

In the future, will news archives no longer be accessible?

En el futuro, ¿los archivos de noticias dejarán de ser accesibles? / No futuro, os acervos de notícias não serão mais acessíveis?

Lucia Santa-Cruz

PhD in Communication and Culture from Universidade Federal do Rio de Janeiro (UFRJ). Adjunct Professor of the Master's Degree Professional in Creative Economy, Strategy and Innovation and courses degree in Journalism, Cinema and Advertising at ESPM Rio, Brazil.
lucia.santacruz@espm.br

ABSTRACT

Humanities research usually relies on journalistic material as documental data. Even if some of these studies use digital material, they were originally physical outlets. As news outlets shift towards digital platforms, it is becoming increasingly challenging to access historical news archives or even current content. This article aims to discuss in what extent this information treatment will affect social science in the future.

Keywords: news archives; digital news; digital media; news research.

RESÚMEN

La investigación en humanidades suele basarse en material periodístico como datos documentales. Aunque algunos de estos estudios utilicen material digital, originalmente eran productos físicos. A medida que los medios de comunicación se trasladan a plataformas digitales, se está volviendo cada vez más difícil acceder a archivos de noticias históricas o incluso contenido actual. Este artículo tiene como objetivo discutir en qué medida este tratamiento de la información afectará a las ciencias sociales en el futuro.

Palabras clave: archivos de noticias; noticias digitales; medios digitales; investigación de noticias.

RESUMO

A pesquisa em humanidades geralmente se baseia em material jornalístico como dados documentais. Mesmo que alguns desses estudos usem material digital, sua origem vem de veículos físicos. À medida que os veículos de notícias migram para plataformas digitais, está se tornando cada vez mais desafiador acessar arquivos de notícias históricas ou até mesmo conteúdo atual. Este artigo tem como objetivo discutir em que medida esse tratamento da informação afetará as ciências sociais no futuro.

Palavras-chave: acervos de notícias; notícias digitais; mídias digitais; pesquisa de notícias.

Introduction

News archives play a vital role in preserving and providing access to historical records of events, insights into past media practices and cultural norms, and valuable resources for research, journalism, and public discourse. They are repositories of journalistic production that gather historical records of events and provide context for current affairs. The increasing digitalization of news archives has made access to these resources easier, but it has also created new threats to their preservation and access.

They are crucial for understanding the evolution of media practices and cultural norms, including changes in reporting styles, editorial policies, and media technologies. Furthermore, news archives are valuable resources for researchers, journalists, educators, and the public, enabling them to analyze trends, patterns, and issues across time and space. They also serve as sources for journalism, providing background and context for current affairs. News archives have been used for research into topics such as political history, social history, and cultural history. They have also been used for journalism on topics such as political scandals, natural disasters, and sports events. Access to news archives is crucial for both research and journalism, as it provides a wealth of information that is not available through other sources.

News archives are also vital for public access to information. They provide a record of events that are of public interest and document the workings of government, business, and society. News archives have been used by individuals, organizations, and government agencies for a range of purposes, such as legal research, advocacy, and public policy development. They are an essential source of information for citizens, providing a means to hold those in power accountable.

For historians, news archives are a great source of data. Few historical sources permit so many investigation possibilities than that media. Carvalho stressed that newspapers in general comprehend two elementary dimensions. First, its discursive dimension, meaning, its ability to organize the world, to establish facts, to produce consensus and to lend meaning to historical experience. Second, but not less important, its capacity to enregister the most distinctive cultural, political, economic, social, and even natural phenomena (Carvalho, 2016, p. 90).

In fact, the relationship between journalism and history sounds clear in sentences like “the journalist is the historian of the instant”, coined during the Second World War by the French writer and philosopher Albert Camus; or “eyewitness to history” the slogan of a famous radio journalistic show in Latin America; or the way the Brazilian reporter Eliane Brum defines herself: A historian of the routine life (Santa-Cruz, 2014).

Taken as important documentary sources, journalistic periodicals establish themselves as the memory of society. A memory available and updated daily, which has grown exponentially since the 2000s, with the digitization of many titles, from small and local newspapers to large ones and of international circulation, from the non-current to those still on the streets.

Recent research conducted in Brazil revealed some challenges in this scenario. The study compared how Brazilian newspapers reported the country's independence celebration throughout different historical periods. Brazil got independent from Portugal on September 7th, 1822. The research compared how Brazilian newspapers talked about independence in 1922, the year of its centenary, with the approach of Brazilian digital journalistic outlets at the bicentenary in 2022.

To collect data from 1922, we used the digitized archives of newspapers hold at Biblioteca Nacional (the most ancient Brazilian library) available online, named *Hemeroteca Digital*. Periodicals are organized by title, year, publication city, and keywords. The corpus was consistent, reliable, and accessible.

The situation differs from the experience of collecting data from 2022. First, we tried to use Google News, a tool from Google Company that aggregates news from sources all over the world (Google News Initiative, 2022). We customized it to select stories about the bicentennial at eleven digital outlets.¹ The platform delivered the select news daily. But the daily email only contained the headlines, not the content, even though the news industry is the providers of the aggregator (Ouakrat; Paterson; Rebillard; Surm; Moreira-Cesar, 2023). To read the complete articles, we had to click on the links. But the content was blocked by the news firewall.

The deadlock led us to formulate some questions. In the future, will it be possible to research media history? How will researchers access news archives? Will digital news compromise journalistic and historical research? Besides that, there is another question: Since a published story may have several daily updates, will it be possible to distinguish those different versions? News articles are documental data for humanities research. To what extent does this information treatment will affect social science in the future?

This article examines the potential loss of access to news archives, due to many factors, and its impact on research and journalism itself, focusing on Brazilian examples. First, we will address the digitalization of news archives and the challenges around their preservation. Following with some issues around the subject, we will discuss digital news reliability. In the sequence, we

¹ Those eleven digital outlets were the most accessed in Brazil according to the July 2022 Comscore ranking (Comscore MMX MP – Key Measures – News/Information – Total Internet – July 2022).

will approach the problem of restricted access to news archives and the consequences of that restriction.

The digitalization of news archives – open access experiences

In recent years, there has been a shift from print to digital news archives. Several news outlets around the world digitized their own archives. At the same time, we notice other experiences, as the Google News Archive, the British Newspaper Archive, the Library of Congress's Chronicling America or Brazilian Hemeroteca Digital. All of them understand news as important pieces of historical data.

This movement enlarged the access to information. Carvalho (2016, p. 91) points out that one of the mainly benefits of digitizing historical collections is the fact that it helps democratize access to knowledge:

For a long time, archives, museums, and libraries were the faithful and exclusive depositaries of documents. Access was limited: one had to personally go to these physical spaces, usually located in major cities, to consult what is public heritage. And even when access was granted to these institutions, one had to rely on the availability of the document, which could already be in use by another user or absent for cleaning or restoration, for example. The logic of the Internet has helped subvert this logic. Once digitized and made available on the Web, the collection of these institutions becomes truly public.

Besides, there is the important issue of preservation. Digitalizing prevents the handling effects of old documents. At Hemeroteca Digital, for instance, it is possible to consult newspapers and pamphlets of the Brazilian Independence period, which are rare and could be damaged if constant handled by researchers. Every document faces the possibility of disappearing. Even when researchers use strict consultation methods, wear and tear can occur due to their handling of the original document. The integrity of the document can also be compromised by factors such as exposure to light, moisture, and manipulation by the researcher. Furthermore, it is important to remember that theft, fires, floods, vandalism, and improper storage can all contribute to irreversible damage or complete destruction of the document.

But the most relevant contribution of digitalization relies on information management. Documents transposed to digital medium can be indexed, that is, a reader may locate information without having to read the entire piece, just defining words for the search. Digital documents turn into database. So, looking

for specific subjects at the newspaper collections becomes easier and quicker, allowing the researcher to cover a more extend corpus or to deepen its study.

Two different experiences between The United States of America and Brazil shows us differences in their approaches to what constitutes the meaning of a comprehensive news archive.

Chronicling America, “a long-term effort to develop an Internet-based, searchable database of U.S. newspapers with descriptive information and select digitization of historic pages” (National Digital Newspaper Program, n.d.) is an example of this shift from printed to digital newspapers. In 2005, the National Endowment for the Humanities (NEH) and the Library of Congress (LC) of the United States of America partnered to form the National Digital Newspaper Program (NDNP). Their goal was to create a digital platform that would bring together newspapers from all American states published between 1770-1963, and also to establish a set of best practices for newspaper digitization (Library of Congress, n.d.).

The project derived from a program that previously funded state-level projects to locate, describe (catalog), and selectively preserve (via treatment and microfilm) historic newspaper collections in that state, published from 1690 to the present (National Digital Newspaper Program, n.d.).

The website became available to the public in 2007 and continue to absorb new materials. There is no cost for accessing or using this database. It now contains selected newspapers from every U.S. state and territory published between 1770-1963. It is a searchable database that is unique among other newspaper digitization projects in the United States in that it provides users with access to the entire corpus of images, recognized words, and descriptive cataloging data for researchers to download and reuse.

The NDNP state partners, who are responsible for selecting and digitizing approximately 100,000 newspaper pages, receive awards to contribute to the *Chronicling America* collection. Participants primarily digitize from microfilm holdings, selecting technically-suitable film, bibliographic completeness, diversity, and “orphaned” newspapers. Each newspaper that has been selected for digitization is accompanied by a “title essay” that provides supplementary description, including basic information about the paper, its geographic area, political and religious affiliations, and specialized audiences, as well as discussions of editors, publishers, and significant events covered by the paper during the relevant time period (National Digital Newspaper Program, n.d.).

Although *Chronicling America* provides free access to information about historic newspapers, it is important to note that it is not a complete collection of newspapers. It contains “historic newspapers pages”. What is preserved and open

to consult has been through a selection process. There is no information available about the criteria of considering an article or story as historic. When research uses the website, they will access a material that has already been filtered by someone else criteria, not his own (not for saying that the journalist work has itself done a previous selection about what facts deserve to appear at the paper).

By the other hand, *Hemeroteca Nacional* displays entire collections of Brazilian newspapers since they are available for digitalization and are free-copyright. The project is part of the *Biblioteca Nacional Digital* (BNDigital), an open, interconnected system focused on preserving Brazil's documentary memory, launched at 2006 (Fundação Biblioteca Nacional, n.d.b).

Biblioteca Nacional is the agency responsible for executing the governmental policy of capturing, storing, preserving, and disseminating the country's intellectual production. With over 200 years of history, it is the oldest Brazilian cultural institution. It has a collection of approximately nine million items and, therefore, has been considered by Unesco (United Nations Educational, Scientific and Cultural Organization) as one of the world's main national libraries.

In 2011, the agency created *Hemeroteca Digital*, a portal of national periodicals that allows broad online access to newspapers, magazines, yearbooks, bulletins, and serial publications. With entirely free access and the possibility of printing the desired pages, researchers from anywhere in the world can consult titles that include everything from the first newspapers created in the country – such as the *Correio Braziliense* and *Gazeta do Rio de Janeiro*, founded in 1808 – to newspapers that were extinct in the 20th century, such as the *Diário Carioca* and *Correio da Manhã*, or that no longer circulate in print form, like the *Jornal do Brasil* (Fundação Biblioteca Nacional, n.d.a).

Even though its physical news archives may achieve 17 km of shelves, not all of them are digitized. BNDigital only provides documents in the public domain or with authorization for publication from the copyright holder (Fundação Biblioteca Nacional, n.d.). Albeit the published material holds the publication authorization, the researchers must ask another permission for use from the copyright holder. This exigence may difficult the study and even delay it due to bureaucratic procedures.

So, as we may notice, those two examples of public archives highlight some questions. The first one is related to the content itself – entire collection or selected parts? When the public agent offer free access to news content, should it be in its entirety or does the institution may deliver filtered pieces? The answer to that question is intertwined with the purpose of releasing the content, as with the context of its production.

The second one is pertinent to the copyright question. As some news organizations may have strict copyright policies that restrict access to their archives, this can reduce the public content at portals like Hemeroteca Digital. This can make it difficult for researchers, journalists, and members of the public to access and use historical news content. In some cases, news organizations may charge high fees for access to their archives or may require permission from the copyright owner before content can be used or republished. Additionally, the copyright status of older news content can be unclear or difficult to determine, which can further complicate efforts to access and use these materials. In the next section, we will present some newspapers owners cases.

There is also aspects of digital archives preservation and the access technology issue, which will be addressed in another section of this article.

The digitalization of news – outlets experiences

The biggest newspapers in the world digitized their entire historical collections but access is not free as it may appear. More than 13 million articles of *The New York Times*, from 1851 to the present are available on the Internet for subscribers only, in a project called Time Machine, made in partnership with Google Cloud² (Ringel, 2023). The British *The Guardian* and *The Observer*, controlled by the same group, also digitized their content from the end of the eighteenth century to the early years of the 2000. Although full access to almost 1,5 million pages of newspapers is paid, it is possible to have a free subscription that allows the person to search, create and share clippings in reduced archive. “Digital reproductions of every page, article and advertisement published in the *Guardian* from 1821-2003 and in the *Observer* from 1791-2003 can be viewed online via the subscription service, newspapers.com” (The Guardian, 2017).

Even with this free account, there is a delay in the editions. Only full subscriptions get the current edition.

In Brazil, the main newspapers followed the same line: *O Globo*, *O Estado de São Paulo*, and *Folha de S. Paulo* digitized their entire historical collection (Santa-Cruz, 2016). *O Globo* was founded in 1925 by journalist Irineu Marinho, whose family later built the largest Latin American media conglomerate, Grupo Globo. *Folha de S. Paulo* was born as an evening newspaper in 1921, and after a fusion with two other papers in 1960, assumed its current name. *O Estado de São*

² The newspaper also offers a free account that gives limited access to articles per month, making difficult to conduct academic searches at its archives without paying.

Paulo, also known as *Estadão*, is one of the oldest newspapers in circulation in the country, being created in 1875.

In none of them, although there is no need of a subscription, non-subscribers have full access to recent editions. At *Folha*, there is a gap of six months in the newspaper's editions. At *O Globo* and *O Estado de São Paulo*, editions are one-month behind. So, the access to newly editions remains paid and traces a problem to academic research.

In spite of these facts, all three have memory projects, in which they use their past editions to elaborate new stories and products. This is a different way of old articles to become news again, but of course do not eliminate tough aspects around the free access to historical edition of newspapers. *O Globo* and *Folha de S. Paulo* organized its collection in 2011, while *O Estado de São Paulo* did it one year later.

O Estadão already emerged with the awareness of the importance of preserving its production. Since the first issue it always archived two copies of each edition. This archive, in 2012, was transformed into an editorial department, changing its name to *Acervo Estadão*. In addition to continuing to physically store the collections of the print newspaper, *Estadão* began digitizing all the material, which can be accessed by subscribers of the vehicle through the website since May 23, 2012. In addition to making available the content published in the 148 years of the periodical, it also allows consultation of unpublished material related to the years 1972 to 1975, a period when the newspaper was under prior censorship during the Brazilian military dictatorship.

Although it calls itself an online archive, it produces content on some topics, such as the Week of Modern Art, Brazilian coins, Cold War, etc., from the material of the vehicle itself. Its homepage, in fact, resembles the homepage of a journalistic website, which clearly indicates its status as a content producer. It is possible to consult texts about personalities of Brazil, the world, or specific areas of knowledge (Science, Politics, Economy, Sports, Culture). It also offers its own material about places (neighborhoods, parks, cities, churches, etc.). And under the news section, it produces journalistic news with archive material, updated from journalistic hooks.

Acervo Estadão not only produces new articles with material from its archives, but also produces specific content for other media. Thus, the *Acervo* maintains accounts on Twitter ([twitter@estadaoacervo](https://twitter.com/estadaoacervo)) and Instagram ([instagram.com/estadaoacervo/](https://www.instagram.com/estadaoacervo/)), as well as a fan page on Facebook ([facebook/arquivos-estadao](https://www.facebook.com/arquivos-estadao)) and a blog updated daily. For this, the editorial department has 11 professionals, including journalists, archivists, and librarians (Santa-Cruz, 2016).

Despite these other uses of the news archives, as source for their own patrons, there is an early digitalization experience. *Jornal do Brasil* (JB) was the first Brazilian newspaper digitized in 2010, as part of the Google News Archive project (Google News, 2011). Google launched that tool in 2008, declaring and ended it three years later, even though the digitized papers are still online. The project intended to digitize the printed editions of newspapers and dispose them freely on a website. It contains titles from various countries, mostly English editions. Related to *Jornal do Brasil*, Carvalho indicates that the results are limited.

Firstly, the JB collection has not been fully digitized. There are no editions available for the period between 1891-1929, except for the solitary issue of December 31, 1910. From the 1930s to the 1990s, most editions have been digitized, but there are abundant gaps in various years. From 2000 to 2010, no editions of the newspaper are available for consultation. Secondly – which sounds quite surprising for a company that revolutionized internet search – the Google News Archive did not index the content, or it was not offered to the user. In other words, it is not possible to search for keywords. The project did not take into consideration aspects such as indexing, complete coverage of the archive, historical context of the digitized material, user participation in the error correction process, among others. Thus, the historical archive of one of the most important newspapers in Brazilian press was brought to the Internet in a precarious way, offering few advantages for researchers and the general public. (Carvalho, 2016, p. 97)

The results confirm what Ventura concluded, when studied the archives of Portuguese newspaper *Público Norte* case: many of the archival aids have migrated to the digital environment without being designed for that purpose (Ventura, 2020). The author emphasizes that studies have shown that the assumptions that have guided the representation of information on digital platforms are traditional normative-instrumental assumptions, characteristic of traditional archival research instruments focused on the description of documents, but which escape the understanding of the user.

Due to these negative results, the owner of *Jornal do Brasil* asked Biblioteca Nacional to redo its digitalization and include it under *Hemeroteca Digital*. Regardless the fact that the library was conducting the digitalization of its physical collections, and JB was not one of those collections, the institution accepted the demand and prioritized the newspaper digital archiving.

Those examples underline the challenges around digital news archives, being public or private ones. In the next section, we will examine the questions about their preservation.

Digital archives preservation

Access to news archives is at risk due to several factors, including funding cuts, technological changes, and business decisions by news organizations. Many news organizations face financial pressures that can lead to cost-cutting measures, including reducing the resources allocated to archiving and digitization. Additionally, the rapid pace of technological change can make older technologies and formats obsolete, making it difficult to preserve and access older archives. Finally, some news organizations may choose to limit access to their archives for business or legal reasons, such as protecting their intellectual property or avoiding liability.

Another significant issue that arises when accessing digital news archives is that of preservation (Formenton; Gracioso, 2020). Physical newspapers and journals were often stored in archives or libraries, ensuring that the information contained within was preserved for future generations. However, in the digital age, where news outlets may shut down or be taken offline, there is a risk that valuable historical data will be lost forever.

As we move towards a more digital media environment, there is a risk that valuable historical data may be lost forever. News outlets may shut down or be taken offline, or data may be lost due to technical failures or human error.

The application of preservation strategies for digital documents is a priority, as without them there would be no guarantee of long-term access, reliability, and integrity of the documents (Arellano, 2004). The main problem related to the preservation of digital archives is technical obsolescence (Verdegem; Slats, 2004). It is necessary to ensure that these files can be read by future technologies. It is well known inside media sector the case of Betacam technology.

The Betacam case is a well-known example of technological obsolescence. Betacam was a professional video format introduced by Sony in 1982, which quickly became the standard format for broadcast-quality video production. However, with the advent of digital video technology in the 1990s, Betacam became increasingly obsolete.

Despite being a high-quality and reliable format, the system was eventually replaced by more advanced digital video formats, which offered higher resolution and better image quality. As a result, many broadcasters and production companies switched to these newer formats, leaving Betacam equipment and tapes largely unused.

This led to a situation where Betacam equipment and tapes became increasingly difficult to find, and maintenance and repair services became more expensive. As a result, many companies were forced to replace their equipment with newer technology, resulting in significant costs.

This case highlights the challenges of technological obsolescence, where technologies that were once dominant can quickly become outdated and unsupported, leading to significant costs and challenges for businesses and industries that rely on them. It also draws attention to the importance of anticipating and planning for technological change, and the need to continually innovate and adapt in order to stay competitive.

Wiggins (2001) analyses digital archives menaces, calling them as digital death. Digital archives preservation can be threatened beyond physical destruction also by the replacement of news information without maintaining the older one; loss of information by a redesign of the web pages; adoption of new formats or languages; overwriting of technology and even by the accidental deletion of content.

As Edwardes (2004) brings out, preserving digital media is complex due to the ephemeral character of internet and electronic archives.

One illustration of those treats is what happens to *O Estado de São Paulo* newspaper. After ten years of Internet, the outlet took some arrangements about its website, and decided to delete the older archives because it changed the warehouse software. So, the media lost a decade of digital content which is irrecoverable.

At that moment, most digital newspapers page was almost an electronic reproduction of its print version. That explains why newspaper executives were not so preoccupied about losing the digital material. They could retrieve most of it by access the print copies of *Estadão*, stored at its own archive. But this context changed a lot, and now the Internet is the mainly platform for news stories, not only at *Estadão*, but at the majority of newspapers. Many content is available exclusively online. Some newspapers do not circulate in print form anymore. The Latin American news media landscape has undergone significant changes in recent years due to the widespread adoption of digital technologies. Being aware of digital preservation threats are mandatory for acknowledging future media studies.

Restricted access to digital news archives

News articles are an essential tool for studying social science, documenting the events and ideas that shape our societies. As our world becomes increasingly digital, news outlets are shifting towards online platforms, which may make it harder for researchers to access historical data.

As we pointed, access to digital news archives may be restricted due to copyright restrictions, limited resources, privacy concerns, legal issues, business models and technical limitations.

Access to news archives is at risk due to a range of factors, including funding cuts, technological changes, and business decisions by news organizations. Funding cuts have impacted public archives and libraries, making it difficult to preserve and provide access to historical records. Technological changes have made it easier to store and provide access to digital archives, but they have also created new challenges, such as the need for ongoing digital preservation and the obsolescence of older technologies. Business decisions by news organizations, such as the closure of newspapers or the decision to focus on current news rather than archival content, can also impact access to news archives.

Another challenge presented by the shift towards digital news is the issue of access. In many cases, digital news outlets are hidden behind paywalls or other restrictions, making it difficult for researchers to access and analyze historical data. This can be particularly problematic for researchers working with limited budgets or resources.

Copyright restricts may apply since news organizations could restrict access to their archives to protect their intellectual property rights, holding their content for their own use or controlling who and for what purposes that material can be used.

Digitizing and maintaining an archive can be expensive and time-consuming, especially for smaller news organizations. Limited resources may make it difficult for them to provide full access to their archives.

Although it is not very common, another explanation for the restriction of access is that news archives may contain sensitive or personal information about individuals, and news organizations may restrict access to protect their privacy. Also, in some cases, access to news archives may be restricted due to legal issues such as court orders, confidentiality agreements, or national security concerns.

The main fact, whatsoever, remains aggregated about business models. Some news organizations may restrict access to their archives in order to monetize their content through paywalls or subscription models. That is the case about Brazilian examples as outlets charge for non-subscribers access to their content.

Technical limitations are, without any doubt, between the main factors that can limit the access to digital news: older archives may be in formats that are not compatible with modern technology, which can make it difficult to access and digitize the content.

Overall, the reasons for restricting access to digital news archives can be complex and multifaceted and may vary depending on the specific news organization and the content of the archives.

Before speaking of the consequences of restricted access to digital news archives, it is important to mention an issue that is at the heart of online journalism.

Digital news reliability

Researchers are now relying on online platforms to access data that was once found in physical outlets such as newspapers and journals. While digital platforms may make data more accessible in some cases, they also pose significant challenges to researchers.

For example, one significant issue that arises when accessing digital news archives is that of reliability. Physical newspapers and journals were generally considered reliable sources of information, and researchers could trust that the information contained within was accurate. However, in the digital age, where news outlets are competing for clicks and views, accuracy is not always a top priority.

There is another question: Since a published story may have several daily updates, will it be possible to distinguish those different versions? The dynamic nature of online news presents further challenges. Since a published story may have several daily updates, researchers must find a way to distinguish between different versions of the same story accurately. The dynamic nature of online news presents another challenge for researchers. As I mentioned earlier, news articles are often updated throughout the day, with new information added or old information revised. While this may be useful for journalists and readers seeking the most up-to-date information, it makes it difficult for researchers to distinguish between different versions of the same story accurately.

Physical newspapers and journals were often considered to be reliable sources of information, but the same cannot always be said for digital news outlets. In some cases, digital news outlets may prioritize sensationalism or clicks over accuracy, leading to misinformation and incomplete or inaccurate data.

News articles are often updated throughout the day, making it difficult for researchers to distinguish between different versions of the same story. This can lead to inaccuracies and inconsistencies in research findings.

Another concern is whether digital news will compromise journalistic and historical research. While digital news may provide more convenient access to information, it may also result in a loss of the original context and nuances of the article. Additionally, there may be a risk of digital news being altered or deleted, which could impact the accuracy and reliability of research findings.

Another issue with digital news is the potential for alteration or deletion. Unlike physical newspapers, digital content can be easily altered or deleted, which can impact the accuracy and reliability of research findings. To mitigate this risk, researchers must be careful to document the version of an article they are using and ensure that the data they are using is as reliable as possible.

Consequences of restricted access to digital news archives

Losing access to news archives can have significant consequences for various stakeholders. Historians and researchers may lose access to critical historical records, limiting their ability to understand the context of past events and the evolution of media practices. Journalists may lose valuable resources for investigating and reporting on current issues, limiting their ability to provide in-depth analysis and commentary. Educators may lose valuable teaching resources, limiting their ability to provide students with insights into media practices and cultural norms. Finally, the public may lose access to information that is essential for informed decision-making, limiting their ability to participate in public discourse and hold institutions accountable.

Losing access to news archives has a significant impact on research, journalism, and public discourse. Research into historical events and trends is limited without access to news archives, making it difficult to understand the context of current affairs. Journalism is also impacted, as journalists rely on news archives for background and context on current events. Without access to news archives, journalists may not be able to provide accurate or comprehensive reporting. Public discourse is also impacted, as access to news archives is essential for citizens to hold those in power accountable.

Restricted access to digital news archives can have significant consequences for society and individuals alike. The first impact is the loss of historical memory. As news archives provide a record of the past, including major events, social trends, and cultural changes, restricting access to these archives means that future generations may not have access to this information, leading to a loss of historical memory and understanding. Even though some news archives allow researchers and public in general to consult older editions, they do not permit they follow facts when they happen, unless as a subscriber.

Access to digital news archives can help individuals and organizations verify information and fact-check claims made by individuals or groups. Without this access, it can be difficult to distinguish between accurate and inaccurate information, which can lead to misinformation and confusion.

Another general issue is related with the lack of accountability. Online news archives can serve as a record of the actions of governments, businesses, and other organizations. Restricting access to these archives can make it easier for these entities to avoid accountability for their actions, as there will be less information available to the public.

Digital news archives provide a wealth of primary sources for humanities researchers, including history, literature, cultural studies and, of course, media

studies, to analyze and interpret. Access-restrictions limits the scope and depth of research that can be conducted, leading to a narrower understanding of historical and cultural events.

Besides, digital news archives provide a record of events as they were reported at the time, which can differ from the way events are remembered or interpreted later on. Restricting access to these archives can limit the ability of humanities researchers to understand how events were perceived and reported at the time, leading to incomplete or inaccurate narratives. By the same way, news can provide insight into cultural trends and attitudes, such as changing views on social issues or popular culture. If researchers have limited access to news archives, it may interfere with the ability of humanities researchers to analyze and understand these trends, leading to a limited understanding of cultural shifts over time.

Another point is that many humanities researchers rely on digital news archives to study events and trends that may have significant social and political implications. If they are not able to access the whole content of newspapers, they may miss important articles or views, raising ethical concerns around censorship and academic freedom.

But mainly, the first and biggest concern is about the media collection preservation. Facing so many challenges, media outlets may be not so worried about collecting and guarding their production towards future researchers. The scenario includes the social media menaces to the extent of traditional news outlets but also how society relates with journalism itself. That framework leads us to examine the question of digital archives preservation for history.

Final notes

As we have seen, the transition towards digital media is changing the way we access and consume news, and this is having a significant impact on humanities research. Historically, newspapers and other print media outlets were the primary sources of information for researchers studying media history. These physical outlets were archived in libraries and other institutions, making it relatively easy for researchers to access the material they needed. Those archives are generally free access, so the information circulates publicly.

However, as more and more news outlets shift towards digital platforms, it is becoming increasingly challenging to access historical news archives. This is because digital content is often subject to paywalls or other forms of content restrictions, which can make it difficult or even impossible for researchers

to access the material they need. Additionally, digital content is often stored on servers that are subject to data breaches or other security vulnerabilities, which can put historical news archives at risk.

Furthermore, it is essential that news outlets themselves take responsibility for preserving their digital archives. This can be accomplished, for instance, through partnerships with digital libraries and archives or by developing their own internal systems for archiving and preserving historical data.

One potential solution to this problem is the development of open-access digital news archives. These archives would be freely accessible to researchers and journalists, ensuring that everyone has access to the same historical data. While this solution may require a significant investment of resources, it would ultimately benefit society by promoting transparency and accountability in media.

In conclusion, the changing media landscape poses significant challenges for humanities research. This article aimed to approach the potential loss of access to news archives, due to many factors, and its impact on research and journalism itself, focusing on Brazilian examples. The paper did not intend to propose ways of solving this problem, since it would demand an in-dept discussion. The goal was to point out some issues about news preservation.

As we move towards a more digital environment, it is crucial to ensure that historical news archives remain accessible to researchers, while also maintaining the reliability and accuracy of the data. Restricted access to digital news archives can have severe consequences for humanities research, including limitations on the availability of primary sources, incomplete narratives, impacts on cultural analysis, challenges to interdisciplinary research, and ethical considerations. Therefore, it is important to maintain access to these archives to ensure that researchers can continue to access valuable information and insights from the past.

References

- ARELLANO, M. A. Preservação de documentos digitais. *Ciência da Informação*, Brasília, v. 33, n. 3, p. 15-27, maio/ago. 2004. Disponível em: <https://revista.ibict.br/ciinf/article/view/1043/1113>.
- CARVALHO, B. L. P. de. Digitalização de jornais: uma reflexão sobre desafios e melhores práticas. *Acervo: Revista do Arquivo Nacional*, Rio de Janeiro, v. 29, n. 2, p. 89-102, jul./dec. 2016. Disponível em: <https://revista.arquivonacional.gov.br/index.php/revistaacervo/article/view/721/730>.
- EDWARDES, E. Ephemeral to enduring: the internet archive and its role in preserving digital media. *Information Technology and Libraries*, v. 23, n. 1, p. 3-6, 2004.

- FORMENTON, D.; GRACIOSO, L. de S. Preservação digital: desafios, requisitos, estratégias e produção científica. *RDBCI: Revista Digital de Biblioeconomia e Ciência da Informação*, v. 18, 2020. Disponível em: <https://doi.org/10.20396/rdbci.v18i0.8659259>.
- FUNDAÇÃO BIBLIOTECA NACIONAL. Hemeroteca Digital Brasileira. BNDigital Brasil, s.d.a. Disponível em: <http://bndigital.bn.gov.br/hemeroteca-digital/>.
- FUNDAÇÃO BIBLIOTECA NACIONAL. Sobre a BNDigital. BNDigital Brasil, s.d.b. Disponível em: <https://bndigital.bn.gov.br/sobre-a-bndigital/historico/>.
- GOOGLE NEWS. *Jornal do Brasil*. 2011. Disponível em: <https://news.google.com/newspapers?nid=oqX8s2kIRwC>.
- GOOGLE NEWS INITIATIVE. 2021 *Impact report: advancing the long-term sustainability of journalism and publishing*. 2022. Disponível em: <https://newsinitiative.withgoogle.com/impact2021/>.
- LIBRARY OF CONGRESS. National Digital Newspaper Program. Disponível em: <https://www.loc.gov/ndnp/>.
- NATIONAL DIGITAL NEWSPAPER PROGRAM. *About chronicling America*. Chronicling America. Disponível em: <https://chroniclingamerica.loc.gov/about/>.
- OUAKRAT, A.; PATERSON, C.; REBILLARD, F.; SURM, J.; MOREIRA-CESAR, C. International news flows through the lens of platformization. *Sur le journalism*, v. 12, n. 1, p. 126-131, 15 jun. 2023. Disponível em: <https://doi.org/10.25200/SLJ.v12.n1.2023.530>.
- RINGEL, S. Digitizing the paper of record: archiving digital newspapers at the New York Times. *Journalism*, v. 24, n. 2, p. 262-279, Feb. 2023. Disponível em: <https://doi.org/10.1177/14648849211023849>.
- SANTA-CRUZ, L. O repórter como historiador do tempo presente. In: PEREIRA, C. A. M.; ANTONIELI, M. E. (ed.). *Desafios do jornalismo: novas demandas e formação profissional*. s.l.: Appris, 2014.
- SANTA-CRUZ, L. Jornalismo feito de notícias velhas: o uso da memória jornalística na produção de conteúdo novo. *Mídia e Cotidiano*, v. 8, 2016. Disponível em: <http://www.ppgmidiaecotidiano.uff.br/ojs/index.php/Midecot/article/view/247/18>.
- THE GUARDIAN. How to access past articles from the Guardian and Observer archive, 15 nov. 2017. Disponível em: <https://www.theguardian.com/info/2017/jun/26/how-to-access-guardian-and-observer-digital-archive>.
- VENTURA, M. I. de A. Desafios da representação da informação nas plataformas digitais dos arquivos: o caso do arquivo do jornal Público Norte. *Páginas A&B: Arquivos & Bibliotecas*, v. 14, p. 52-73, 2020. Disponível em: <https://doi.org/10.21747/21836671/pag14a5>.
- VERDEGEM, R.; SLATS, J. Practical experiences of the Dutch digital preservation test-bad. *VINE: Very Informal Newsletter on Library Automation*, v. 34, n. 2, p. 56-65, 2004.
- WIGGINS, R. Digital preservation: paradox & promise. *Library Journal*, v. 12, p. 12-16, 2001.

Received on 17 July, 2023

Approved on 24 October, 2023