. Some research has progressed in this direction, raising important questions about this archival activity (Ben-David; Amram, 2018; Summers, 2020; Summers; Punzalan, 2017). However, further discussion is needed to understand how decisions about what to archive shape the composition of a web archive.

Although the *Internet Archive* rejects the notion of archiving everything (Masanès, 2005), it is clear that content selection inherently involves choosing which records will be excluded from preservation and, consequently, forgotten (Harris, 2014). This process of forgetting is driven by necessity due to the impossibility of storing all records. Thus far, this process mirrors the traditional evaluation process in Archival Studies.

The advent of web archives prompted the development of tools, protocols, standards, collaborative networks, and specialized expertise. Since 1996, the *Internet Archive's* automated agents (web crawlers) have been tracking link by link on the web, archiving everything encountered along the way (Mohr *et al.*, 2004). Following this lead, organizations within the *International Internet Preservation Consortium* (IIPC) began building their own collections, either based on internet domain logic (Gomes, Freitas, Silva, 2006; Phillips, 2005) or specific sites chosen according to collection development policies (Schneider *et al.*, 2003). Regardless of the techniques employed, it's unlikely that the evaluation process can be conducted without the aid of specialized technologies and automated agents to retrieve the selected content for archiving (Summers; Punzalan, 2017).

Therefore, to establish a framework of roles, policies, and technologies for evaluating web content, it's essential to understand current evaluation practices. From this understanding, a theoretical foundation can be built upon critical archival studies to avoid merely replicating traditional practices without a nuanced analysis of web archive evaluation. Additionally, the

political dimension of the archive must be emphasized in discussions about web archiving evaluation, as the process of documenting selection decisions carries inherent political implications that must be addressed when considering how web content is chosen.

Political decisions also play a role in recognizing the web as an essential source for studying cultural and social phenomena of the last decades (Schneider & Foot, 2005). The rising interest in community archives, particularly regarding preserving previously erased voices, could serve as a starting point for debates on archival composition on the web. Community archives involve activities such as creation, collection, processing, curation, preservation, and dissemination of materials related to a specific community or subject.

In this way, these archives can help counteract social inequality by providing access to information and content that are often excluded from formal records. By preserving information from the digital realm, web archives can offer a deeper understanding of society and help document social struggles and efforts for change. Therefore, we will next explore some examples of web archives that can serve as references for studies on selection and evaluation.

## WEB ARCHIVES AS PROMOTERS OF SOCIAL JUSTICE - SOME CASES

*Documenting the Now*, also known as *DocNow*<sup>7</sup>, established in 2016, is a project dedicated to developing open-source tools and community-driven practices for the ethical collection, use, and preservation of publicly available web and social media content. This initiative arose in response to the pervasive use of social media as a means of documenting historically significant events. One of *DocNow*'s primary concerns is ensuring ethical practices when working with social media content, particularly regarding long-term collection and preservation. This involves respecting user intent and the rights of content creators, aligning with the policies of platforms like X (formerly *Twitter*). The project develops open-source tools that empower archivists, activists, and researchers to effectively engage with social media data.

In 2017, the *Community Webs*<sup>8</sup> program, an education, training, and service initiative funded by the Institute of Museum and Library Services and the Internet Archive, was launched. Its aim was to empower public libraries to create collections of historically significant web-published materials that document their local communities. With support from the Mellon Foundation, *Community Webs* expanded in 2020 to include additional public libraries across the United States. In 2021, the Internet Archive broadened the program's scope to accept applications from cultural heritage organizations in Canada and worldwide. Since its inception, *Community Webs* has attracted over 150 public libraries and other cultural heritage institutions. Collectively, these organizations have archived over 100 terabytes of web-based community heritage materials, encompassing collections on local citizens' lives, elections, local blogs,

---

7 *Site:* <http://www.docnow.io/>.

8 *Site:* <https://communitywebs.archive-it.org/>.

social justice movements, crises and disasters, the COVID-19 pandemic, and more. These collections represent marginalized voices and groups often absent from traditional historical records.

The *Stanford Web Archive Portal*<sup>9</sup> prioritizes archiving web content prone to rapid disappearance. This includes political candidates' campaign websites, which are typically accessible only during the election cycle; funded projects that may remove their websites after grant funding ceases, even if the project continues; and dissident political speeches, which may face government censorship (Ferreira, Martins, Rockembach, 2018).

The MFigure project, initiated in Malaysia, aims to develop a database and repository of web archives that serve as a storage platform with descriptive and preservation metadata related to the country's social and cultural heritage (Saiful Bahry *et al.*, 2022). The project focuses on preserving web-published information about notable Malaysian figures, including acclaimed athletes, successful entrepreneurs, prominent researchers across various fields, politicians, and other distinguished individuals. By collecting stories, speeches, event participation, daily activities, and thoughts of these public figures, the project recognizes their impact on future generations. Additionally, the project has developed a metadata framework that facilitates the expansion and dissemination of web content, with the goal of creating a virtual repository of social and cultural heritage. The web archiving process utilizes tools like *Htttrack* and *Conifer*, generating both a *WARC* file and an actual web document. To ensure the integrity of the information, both file types are deposited in the MFigure repository, resulting in two associated file types for each web archive content.

*Huellas Incómodas*<sup>10</sup> is a digital preservation initiative dedicated to documenting, contextualizing, and creating a public legacy of collective memory and the right to truth, grounded in feminist social struggles across Latin America. The project addresses the ephemeral nature of protest expressions in physical and digital public spaces, intervening against the urgency to erase them. This platform comprises digital collections of activism from various parts of Latin America. The initiative encourages public participation in preserving digital activism by allowing anyone to submit feminist websites for archiving.

ATBW<sup>11</sup> is an initiative striving to establish a more equitable and accessible web archiving practice to document the Black experience. The vast growth of the web and social media, combined with their widespread use by Black people, presents significant opportunities and responsibilities for collecting institutions interested in documenting Black life and experiences online. The initiative received a \$2.5 million grant from the Mellon Foundation<sup>12</sup> to develop a continuing education training program in web archiving for memory professionals who collect archival content documenting Black life, history, and culture. The goals of ATBW are multifaceted: to diversify and increase the number of professionals active

---

9 Site: <https://swap.stanford.edu/was/>.

10 Site: <https://idrku.org/huellasincomodas/webarchive>.

11 Site: <https://archivingtheblackweb.org/>.

12 Site: <https://www.mellon.org/>.

in web archiving, to expand collections focused on the Black experience, and to bring together scholars and archivists interested in deepening the understanding, collections, and studies on the use of digital communication technologies by Black people.

The *New York Public Library* maintains the *#SchomburgSyllabus*<sup>13</sup>, a web archive comprising educational resources related to Black studies, movements, and experiences. By connecting these materials to the collections of the Schomburg Center for Research in Black Culture, the initiative honors and recognizes the sources and strength of Black self-education practices, collective study, and librarianship. The *#SchomburgSyllabus* is organized into 27 themes, including Black feminism, disability, fashion, film and TV, gentrification, monuments, music, and more, all aimed at fostering a greater understanding of the Black experience.

The University of California, Berkeley maintains the *Archiving the LGBT Web: Eastern Europe and Eurasia* project<sup>14</sup>, a web archive dedicated to exploring strategies for preserving online content related to the LGBT community in Russia and Eastern Europe, a region with constantly shifting legal and social attitudes. Archives focused on this community are not a recent phenomenon. For instance, the *ONE National Gay & Lesbian Archives*<sup>15</sup>, currently housed at the University of Southern California Library, was founded in 1952. There are numerous other examples of LGBT archives worldwide, such as the *Canadian Lesbian+Gay Archives*<sup>16</sup> and the *International homo/lesbisch centrum voor informatie en archivering*<sup>17</sup> in the Netherlands. However, while these archives collect materials related to the LGBT movement, they often lack systematic web archiving policies that would ensure the preservation of the LGBT web for future generations.

The *University of British Columbia* maintains the *First Nations and Indigenous Community Websites*<sup>18</sup>, a collection of information related to Indigenous peoples and community content. Websites from various institutions associated with this population are preserved, as well as news that circulates on websites in the United States and Canada.

A notable government initiative in preservation promoting social justice is the UKWA<sup>19</sup> collections, considered the largest of its kind. Beyond government websites, the initiative features collections related to diverse societal aspects, including aging, Black and Asian Britons, Caribbean communities in the United Kingdom, gender equality, LGBTQIAP+ lives online, religion, and many others.

Fortunately, the examples of initiatives promoting social inclusion through web archiving are numerous and diverse. Here, we have highlighted just a few, representing a subset of socially excluded communities. These examples aim to illustrate the wide range of possibilities for web archiving and selection criteria that go beyond the application of evaluative techniques, which can often be exclusionary and limiting, perpetuating existing

---

13 *Site:* <https://www.nypl.org/schomburgsyllabus>.

14 *Site:* <https://archive-it.org/collections/6165>.

15 *Site:* <http://www.onearchives.org/>.

16 *Site:* <https://arquives.ca/>.

17 *Site:* <https://ihlia.nl/>.

18 *Site:* <https://archive-it.org/collections/4069>.

19 *Site:* <https://www.webarchive.org.uk/en/ukwa/category>.

power structures within the archival record. While progress is being made, archivists still have much work to do to ensure that archives represent all voices, not just those in positions of power. It is widely acknowledged that the selection, organization, and access to archival records are pivotal activities that shape public memory by determining what is preserved (remembered) and what is discarded (forgotten). In this context, the stakes are high, as the choices made by archivists have a profound impact on the construction of history and the preservation of cultural and social diversity.

## **FINAL CONSIDERATIONS**

Based on our review of relevant literature, we have presented numerous elements from archival discourse that consider social issues as a crucial evaluative criterion. We have also introduced ten web archiving initiatives and projects that prioritize social justice. Analyzing this dataset reveals a strong connection between document evaluation, web archiving, and social justice.

There is substantial engagement from archivists and professionals across various disciplines in discussions surrounding website preservation. They are actively seeking methods and tools to achieve the goals of preserving and providing access to web pages. Similarly, numerous studies advocate for the evaluation of archives based on social criteria. Despite these collaborative efforts within the research community, there remains a significant need to bridge the gap between these studies and the actual practices of archivists involved in evaluating web content for archiving. It is essential to further explore how this socially conscious approach can help reduce social inequalities by ensuring that memory spaces include communities previously underrepresented.

The first step in these discussions is acknowledging that systemic inequalities and discrimination against individuals or groups can, over time, lead to their erasure from the historical record. Next, recognizing the power of archivists to intervene in this process is paramount. Finally, it is essential to rethink selection processes beyond the Anglo-Saxon archival tradition that has been internalized in many archiving practices, particularly in the Brazilian context. These actions can have a wide range of impacts, from immediate outcomes like legal victories or increased awareness of inequality and discrimination among exhibition visitors, to long-term effects that contribute to meaningful changes in justice and social equality.

Document evaluation techniques are inherently rooted in specific cultural realities that promote particular “moral and technical orders,” concretely shaping how records are preserved. By critically examining and adapting these techniques, we can ensure a more diverse and inclusive documented heritage for both present and future generations, fostering a stronger sense of community and belonging. Classical evaluation theories were developed within the context of the 19th and 20th centuries. Web archives, like other born-digital documents, are products of the contemporary era and demand new approaches to management and evaluation.

Web archiving is a multifaceted endeavor, encompassing critical decisions regarding which aspects of the web are preserved and how they are organized and accessed in the future. Given that the web mirrors the world's cultural and social diversity, considering the cultural dimensions that constitute it is paramount when evaluating what content merits preservation. Online documents encapsulate a myriad of perspectives, ideas, and cultural expressions vital for comprehending contemporary society in its entirety. When archiving the web, recognizing and respecting this diversity is essential to ensure the equitable representation of various voices and cultures.

It's crucial to acknowledge that archiving decisions are not inherently neutral; they are inevitably influenced by underlying values, beliefs, and power structures within society. Neglecting cultural dimensions risks perpetuating inequalities and injustices by privileging certain narratives over others. Therefore, carefully considering these dimensions is vital to ensure that web archiving actively promotes social justice. We argue that taking cultural dimensions into account is fundamental to web archiving because it allows us to accurately and inclusively preserve the cultural diversity of contemporary society, ensuring that the choices made today foster justice and equity for future generations.

While acknowledging the limitations of this study, which examined a sample of only ten web archives and is not intended to be exhaustive, this article offers valuable insights into the curatorial and content evaluation aspects of web archiving. Despite its limitations, it serves as a starting point for more extensive and qualitative research, and we anticipate that this conversation will continue to evolve in the coming years. Our aim is to initiate a broader dialogue to identify alternative pathways for fruitful engagement and propel this discourse forward.

In conclusion, web archiving holds immense potential as a tool for safeguarding cultural heritage and promoting social inclusion. The internet has become an increasingly integral part of people's daily lives, transforming into a primary medium for communication and cultural expression. Web archiving ensures that cultural and historical manifestations, including events and digital content published online, are preserved for future generations. Furthermore, web archiving can champion social inclusion by guaranteeing access to information and content for all, regardless of geographical location or socioeconomic status.



## **BIBLIOGRAPHY**

BARRIAULT, M. Archiving the queer and queering the archives: a case study of the Canadian Lesbian and Gay Archives (CLGA). In: Bastian J. A.; Alexander B. (ed.). Principles and practice in records management and archives. **Community archives: the shaping of memory**, London: Facet Publishing, p. 97-108, 2009. DOI: <https://doi.org/10.29085/9781856049047.008>.

BASTIAN, J. A. Flowers for homestead: a case study in archives and collective memory. **The American Archivist**, Chicago, IL, v. 72, n. 1, p. 113-132, Spring - Summer, 2009.

BASTIAN, J. A. **Owning memory**: how a Caribbean community lost its archives and found its history. Westport, Conn.; London: Libraries Unlimited, 2003. ISBN: 031332008X.

BASTIAN, J. A. Reading colonial records through an archival lens: the provenance of place, space and creation. **Archival science**, Dordrecht, v. 6, p. 267-284, 2006. DOI: <https://doi.org/10.1007/s10502-006-9019-1>.

BASTIAN, J. A.; ALEXANDER, B. Introduction: communities and archives: a symbiotic relationship. **Community archives: the shaping of memory**, London: Facet Publishing, p. xxi-xxiv, 2009. DOI: <https://doi.org/10.29085/9781856049047.002>.

BAUMGARTNER, J. N.; BURNS, J. K. Measuring social inclusion: a key outcome in global mental health. **International journal of epidemiology**, [s. l.], v. 43, n. 2, p. 354-364, 2014. DOI 10.1093/ije/dyt224.

BEARMAN, D. An indefensible bastion: archives as repositories in the electronic age. **Archival management of electronic records**, [s. l.], v. 13, p. 14-24, 1991.

BEN-DAVID, A.; AMRAM, A. The Internet Archive and the socio-technical construction of historical facts. **Internet Histories**, [s. l.], v. 2, n. 1-2, p. 179-201, 2018.

BERNERS-LEE, T. **Cool URIs don't change**. W3C Style, [s. l.], 2013. Available at:<http://www.w3.org/Provider/Style/URI>. Access in:10 maio 2023.

BOERES, S. A. A.; SAAD, R. G. Arquivamento da Web: definições, estratégias, fluxos e iniciativas. **Revista Brasileira de Preservação Digital**, Campinas, SP, v. 4, n. 00, p. e023005, 2023. DOI 10.20396/rebpred.v4i00.17934. Available at:<https://econtents.bc.unicamp.br/inpec/index.php/rebpred/article/view/17934>. Access in:21 fev. 2024.

BNF. **Bibliothèque Nationale de France**. Internet archives. 2017. Available at:[http://www.bnf.fr/en/collections\\_and\\_services/book\\_press\\_media/a.internet\\_archives.html](http://www.bnf.fr/en/collections_and_services/book_press_media/a.internet_archives.html). Access in:26 maio 2023.

BOOMS, H.; JOLDERSMA, H.; KLUMPENHOUWER, R. Society and the formation of a documentary heritage: issues in the appraisal of archival sources. **Archivaria**, Ottawa, v. 24, p. 69-107, 1987.

BRAGG, M.; HANNA, K.; DONOVAN, L.; HUKILL, G. PETERSON, A. The Web Archiving Life Cycle Model. White Paper. **The Archive-It team**: Internet Archive, [s. l.], 2013. Available at:[http://ait.blog.archive.org/files/2014/04/archiveit\\_life\\_cycle\\_model.pdf](http://ait.blog.archive.org/files/2014/04/archiveit_life_cycle_model.pdf). Access in: 10 fev. 2023.

BROWN, A. **Archiving websites**: a practical guide for information management professionals. London: Facet Publishing, 2006.

BRÜGGER, N. Web archiving—Between past, present, and future. **The handbook of Internet studies**, [s. l.], p. 24-42, 2011. DOI <https://doi.org/10.1002/9781444314861.ch2>.

BRÜGGER, N. When the present web is later the past: Web historiography, digital history, and internet studies. **Historical Social Research/ Historische Sozialforschung**, [s. l.], v. 37, n. 4 (142), p. 102–117, 2012.

BRÜGGER, N. The archived web: doing history in the digital age. **MIT Press**, Cambridge, Massachusetts, EUA, 2018.

CAIDI, N.; ALLARD, D. Social inclusion of newcomers to Canada: An information problem? **Library & Information Science Research**, [s. l.], v. 27, n. 3, p. 302-324, 2005. ISSN 0740-8188.

CASWELL, M. Khmer Rouge archives: accountability, truth, and memory in Cambodia. *Archival Science*, Dordrecht, v. 10, p. 25-44, 2010.

CASWELL, M. "Not Just between Us: A Riposte to Mark Greene". **The American Archivist**, Chicago, IL, v. 76, n. 2, p. 605–8, 2013.

CASWELL, M. Inventing New Archival Imaginaries: theoretical Foundations for Identity-Based Community Archives. In: DANIEL, D.; LEVI, A. S. (Ed.). **Identity Palimpsests**: Archiving Ethnicity in the U.S. and Canada. Sacramento, CA: Litwin Books (Archives, archivists and society: book 5), 2014. p. 35-58. ISBN 978-1-936117-85-7.

CASWELL, M.; BROMAN, G.; KIRMER, J.; MARTIN, L.; SOWRY, N. Implementing a social justice framework in an introduction to archives course: lessons from both sides of the classroom. **InterActions**: UCLA Journal of Education and Information Studies, California, v. 8, n. 2, 2012. DOI: <https://doi.org/10.5070/D482011846>.

COOK, T. "Many Are Called, but Few Are Chosen": Appraisal Guidelines for Sampling and Selecting Case Files. **Archivaria**, Ottawa, 1991.

COOK, T. Mind over matter: towards a new theory of archival appraisal. **The Archival Imagination: Essays in Honour of Hugh A. Taylor**, (Ottawa: Association of Canadian Archivists), p. 38–70, 1992.

COOK, T. Macro-appraisal and functional analysis: documenting governance rather than government. **Journal of the Society of Archivists**, [s. l.], v. 25, n. 1, p. 5-18, 2004.

COOK, T. Evidence, memory, identity, and community: four shifting archival paradigms. **Archival science**, Dordrecht, v. 13, p. 95-120, 2013.

COOK, T.; SCHWARTZ, J. M. Archives, records, and power: from (postmodern) theory to (archival) performance. **Archival science**, Dordrecht, v. 2, p. 171-185, 2002.

COX, R. J. Archivists and collecting. *In*: **Encyclopedia of library and information sciences**. CRC Press, Boca Raton, Florida, EUA, 2009. p. 208-220.

CUSHING, L. **All of us or none**: social justice posters of the San Francisco Bay area. Heyday, [s. l.], 2014.

DANIEL, D. Documenting the immigrant and ethnic experience in American archives. **The American Archivist**, Chicago, v. 73, n. 1, p. 82-104, 2010.

DOUGLAS, J. Origins: evolving Ideas about the principle of provenance: *In*: EASTWOOD, T.; MACNEIL, H. (org.). Currents of archival thinking. London: **Libraries Unlimited**, 2010.

DUFF, W. M.; FLINN, A.; SUURTAMM, K. E.; WALLACE, D. A. Social justice impact of archives: a preliminary investigation. **Archival Science**, Dordrecht, v. 13, ed. 4., p. 317-348, 2013. DOI:10.1007/s10502-012-9198-x.

EASTWOOD, T. Reflections on the goal of archival appraisal in democratic societies. **Archivaria**, Ottawa, p. 59-71, 2002.

FERREIRA, L. B.; MARTINS, M. R.; ROCKEMBACH, M. Usos do arquivamento da web na comunicação científica. **Prisma.com**, [s. l.], n. 36, p. 78-98, 2018.

FLINN, A.; STEVENS, M. 'It is noh mistri, wi mekin histri': telling our own story: independent and community archives in the UK, challenging and subverting the mainstream. **Community archives: The shaping of memory**, London: Facet Publishing, p. 3-28, 2009.

FLINN, A.; STEVENS, M.; SHEPHERD, E. Whose memories, whose archives? Independent community archives, autonomy and the mainstream. **Archival science**, Dordrecht, v. 9, p. 71-86, 2009.

FRITZ, S.; MILLIGAN, I.; RUEST, N.; LIN, J. *Fostering Community Engagement through Datathon Events: The Archives Unleashed Experience*. DHQ: Digital Humanities Quarterly, [s. l.], v. 15, n. 1, 2021.

FUNDAÇÃO NELSON MANDELA. *Memory for justice: report on a colloquium*. Nelson Mandela Foundation, Johannesburg, [s. l.], 2005. Disponível em <http://faculty.wvu.edu/~jimerson/ArchivesPower-Appendix.htm>. Access in: 15 maio 2023.

GLANVILLE, L. Web archiving: ethical and legal issues affecting programmes in Australia and the Netherlands. **The Australian Library Journal**, [s. l.], v. 59, n. 3, p. 128–134, 2010. DOI 10.1080/00049670.2010.10735999.

GOMES, D.; FREITAS, S.; SILVA, M. J. Design and selection criteria for a national web archive. *In: Research and Advanced Technology for Digital Libraries: 10th European Conference, ECDL 2006, Alicante, Spain, Proceedings 10*. Springer Berlin Heidelberg, Sep. 2006. p. 196-207.

GREENE, M. Ethics in Archives: is social justice an archival ethic? **Archival Outlook**, 2007. Available at: <http://files.archivists.org/periodicals/Archival-Outlook/Back-Issues/2007-6-AO.pdf>. Access in: 10 maio 2023.

GREENE, M. A critique of social justice as an archival imperative: what is it we're doing that's all that important? **The American Archivist**, Chicago, v. 76, n. 2, p. 302-334, 2013.

HAM, F. G. Archival strategies for the post-custodial era. **The American Archivist**, Chicago, v. 44, n. 3, p. 207-216, 1981.

HARRIS, V. S. The archival sliver: power, memory, and archives in South Africa. **Archival science**, Dordrecht, v. 2, p. 63-86, 2002.

HARRIS, V. S. Archives and justice: a South African perspective. **Society of Amer Archivists**, [s. l.], 2007.

HARRIS, V. S. Antonyms of our remembering. **Archival Science**, Dordrecht, v. 14, n. 3-4, p. 215-229, 2014.

HEDSTROM, M. Archives, memory, and interfaces with the past. **Archival Science**, Dordrecht, v. 2, p. 21-43, 2002.

HEDSTROM, M. Archives and collective memory: more than a metaphor, less than an analogy. **Currents of archival thinking**, [s. l.], p. 163-179, 2010.

HURLEY, C. Parallel provenance [Series of parts]: Part 1: What, if anything, is archival description? [An earlier version of this article was presented at the Archives and Collective Memory: Challenges and Issues in a Pluralised Archival Role seminar (2004: Melbourne)]. **Archives and Manuscripts**, [s. l.], v. 33, n. 1, p. 110-145, 2005.

IIPC. Why archive the web? **Website IIPC**. [s. l.], 2017. Available at: <http://netpreserve.org/web-archiving/>. Access in: 20 maio 2023.

INTERNET ARCHIVE. **About the internet archive**, San Francisco, 2023. Available at: <https://archive.org/>.

Access in: 10 maio 2023.

JACKMAN, T. Activist archivists and the limits of postmodern thought. **A paper for ARST**, [s. l.], v. 573, p. 1-22, 2012.

JACOBSEN, T.; PUNZALAN, R. L.; HEDSTROM, M. L. Invoking “collective memory”: Mapping the emergence of a concept in archival science. **Archival Science**, Dordrecht, v. 13, p. 217-251, 2013.

JIMERSON, R. C. Archives for all: professional responsibility and social justice. **The American Archivist**, Chicago, IL, v. 70, n. 2, p. 252-281, 2007.

JIMERSON, R. C. **Archives power: memory, accountability, and social justice**. Chicago: Society of American Archivists, 2009.

JIMERSON, R. C. Archivists and social responsibility: a response to Mark Greene. **The American Archivist**, Chicago, IL, p. 335-345, 2013.

JIMERSON, R. C.; HARRIS, V.; WALLACE, D. A. *Global issues forum: archival ethics and social justice: what is our professional responsibility?*. In: SOCIETY OF AMERICAN ARCHIVISTS. **Archives 2008: Plan now to attend SAA's 72nd Annual Meeting**. Hilton San Francisco, 2008.

JOSIAS, A. Toward an understanding of archives as a feature of collective memory. **Archival science**, Dordrecht, v. 11, p. 95-112, 2011.

KAPLAN, E. We are what we collect, we collect what we are: Archives and the construction of identity. **The American Archivist**, Chicago, v. 63, n. 1, p. 126-151, 2000.

KETELAAR, E. Tacit narratives: the meanings of archives. **Archival science**, Dordrecht, v. 1, p. 131-141, 2001.

KHAN, M.; RAHMAN, A. U. A systematic approach towards web preservation. **Information Technology and Libraries**, [s. l.], v. 38, n. 1, p. 71–90, 2019. DOI: <https://doi.org/10.6017/ital.v38i1.10181>.

KOLISH, E. Sampling Methodology and its Application: an illustration of the tension between theory and practice? **Archivaria**, Ottawa, 1994.

LEVITAS, R. Let's hear it for Humpty: social exclusion, the third way and cultural capital. **Cultural trends**, [s. l.], v. 13, n. 2, p. 41-56, 2004.

MACKINNON, K. Ethical approaches to youth data in historical web archives (dispatch). **Studies in Social Justice**, Hamilton, Canada, v. 15, n. 3, p. 442–449, 2021. ISSN 1911-4788.

MAEMURA, E.; WORBY, N.; MILLIGAN, I.; BECKER, C. If these crawls could talk: studying and documenting web archives provenance. **Journal of the Association for Information Science and Technology**, [s. l.], v. 69, n. 10, p. 1223-1233, 2018. DOI: <https://doi.org/10.1002/asi.24048>.

MANDER, D. Special, local and about us: the development of community archives in Britain. **Community archives: the shaping of memory**, London: Facet Publishing, p. 29-46, 2009.

MASANÈS, J. Web archiving methods and approaches: a comparative study. **Library trends**, Baltimore, Maryland, EUA, v. 54, n. 1, p. 72-90, 2005.

MELO, J. F. **Arquivamento dos websites do Governo Federal Brasileiro - preservação do domínio GOV.BR**. 2020. 133 f. Dissertação (Mestrado – Programa de Pós-Graduação em Comunicação e Informação) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2020.

MELO, J. F.; OLIVEIRA, C.; ROCKEMBACH, M. História do arquivamento da web no Brasil: um percurso entre a academia, o legislativo e executivo brasileiro. Simpósio da História dos Arquivos e da Arquivologia (2023 maio. 25-26: Niteroi, RJ). **Anais** [...] [recurso eletrônico]. Niteroi: UFF, 2023.

MILLAR, L. The death of the fonds and the resurrection of provenance: archival context in space and time. **Archivaria**, Ottawa, p. 1-15, 2002.

MILLIGAN, I. Historiography and the Web. **The SAGE handbook of web history**, [s. l.], p. 3–15, 2018.

MILLIGAN, I. **History in the age of abundance?**: how the web is transforming historical research. McGill-Queen's University Press, [s. l.], 2019. p. 328. ISBN 978-0-7735-5697-3.

MOHR, G.; STACK, M.; RANITOVIC, I.; AVERY, D.; KIMPTON, M. An Introduction to Heritrix: An open source archival quality web crawler. *In: 4th International Web Archiving Workshop (IWAW04)*, 4, 2004, Bath, UK, p. 109-115, 2004.

NÉMETH, M.; DRÓTOS, L. A blended learning-based curriculum on Web archiving in the national Széchényi library. **Digital Library Perspectives**, [s. l.], v. 35, n. 2, p. 97-114, 2019. DOI: <https://doi.org/10.1108/DLP-03-2019-0012>.

NESMITH, T. Still fuzzy, but more accurate: Some thoughts on the "ghosts" of archival theory. **Archivaria**, Ottawa, p. 136-150, 1999.

NESMITH, T. Documenting Appraisal as a Societal-Archival Process: theory, practice, and ethics in the wake of Helen Willa Samuels. *In: COOK, T. (ed.). Controlling the Past: Documenting Society and Institutions - Essays in Honor of Helen Willa Samuels*. Chicago: Society of American Archivists, 2011. p. 31-50.

NYE, B. E. Records retention schedules. *In: Encyclopedia of Library and Information Sciences*. CRC Press, Boca Raton, Florida, EUA, p. 4465-4469, 2009.

OGDEN, J. "Everything on the internet can be saved": Archive Team, Tumblr and the cultural significance of web archiving. **Internet Histories**, [s. l.], v. 6, n. 1-2, p. 113-132, 2022.

OGDEN, J.; HALFORD, S.; CARR, L. Observing web archives: the case for an ethnographic study of web archiving. **WebSci '17: Proceedings of the 2017 ACM on Web Science Conference**, Troy, NY, USA, p. 299-308, 2017. DOI 10.1145/3091478.3091506.

PEARCE-MOSES, R. **A glossary of archival and records terminology**. Chicago: Society of American Archivists, 2005. ISBN 1-931666-14-8.

PHILLIPS, M. E. What should we preserve? The question for heritage libraries in a digital world. **Library trends**, Baltimore, v. 54, n. 1, p. 57-71, 2005.

PUNZALAN, R. L. All the things we cannot articulate': colonial leprosy archives and community commemoration. **Community archives: the shaping of memory**, London: Facet Publishing, p. 197-219, 2009.

PUNZALAN, R. L.; CASWELL, M. Critical directions for archival approaches to social justice. **The Library Quarterly**, [s. l.], v. 86, n. 1, p. 25-42, 2016.

RIBEIRO, F. Da arquivística técnica à arquivística científica: a mudança de paradigma. **Ciências e Técnicas do Patrimônio**: Revista da Faculdade de Letras da Universidade do Porto, 2002. p. 97-110.

RIBEIRO, F. Novos caminhos da avaliação de informação. **Seminário Avaliação de Documentos a Serviço da Administração**, Santa Catarina, 2005. 22 p.

ROCKEMBACH, M. Inequalities in digital memory: ethical and geographical aspects of web archiving. **The International Review of Information Ethics**, Edmonton, v. 26, 2017. DOI 10.29173/irrie286. Available at: <https://informationethics.ca/index.php/irrie/article/view/286>. Access in: 26 feb. 2024.

ROCKEMBACH, M. Arquivamento da Web: estudos de caso internacionais e o caso brasileiro. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 16, n. 1, p. 7-24, 2018a.

ROCKEMBACH, M. A web retrospectiva como campo de pesquisa: arquivamento da web e preservação digital. **Pesquisa e Perspectivas de Comunicação e Informação**. Porto Alegre: Sulina, p. 240-256, 2018b.

ROCKEMBACH, M. A web como memória: preservação digital e arquivamento da web no Brasil e no exterior. In: **Congresso Internacional em Humanidades Digitais** (1.: 2018: Rio de Janeiro, RJ). Caderno de Resumos HDRio 2018. Rio de Janeiro: FGV, 2018c.

ROCKEMBACH, M. Arquivamento da Web no contexto das humanidades digitais: da produção a preservação da informação digital. **LIINC em revista**, Rio de Janeiro, RJ, v. 15, n. 1, p. 131-139, 2019.

ROCKEMBACH, M.; FERREIRA, L. B. Web archiving of elections and Brazilian possibilities. In: RIBEIRO, F.; CERVEIRA, M. E. (ed.). **Challenges and Opportunities for Knowledge Organization in the Digital Age: Proceedings of the Fifteenth Interantional ISKO Conference 9-11 July 2018**. Ergon-Verlag, Porto, v. 16, 2018. p. 969-972.

ROCKEMBACH, M.; PAVÃO, C. M. G. Políticas e tecnologias de preservação digital no arquivamento da web. **Revista Ibero-Americana de Ciência da Informação**, Brasília, v. 11, n. 1, p. 168-182, 2018.

RODRIGUES, V. L. D.; ROCKEMBACH, M. Arquivos da web como fonte historiográfica. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, 2021.

ROGERS, R.; BRÜGGER, N.; MILLIGAN, I. Periodizing web archiving: Biographical, event-based, national and autobiographical traditions. **The SAGE handbook of web history**, [s. l.], p. 42–56, 2018.

SAIFUL BAHRY, F. D.; AMRAN, N.; PUTRI, T. E.; RAMLI, M. I. Database design of the Malaysia public figures web archive repository: a social and cultural heritage web collections. **Collection and Curation**, [s. l.], v. 41, n. 4, p. 133-140, 2022. ISSN: 2514-9326. DOI 10.1108/CC-09-2021-0025.

SAMUELS, H. Who controls the past? **The American Archivist**, Chicago, v. 49, n. 2, p. 109-124, 1986.



SAMUELS, H. W. **Varsity letters**: documenting modern colleges and universities. [S. l.]: Scarecrow Press, 1998.

SAUER, C. Doing the best we can? The use of collection development policies and cooperative collecting activities at manuscript repositories. **The American Archivist**, Chicago, v. 64, n. 2, p. 308-349, 2001.

SCHNEIDER, S. M.; FOOT, K. A.; KIMPTON, M.; JONES, G. Building thematic web collections: challenges and experiences from the September 11 Web Archive and the Election 2002 Web Archive. *In*: **3rd Workshop on Web Archives (IWAW'03)**, Digital Libraries, ECDL, Trondheim, p. 77-94, 2003.

SCHNEIDER, S. M.; FOOT, K. A. Web sphere analysis: An approach to studying online action. **Virtual methods**: Issues in social research on the Internet, Berg Publishers, Oxford, p. 157-170, 2005.

SCHNEIDER, S.; FOOT, K. Archiving of internet content. *In*: DONSBACH, W. (ed.). *The International Encyclopedia of Communication*. Wiley Publishing, Malden, 2008.

SCHROEDER, R.; BRÜGGER, N. **Introduction**: The web as history. *In*: BRÜGGER, N. SCHROEDER, R. (ed.). *The Web as History: Using Web Archives to Understand the Past and the Present*. UCLPRESS, 2017.

SHILTON, K.; SRINIVASAN, R. Participatory appraisal and arrangement for multicultural archival collections. **Archivaria**, Ottawa, p. 87-101, 2007.

SILVA, A. M. **A Informação**: da compreensão do fenómeno e construção do objecto científico. Porto: Edições Afrontamento, 2006.

SILVA, A. M.; RIBEIRO, F. **Paradigmas, serviços e mediações em ciência da informação**. Recife: Nectar, 2011.

SILVA, A. M.; RIBEIRO, F. Documentation/Information and their paradigms: characterization and importance in research, education, and professional practice. **Knowledge organization**, [s. l.], v. 39, n. 2, p. 111–124. 2012.

SILVA, A. M.; RIBEIRO, F. *Ciência da informação trans e interdisciplinar: para a superação de equívocos*. **Ciência da Informação**: visões e tendências, Coimbra, 2020.

SOCIETY OF AMERICAN ARCHIVISTS. **In pursuit of the moral imperative**: exploring social justice and archives. Annual Conference, 2012a.

SOCIETY OF AMERICAN ARCHIVISTS. Spotlight: oral history and social justice. *In*: SOCIETY OF AMERICAN ARCHIVISTS. **Dialogue**: The Newsletter of the Oral History Section. v. 8, ed. 1 (Spring, 2012), p. 4–9, 2012b. Available at:<https://www2.archivists.org/sites/all/files/spring2012final%20draft.pdf>. Access in: 14 maio 2023.

SUMMERS, E. Appraisal talk in web archives. **Archivaria**, Ottawa, v. 89, n. 1, p. 70-102, 2020.

SUMMERS, E.; PUNZALAN, R. Bots, seeds and people: Web archives as infrastructure. *In*: LEE, C. P.; POLTROCK, S.; BARKHUUS, L.; BORGES, M.; KELLOGG, W. CSCW '17: **Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing**. New York, NY: Association for Computing Machinery, 2017. p. 821-834. ISBN: 978-1-4503-4335-0.

UNESCO. **Digital Heritage**. Website. [s. l.], 2021. Available at:<https://en.unesco.org/themes/information-preservation/digital-heritage>. Acesso em 26 maio 2023.

WALLACE, D. A. Locating agency: interdisciplinary perspectives on professional ethics and archival morality. **Journal of Information Ethics**, [s. l.], v. 19, n. 1, p. 172, 2010. DOI 10.3172/JIE.19.1.172.

WEBER, M. S. Web archives: a critical method for the future of digital research. **WARCNet Papers**, Aarhus, Dinamarca, v. 17, 2020.

WURL, J. Ethnicity as provenance: in search of values and principles for documenting the immigrant experience. **Archival issues**, [s. l.], p. 65-76, 2005.

YACO, S.; HARDY, B. B. Historians, archivists, and social activism: benefits and costs. **Archival Science**, Dordrecht, v. 13, p. 253-272, 2013.

YACO, S.; HARDY, B. B. Archivists, Social Activism, and Professional Ethics. *In*: Society of American Archivists (Society of American Archivists – 2011 Research Forum), 2012. Available at:<http://www2.archivists.org/sites/all/files/SoniaYacoBeaHardy-SAA-ResearchPaper-2011.pdf>. Access in: 10 maio 2023.

YAKEL, E. Digital curation. **OCLC Systems & Services: International digital library perspectives**, [s. l.], v. 23, n. 4, p. 335-340, 2007.

ZINN, H. Secrecy, archives, and the public interest. **The Midwestern Archivist**, [s. l.], v. 2, n. 2, p. 14-26, 1977.