University of Mostaganem-Algeria
VOL:10 / N°: 03 / (2023) , p. p. 159/178

International Journal of Social Communication
ISSN: 2437 – 1181
EISSN: 2710 – 8139

The Online Archive and the Internet Archive: Challenges and Stakes

BOUCHETA Rahmouna *

ORAN 1 - Ahmed ben Bella University , bouchetarahmouna@yahoo.fr

Reçu le: 18/05/2023 *Accepté le: 13/09/2023* *Publié le: 30/09/2023*

DOI: 10.53284/2120-010-003-011

Abstract:

The present study seeks to identify the Internet Archive, its goals, and the significance it provides to beneficiaries and scholars. It also emphasizes the importance of open archives and their diversity across Arab and European national archive websites that have lately acquired popularity. To investigate the most relevant studies on the topic, the study used a descriptive-analytical approach.

The research revealed that internet archiving had the most important findings. In the face of information security risks and dangers, on the other hand, several firms strive to identify methods, preventative measures, and processes to fight security concerns. Their purpose is to perform the job of information security and safeguarding successfully, with an increased emphasis on assuring ongoing operations and supplying high-quality information. This attempt strengthens the reliability of the information presented. It improves the stability of organizations, allowing them to carry out their role in providing services, the majority of which are now automated.

Keywords: Internet Archive, Repositories, Archivist, Archive Sites, Open Access

* **Corresponding author**

1. INTRODUCTION

Major advancements in information and communication technology occurred in the late twentieth century, resulting in the construction of networks and computers which enabled the quick exchange, copying, use, preservation, and accessibility to information and data.

Database systems were critical in rapidly obtaining and accessing a large volume of digital material, therefore assisting the Open Access Initiative for Archives (OAI) via research initiatives.

Geographic obstacles that previously prevented researchers and users from accessing desired archives were also removed, enabling numerous beneficiaries to participate with the archives.

When archives were accessible on the internet, their importance expanded in response to rising information demand and the advancement of document preservation technologies.

Traditional methods should be modified to meet the needs of today's recipients. Interactive techniques built around hypertext that depend on links have become vital. As archiving websites became more visible on the internet, standards such as the Dublin Core Metadata standard for defining internet-related material and standards for document and information protection and security were established to assure optimum performance.

Problem statement

The Internet's discovery was a remarkable sign in society as a whole, particularly in the area of research and information organizations, however its influence on the archive and its institutions beyond all expectations.

Furthermore, owing to digitalization, certain historical papers or records that the researcher had limited accessibility and had traveled vast distances were visible on the computer screen via the Internet with the click of a mouse.

In contrast, the archivist's passion for utilizing the Internet does not allow him to forget his responsibilities to records and to the preservation of private life, state security, and the non-dissemination of sensitive papers...etc.

As a result, the archivist is confronted with a slew of legal and technological issues. What are the archivist's obstacles, whether they are connected to the internet archive or the online archive, and how can he win the bet?

Study importance

The research is significant for the following factors:

- The study's topic is closely connected to an essential component of online archives and their function in meeting the rising demand for information, as well as the creation of and methods for conserving and exploiting records.
- The Internet archive is important because it assists institutions of different specialties in retaining their balances, attaining a higher position, and generally enhancing the quality of preservation and retrieval.

- The Internet archives' crucial and successful function in providing unfettered access to knowledge without limits or impediments.
- The worldwide trend toward developing open archive repositories owing to its numerous advantages, which include free storage, organization, and distribution to users and researchers via the Internet.

Study objective

Recognizing the Internet Archive as a new system for scientific study and presenting the Internet Archive, its pros and disadvantages.

Controlling the ways of free access expressed in open access journals and open archives.

The Internet Archive is introduced, along with its goals and advantages.

Defining the goals and significance of the open archive for beneficiaries and scholars.

The degree of the magnitude and brightness of current European and Arab national archives.

The Internet Archive's significance on the Internet.

Previous studies

1) Mr. AMANI Mohamed's research is entitled "Archiving Self as a Channel for Knowledge Communication on the Web: **A Study of Its Applications in the Field of Libraries and Information**." Volume 2, No. 29, April 2009. Journal of Arab Libraries and Information. The study seeks to determine the views of researchers in Arab libraries and information departments toward protecting their scientific production on their official and unofficial personal sites, evaluate the challenges and impediments that prevent self-preservation in the Arab world, to determine the general characteristics of scientific production that exist on researchers' sites in the field of libraries and information,and to expose the self-preservation tendencies of Arab electronic periodicals producers in the sphere of libraries and information.

2) Wahid Qaddoura's research titled "**Scientific Communication and Free Access to Scientific Information**" Arab University Researchers and Libraries. Arabic Educational, Cultural, and Scientific Organization, Tunisia, 2006. The study demonstrates the variations generated by the Internet in Tunisian researchers' scientific communication. According to the report, 36.63% of the researchers are aware with the notion of open access, and the majority of them work in medical and biological sciences. In addition, only a small percentage of researchers have published their work in open archives.

3) Younes Ahmed Al Shouabka and Abdalmajid Salah Bouazza's research, "**Attitudes of Faculty Members at the United Arab Emirates University Towards a System of Free Access to Scientific Information."** The Arab Federation for Libraries and Information's Eighteenth Conference, 17-20 November 2007, Jeddah. The research looked at the attitudes of faculty members at the UAE University toward the free access system to scientific information, and the most important findings were that 62.9% of the participants are familiar with the concept of free

accessibility, 68.6% have never published in free access journals, and 31.4% have already published research in those periodicals. According to the results of the study, 82.9% of participants had never published in digital archives, whereas 17.1% had placed research in electronic archives.

4) Abd Al Majid Bouazza's study, "**Attitudes of Arab Researchers Towards Open Archives and Periodicals Available for Free Through the Internet**": Arab faculty members at Sultan Qaboos University as a paradigm. King Fahd National Journal, Volume 12, Number 1 (2007), pp. 145-174. According to the findings of this study, researchers at Sultan Qaboos University have yet to crystallize their knowledge of the notion of open access to scientific information. The Arab researcher at Sultan Qaboos University still does not embrace the open access concept; 78.8% said that they did not publish research in free access publications, and 77.8% were unfamiliar with worldwide free access programs.

Free access to information: a study of academics' use of the Internet to disseminate their intellectual output in Saudi universities . The Arab Federation for Libraries and Information's Eighteenth Conference, held in Jeddah from November 17 to November 20, 2007. The research seeks to show the degree to which professors at Saudi universities in the city of Riyadh are committed to publishing their scientific output and making it freely accessible on the Internet. The study indicated that a significant percentage of participants support this trend, with 63.4% of the total study participants expressing a desire to have scientific production published on the Internet in the future, in comparison with 22.4% who did not support the idea of publishing a scientific production. The researcher discussed the archive, its sections, ages, physical form and preservation, arrangements, challenges of the archive in the information age, the effect of technology on vocational training demands, and the importance of the Internet in changing the reality of dealing with national archives.

5) Ashraf Mohammed Abdo's research titled "**Electronic archive: theoretical foundations and practical applications**." House of Books and Arabic Studies, Alexandria, 2017. 565 pp. Its goal was to define a collection of ideas related to archive science, electronic document management, and digitalization. From an archiving standpoint, the author also encouraged how to establish an electronic archiving system in businesses and institutions, as well as a recommended idea for developing an automated program for documents and files.

2. INTERNET ARCHIVE

Definition

The concept for the Internet Archive originated with the American scientist (Kahle), who is one of the most well-known Internet academics. In 1996 AD, he founded the Internet Archive, an organization dedicated to preserving and archiving Internet pages. However, what exactly is the Internet Archive? Some experts worked hard to develop the following definitions:

It is a virtual time machine that serves to preserve Internet sites from disappearance and non-appearance difficulties, as it is regarded as an essential medium in preservation, blogging, and many historical objectives.

The Internet Archive is the stored memory of all web pages published on the Internet since its establishment in 1996.

It is characterized as an archive founded in 1996 with the goal of developing a digital library on the web to provide a continuous and free access service for research, study, and historical reasons, as well as for general use. . (Khalifa, 2005, p. 107)

Its concept

The Internet Archive is an electronic entry point for the (Internet Archive Organization), a non-profit organization that is held accountable. It was created in 1996 and is based in the United States state of California, specifically in the city of San Francisco. A wide range of governmental and non-governmental bodies, organizations, and institutions back it up. The Internet Archive is a non-profit organization dedicated to the preservation of all Internet content. It archives practically all sites and for almost all days of the year, whether it be movies, music, texts, images, books, or even a time machine for sites, which is a method through which any site may know how it looked in a prior period of time.

According to the popularity and significance of the website, it is pretty handy for those who want to view what a website looked like five years ago. You may even use the gadget to retrieve a recently deleted topic. Unlike Google's constantly changing archive, its permanent archive remains untouched.

3. INTERNET ARCHIVE FEATURES

- The Internet Archive includes all Internet sites in the world, with the exception of a handful that it does not archive, and it is not constrained by subjective, geographic, technological, or language boundaries.
- The Internet Archive offers an excellent method for searching its contents, using a program called the "Wayback Machine," which is the most effective method for searching since it allows you to search by URL.
- The Internet Archive builds an intellectual property rights policy that respects website owners' right to have their websites removed from the archive.
- The Internet Archive is regarded as the United States of America's national archive. It conducts certain efforts to archive American websites, with a primary focus on the United States, such as producing an archive of US presidential election websites, archives of state election websites, and archives of September 11 events.

In addition to Internet sites, the Internet Archive restricts the following types of information sources: music, videos, texts and documents, software, and educational resources.

The Internet Archive has a problem in that it does not provide a search engine for the contents of the sites since the existing technology does not operate as well as it should. Because it is feasible to view pornographic websites while using Internet censoring software. The Internet Archive does not give any thorough or succinct bibliographic information on the sites that it stores.

- The archive does not disclose numerous details on the sites it preserves, such as the number of sites, the nature of their domain names, the country they belong to, and so on, instead simply providing information regarding the size of the sites.
- Because the Internet Archive does not give any objective categorization or classification of the sites it provides, the user must be acquainted with the URL in order to successfully search and utilize the archive.
- There is no regulation specifying the timeframes for retaining site copies, which resulted in the retention of numerous copies of static sites that do not talk, while not enough copies were maintained for sites that are continually continually updated, such as newspaper and magazine sites.

4. INTERNET ARCHIVE ROLE AND FEATURES:

- The World Internet Archive is distinguished by a number of features, the most prominent of which are as follows:

- Text preservation: The Internet Archive is a service that saves copies of books and text documents. It is one of the most established services for this objective. It predates Google's (Google Books) service, which was founded in 2004. The Internet Archive offers a number of copy facilities for books that are particularly equipped and dispersed for this aim. It covers five different nations, and the copying pace reaches thousands of volumes per day, providing researchers with a digital library without jeopardizing the content that is stored in an environment ideal for preservation.

Another benefit of the Internet Archive is that it only provides entire versions of books that are not copyrighted and are in the public field. They often include information that you do not need to view. There are also no unpleasant "Snibit" pages that reveal just the book's information but none of its contents.

The majority of the content stored in the Internet Archive is accessible in many file formats or extensions, giving users additional choices for reading or browsing on various devices.

- Searching the Internet Archive is simple because you may filter your search depending on the kind and extent of content you want.
- The Wayback Machine is one of my preferred aspects of the Internet Archive. It allows you to view what many websites looked like in the past.
- The Internet Archive offers a free virtual library card. One benefit is that after establishing the account, you can then add tags of objects inside the group that you may utilize to

search. You may also sign up for the Internet Archive's monthly newsletter, that will continue to keep us up to date on new additions to the archived collections.

5. WEB ARCHIVE... SAVE INTERNET MEMORY

Preamble

The globe entered the twenty-first century via its broad gateway, bringing with it new cultural practices, diverse desires, and proportionate services to fulfill the requirements of this era and the needs of all users. The Internet, with its limitless capacity and variety of tasks and applications, has managed to occupy the most prominent position in human existence, assisting it in scientific, practical, and personal dimensions. We document our lives practically everyday on social media platforms such as Instagram and Facebook.

By posting words, photographs, and videos and expressing one's thoughts by leaving comments here and there. However, the network's broad, deep, and limitless universe conceals a fragile nature that necessitates a solution that protects the history of the Internet and assists in returning to all of its elements when required. The web archiving method is one solution. (chabin, 2020, p. 139)

Therefore, what exactly is web archiving? And what are the root causes? How is this form of archiving carried out? And who are the most significant entities in the world that do this?

What is web archiving?

The British National Archives describes web archiving as the practice of gathering websites and information from the World Wide Web and preserving them in an archive. Web archiving is a procedure that is identical to conventional paper document archiving in that it involves the selection, storage, preservation, and availability of records.

This procedure enables governments, enterprises, organizations, academics, historians, and the general public to access preserved sites (chabin, 2020, p. 140). Web archives, like conventional archives, are gathered and managed by archivists; they may be referred to as "web archives" in this context.

Why archive the web?

Changing information practices

The online area is expanding in information practices at a rapid speed, and the Internet looks to be the most significant, comprehensive, and responsive information source. The second web caused a shift in Internet usage practices, resulting in interaction transforming into the master of the situation, and the user now adds texts, writes in web pages, represents his opinion, and interacts with other users via numerous programs and applications, as well as text and various forms of media. It may not be accessible outside the network. As a result, it is critical to preserve and archive documents for future reference.

Preserving the digital heritage

With the Internet, a "digital heritage" has emerged. In and of itself, the release of more and more documents in digital form constituted a legacy that needed to be gathered and conserved in order to provide a complete census of the country's cultural output. Deposit laws have been created in various nations throughout the globe to encompass the gathering and preservation of this new form of document. As a result, digital heritage is given the same significance as paper heritage, and web archiving is seen as a crucial need to address the issue of digitization fragility, as the nature of digital documents necessitates serious solutions to prevent their disappearance in all forms, including web pages, texts, and audio-visual documents.. (Chaimbault)

How does web archiving work?

Archivists utilize electronic technologies to gather sites since the web has a large number of sites and information. Using specialized software, web archiving captures websites from their current positions on the live web. This sort of application is known as a crawler because it "roams" between sites and across the internet, copying and saving information as it goes. As part of the Web Archive Collection, archived sites and the information they represent are made accessible online. These sites and the gathered information are also visible, legible, and browsable, just as they were when they were directly on the web, but they are kept in the form of web snapshots of information at a certain moment in time. (Archives., p. 5)

Who are the web archivists?

National libraries, national archives, information organizations, technological groups, and a variety of other organizations. These parties are all involved with the careful preservation of the most significant online material. (Al Sharif A. A., 2015, p. 110). Commercial online archiving solutions and services are now accessible to all companies who need to preserve their web material for business objectives, history preservation, legal reasons, and so on.

6. INTERNET ARCHIVE

A- About the organization

The Internet Archive is one of the most prominent organizations involved with online archiving. . (https://archive.org). On the home page of its website archive.org, it is described as "a non-profit library of millions of books, movies, software, music, websites, and more, all free of charge." Academics, people with disabilities, and the general public have permanent access to historical resources accessible in digital format, and the project aspires to establish the world's biggest digital library.

The organization was established in 1996 and is based in San Francisco. In the beginning, it got data contributions from a variety of sources, most notably Alexa internet. (Al Sharif A. A., 2015, p. 110)

The organization started to develop and expand in 1999, and it continues to do so today, with the most information accessible on the Internet via its website archive.org.

b- Contents of the Internet Archive

On the Internet Archive's website, there are presently 279 billion stored web pages, 11,331,882 books and texts, 3,084,460 animated pictures (video), 3,260,704 audio files, 1,323,000 TV series, and 154,853 results. It allows you to access millions of software, CD images, documentation, and multimedia items, as well as 1,391,735 photographs and 168,837 audio files.

The site offers the possibility to search all its contents, and the ability to evaluate the results and obtain the desired document, of any form. Furthermore, all forms of archived and accessible documents may be searched by title, date of publication, and originator's name. It also gives users access to 129,989 Arabic texts that they have submitted.

Archived materials from libraries all across the globe make up a significant portion of the archive's content. We locate materials from American libraries, Canadian libraries, the University of Toronto's Roberts Library, the California Digital Library, European libraries, and others.

C- Wayback machine:

The Internet Archive provides the Wayback Machine project, which is a tool for searching for archived online pages. It is accessed simply by typing the name of the site from which one of its pages is being searched into the search box, or by entering a term that expresses the researcher's goal. This program summons all archived versions of the same site's content, or the researcher's keyword, in chronological sequence.

In 2002, the Bibliotheca Alexandrina, that maintains a backup copy of the archive, and two similar sites for the Internet Archive are created. The agreement stated that the original Internet Archive will provide to the Bibliotheca Alexandrina the first generation of website archiving equipment in 2002, involving a copy of the Internet Archive, which operated from 1996 to 2001 at the time (Sameh Zeinhom, 2012, p. 650)

The Bibliotheca Alexandrina replaced the original 100 terabyte Internet archiving infrastructure in 2006 with a newer design acquired from the original Internet Archive to store data of up to 1.5 petabytes (1.5 million gigabytes) in 23 holders; at the time, it encompassed a decade of archival material dating back to 1996. In 2008, the library added another 4.1 petabytes, which were made possible by locally manufactured and installed technology. Since then, the capacity has been raised twice more, always using locally acquired hardware. (Sameh Zeinhom, 2012, p. 650)

The archive now has the ability to hold 4.9 petabytes of diverse data; it is utilized not just for archiving the Internet, but also for archiving other digital assets, the majority of which are the library's huge collection of digitized books.

The Internet Archive at the Bibliotheca Alexandrina is accessible via the Wayback Machine via the website (web.archive.bibalex.org), and the collection is extensively viewed locally, regionally, and worldwide; the number of site visitors grows each year.

D- Internet Archive in the Bibliotheca Alexandrina

The Internet collection signed a contract with the Library of Alexandria in 2002, under which the latter holds a backup copy of the collection, (Al Sharif A. A., 2015, p. 200)and two identical sites will be constructed for the Internet Archive. The initial generation of website archiving devices, in addition to a copy of the Internet Archive from 1996 to 2001, were presented to the Bibliotheca Alexandrina.

The Internet Archive is now regarded the first of its type outside of the United States of America, with a capacity of 4.9 petabytes of diversified data. It is used for archiving Internet and other digital information, primarily the library's huge collection of digitized books. The Internet Archive for the Bibliotheca Alexandrina may be found at archive.bibalex.org. (Projects and Activities, 2017)

According to MacArthur, an employee of the Internet Archive, digitization is the answer today for protecting the global knowledge history, and that preservation must also involve saving digital pages from disappearance. (Internet archives, 2016)As a result, it is vital to concentrate on eliminating barriers to knowledge digitization, not necessarily as an alternative to what it is today, however as a solution that saves it from extinction by overcoming financial, legal, and technological barriers.

It will unavoidably take libraries and all those concerned with knowledge and information into the digital era, where they will strive to conserve the knowledge legacy created by the human intellect in our time in all contemporary methods.

7. PRACTICAL USES OF THE INTERNET ARCHIVE

Checking citations from internet sources

Internet sites represent one of the sources of information, and also the specialization of libraries and information has addressed it from this perspective, resulting in the establishment of bibliographic description guidelines for Internet resources. Today, many libraries collect and index Internet sites, much as they do books and periodicals. (Khalifa, 2005, p. 115)

Since the Internet has developed into a source of information, it is normal for researchers to rely on it in preparing their studies and research, as well as for master's and doctoral students, where researchers draft reference citations to Internet sources in accordance with the rules in effect, and it is no longer an area of explanation. The problem arises when it becomes impossible to check the reference owing to probable changes in the specified location or the contents of the page, without any accessible resource referred to in the citation. The Internet Archive comes into play here, since it can be searched for existing news and retrieved as it was on the Internet on the date specified by the researcher inside the data.

Website Indexing

A growing number of libraries are now indexing Internet resources as one of its collections. Of course, the requirements for bibliographic description have been devised to assist libraries in

indexing Internet sites. Without getting into the standards for defining Internet sites in detail, one of the most significant description areas used for indexing sites is the availability domain. (Yusriya, 1999)

In which the address of the site is provided, which is known as the URL (Uniform (Resource Locator), and for machine-readable indexing, the MARC 21 standard has allotted field No. 856 in which to record the address of the site.

What we are concerned with here is machine-readable indexing, that could be used by the Internet Archive for connecting the beneficiary to old copies of the site in the Internet Archive by employing the local fields provided by Mark, which are free to be utilized according to the needs of each library, and thus one of these fields can be used and an address is written in it. The site is listed in the Internet Archive as follows:

If the website address is

http://www.darelkotob.org.eg

Its address will be in the Internet Archive:

http://web.archive.org/web/*/http://www.darelkotob.org.eg

http://web.archive.org/web

Then we enter the address of the site whose address we wish to indicate in the Internet Archive, for example:

www.alexa.com

Its address in the Internet Archive is:

http://web.archive.org/web/*/http://www.alexa.com

To begin, we define the web as a collection of digital materials on a certain subject that are accessible in print or on the Internet.Subject Gateway

The word webography has just recently been employed in Arab intellectual creation, first appearing in an essay by Abdel Rahman Farraj. (Farraj, Libraries and digital collections in directories and papers available on the World Wide Web, 2002)

In 2002, he limited Internet guides to the issue of digital libraries, and the word was utilized in SherifShaheen's renowned metadata research the same year. (Shaheen, 2002)In 2004, a specialized electronic periodical (The cybrarians website has issued a specialized electronic journal in the field of libraries and information called)was published, and one of its chapters was titled "Webbiographical," and it is dedicated to publishing webographic lists on a specific topic in each issue.

Concerning the Internet Archive's participation in compiling webographic listings, we have explained that among the data denied about digital sources is the URL address, which is susceptible to change. As a result, if the URL changes and the page to be referred to is no longer available, its address is recorded in the Internet Archive, for example, if the title of one of the articles was: Www.w3.org/TR/2002/WD-xpath20-20021115

Its direct address will be in the Internet Archive:
http://web.archive.org/web/20030210052440/www.w3.org/TR/2002ND-xpath 20 -20021115

Thus, we may have surmounted the difficulty of the page's disappearance, since we refer to the title of the article in the Internet Archive on the site, despite the fact that its original URL no longer exists.

History and evaluation of websites

Many studies and research in all disciplines have been conducted on the evaluation of Internet sites, and examples in our specialization include studies on evaluating the sites of various types of libraries, such as public libraries, academic libraries, specialized libraries, and school libraries, as well as evaluating search engines; all of these studies present the Internet Archive. It includes the primary tool, which allows you to identify multiple versions of the same site and so compare an old version to a newer one, as well as track the evolution of the site over time in terms of content, form, and design.

Website design uses

These applications are mainly aimed at site creators and administrators, and we demonstrate several areas that may benefit from the Internet Archive.

1. When the site address is changed, the old site is utilized to lead visitors to the new URL. The visitor may also be referred to the site URL in the Internet Archive to identify the previous site.
2. By referring to the many copies that the Internet Archive keeps from the site Arabic Archives on the Internet, the Internet Archive may remove the need for a separate archive for each site, since the Internet Archive may be utilized and regarded the site's archive.

The Arab nations are among those with a large number of documents and manuscripts stored in state archives. Arab countries have started to collect and organize this heritage, as well as provide information about it on the Internet, with the goal to educate current and future generations about their country's history, and to make this heritage available to all after it was previously limited to a group of heritage researchers.

Technological advancements have assisted archive devices and institutions in departing from their old traditional role and transforming them into centers of radiating facts and information, and they are required to work according to various directions within a framework of coordination and methodology that starts with the raw material (A document created in state agencies and temporarily settled in preservation units before being transferred to the final archive, the National Document Arcade). in an integrated and disciplined way, resulting in the development of documented inputs that are easily dealt with by technical methods.

This period, known as the information age, has seen the rise of significant technical breakthroughs, which are represented by three primary technologies: (Kamel, 2017, p. 215)

1-Computer technology, which is distinguished by its exceptional capabilities in gathering, storing, updating, analyzing, and sending massive volumes of archived data at high accuracy and speed.

2- Multimedia technology: Due to the technological advancements of these media, the transformation of archival containers from paper or film mediums to non-traditional media such as magnetic and digital media such as CDs and DVDs that deal with text, image, and sound is currently a reality and urgent.

3- Communication technology: The development of digital communication techniques, as well as the appearance of local and global information networks, aided in the speed and accuracy of the transmission of archival information between institutions and individuals, providing researchers with numerous opportunities to view archival documents, regardless of how far apart their locations are.

In light of the preceding, these archive institutions are in charge of planning and creating sophisticated scientific methods that will allow them to realize and support these aims in the Internet age and widespread use of new technology. These are the plans and strategies: (Kamel, 2017, p. 216)

1- Attempting to upload, at the very least, archival papers of high research interest on electronic or digital information medium using scanners, in order to accomplish a twin purpose of preserving the paper assets of these vessels, protecting them, and rapidly retrieving them upon request.

2- Creating integrated archive databases that comprise search and retrieval procedures, lists, indexes, directories, and indexes that aid in the identification of archival records, their locations, and methods of access.

3- Using global information networks to help beneficiaries access archives by constructing and developing archival websites for those centers that give the aspects of introducing these centers and assisting in the retrieval of their collections of archived documents.

Because of the fast technical changes in the realm of information and communication technologies, Arab archives confront significant problems.

Pages on the Arab National Archives website:

In terms of page count, the richness of the Arab national archives' websites differs substantially. While the Sultanate of Oman's site has over two thousand spider pages, some sites have just one or two pages.

According to Table (1), the total number of archive website pages (according to the search engine Google) is 5852 spider pages. These sites are classified into three categories: those that account for more than 10% of the total, those that account for 1-20%, and those that account for less than 1% of the total. The archives of the Sultanate of Oman, the Emirates, and Bahrain are in the top category, with the Sultanate of Oman's website accounting for over two-thirds of al

archive pages. In general, this reflects the depth of content that distinguishes the archives of these three nations.

Table 1. the total number of archive website pages

N	Nation	Language	Domain	The national archive Name	How effective is the site?	Availability on social media
1	Jordan	Arabic – English	www.nl.gov.jo/Ar/archivedirectorate.aspxmy#cont	Directorate of Documentation and Documentation of the National Library Department	×	FB / Youtube Linkedin / Instagram
2	UAE		www.na.ae	The national archive	✓	FB/ YT/ TWITTER
3	AlBahrin		www.iccc.gov.bh	Aissa cultural Center	✓	FB
4	Tunisia	Arabic – English - French	www.archives.nat.tn	National Archive of Tunisia	✓	
5	Algeria	×	http://www.archives-gdan.gov.dz	Algeria Archive	×	FB
6	Comoros	French	www.cndrs-comores.org/index.php/documentation/archives	National Archive	✓	
7	Jibouti	×	×	National Archive	×	FB
8	Saudia Arabia	Arabic-English	Ncar.gov.sa	National Center for Documents and Archives	✓	FB. YT. TWITTER
9	Oman	١	www.nraa.gov.om	National Records and Archives Authority	✓	FaceBook Youtbe Linkedin Instagram
10	Sudan	×	http://www.nro.gov.sd	Sudanese National Records House	×	FACEBOK
11	Syria	Arabic	Www.dgam.gov.sy/index.php?p=257	Historical Documents Center		
12	Soumalia	×	×	National Archive	×	
13	Irak	Arabic	https://www.iraqnla-iq.com	National Books and Archives Center	×	FB YT
14	Palestine		www.pnac.pna.ps/ar_new/index.php?p=home	National Archive	✓	FB
15	Qatar	×	×	Department of Documents and Research at the AmiriDiwan	×	
16	Kuweit	×	http://www.da.gov.kw/ara/historiccenter/historix/doc-center.php	Documents Center and AmiriDiwanLibraries	×	
17	Lebanon	×	http://www.can.gov.lb	National Archives Foundation	×	
18	Libya	Arabic	www.libsc.org.ly/mrkaz	Libyan Center for Archives and Historical Studies	×	Facebook
19	Egypt	×	http://www.nationalarchives.gov.eg/nae/ar/home.jsp	National Archives	×	
20	Morocco	Arabic French	http://www.archivesdumaroc.ma	Morocco Archives	✓	Facebook youtube
21	Mauritania	×	http://www.archives-mauritanie.mr	National Documents	×	
22	Yemen	×	www.ndc.gov.ye	The National Center for Documents and Archives	×	

Source: (Farraj, 2020, p. 22)

In addition to the Documents Management Section, the Sultanate of Oman's archive website is distinguished by the inclusion of main sections, namely the adequate definition of the Documents Authority, its news, and services, and the media library, which includes scientific

media sources (represented by the Authority's publications). The site allows users to search both the site and the archives included inside it. There are numerous photos and movies on the site, however the amount of rich files (accessible in pdf format) pales in contrast to the number of pages and the authority's same scientific activities.

The Emirates Archives website, which has nearly a quarter of all archives' pages, includes many sections, including the definition of the National Archives, the founding fathers, available jobs, the Emirates Library, a tour of the archive, including texts, documents, and images, and the availability of documentary services, and the media center, which includes electronic publications, including a refereed scientific journal. In general, this site is an archives of the most comprehensive documents accessible in pdf format. (Farraj, 2020, p. 22)

Table 2. The extent of the richness of the Arab national archives' websites according to the number of pages each

National Archives	Number of the sitepages	% of total website pages	Number of rich files available in format Pdf	% of website pages
Oman	2040	34.8	47	2.3
UAE	1360	23.2	132	9.7
Albahrin	775	13.2	5	0.6
Morocco	619	10.5	82	13.2
Libya	526	9	20	3.8
Tunisia	333	5.6	106	30.8
Arabia Saudia	140	2.4	1 (*)	0.7
Palestine	40	0.7	-	-
Syria	16	0.3	-	-
Comoros	2	0.03	-	-
Jordan	1	0.01	-	-
Total	5856			

Source: (Farraj, 2020, p. 23)

We conclude from this study that national archives play a significant part in the social, economic, and political development of any society in the age of information and digital technologies on the Internet, which includes the platforms they contain, such as sites, portals, social networks, Twitter, YouTube, and other media platforms that have spread in recent years.

The prior research, as provided in the tables, revealed that the electronic presence of Arab National Archives sites via the Internet has a severe limitation that differs from nation to country, and this is what came with this previous study.

Criteria for evaluating archives sites on the Internet:

Assessing Internet sites in general is a fundamental and necessary activity for a variety of reasons, the most important of which is determining the degree to which such sites fulfill their primary duties and meet the purposes for which they were created.

There are several approaches and methods for evaluating online archive sites, however these methods may be divided into two categories. The first is assessment studies that concentrate on the site as an information system. System-centered research. (Farraj, 2020, pp. 10-11)

The second is those that focus on the beneficiaries of those sites

User-centered studies.

Each pattern has its own set of methods and approach. However, studies that concentrate on the website as a system mostly use quantitative approaches to assess the website's performance and information retrieval methods. It investigates, for example, the behavior of the beneficiaries in terms of the rates at which they access the site's pages or the various sections on those pages, as well as how they use the search terms in the system or the site, and the characteristics used in developing the system to meet their information needs. The most notable methodological approaches for this approach are the questionnaire and Web Log Analysis.. (Farraj, 2020, p. 14)

In studies that concentrate on the beneficiary, qualitative approaches are used in addition to quantitative ones. To explore, for example, the behaviors of beneficiaries while doing certain searches, the thoughts and choices made throughout the research process, as well as the variables that influence search behaviors. In addition to the questionnaire, this methodology employs personal interviews, content analysis, the so-called Think-aloud Protocols, observation, grounded theory, validity, and relevance.

The last category of studies is known as user studies, which are one of the categories of what is now referred to as information seeking behavior studies in the archiving area.

In the sphere of libraries and information, this is a rather modest phenomena. Until the mid-1980s and early 1990s, archive studies were dominated by preservation rather than access to historical records. Some say that there are few user studies from archives accessible on the Internet, and that there is no common framework for the evaluation components to gather them. (Farraj, 2020, p. 15)

The importance of information security and protection:

The importance of the reasons for document security and protection can be clarified as follows: (Al Sharif A. A., 2015, p. 243)

The need to connect to communication systems and the Internet, as well as the impossibility to separate devices from local and large-scale networks in order to offer information to people in need.

The reliance of diverse organizations on information effectiveness.

Due to the absence of geographical borders while using the Internet and electronic communications, it is difficult to recognize and regulate threats or trace and punish offenders; it gives the possibility to pierce spatial barriers.

The continual increase in electronic uses and applications, as well as the rise of e-government and e-management, all of which need an informational environment.

Securing the safety and protection of documents achieves the following: (Al Sharif A. A., 2015, p. 243)

Privacy and Confidentiality

It is to keep information confidential about everyone except for those who have access to it.

Data Integrity

Entity Authentication: This is the verification of the identity of the entity (person, terminal, calculator, credit card).

Standards for the security and protection of documents and information:

a. ISO standards: The most famous standards affiliated to it and related to the rest of the information, including:

1. ISO 27002: (Al Sharif A. A., 2015, p. 245)

This standard includes some policies and directives, including:

• Security Policy.
• Organization of information security.
• Human resources security.
• Physical and environment security.
• Communication and operation management
• Access control.
• Acquisition, development and maintenance of information systems Acquisition development and maintenance
• Information security incident management
• Business continuity management
• Compliance management.

2. ISO: 27001: (Al Sharif A. A., 2015, p. 246)

This standard introduces a periodic model known as PDCA

Its objective is to identify the requirements for establishing, implementing, operating, monitoring, reviewing, maintaining, improving, and documenting an information security management system within an organization. It is typically applicable to all types of organizations, including

commercial enterprises, government agencies, and others. technology The information and corresponding control goals.

3. ISO 15408:

This standard aids in the evaluation, verification, and certification of security assurances for technology goods, as well as the evaluation of hardware and software to fight climate change in recognized certification labs.

4. ISO 13335:

It consists of a series of principles and guidelines, namely:

- ISO 13335-1: Documentation of concepts and models for ICT security management.

ISO 13335-3: Documentation of techniques for managing information technology security.

- ISO 13335-4: includes the choice of safeguards such as technical security controls.

It is a control and oversight framework that relates information technology to work needs, integrates information technology operations in an acceptable process model, and identifies the primary resources of information technology and the management control goals in which it is awaited.

Conclusion:

The historical significance of archive institutions cannot be overstated, since they have maintained the history of earlier generations, and so the memory of countries is becoming more significant in its varied forms. This period, known as the information age, is also distinguished by the rise of massive technologies, which are represented in three categories.

Main axes:

1- Computer technology.

2- Multimedia technology.

3- Communication technology.

Consequently, all three of these components provided a great opportunity to see archived papers, regardless of their size or location.

Therefore, the information and Internet era was distinguished to have a distinct role and a technical impact on archival centers in safeguarding their documentary heritage in different media, information, electronic and digital media, in addition to designing integrated databases and thus benefiting from global information networks to access archival vessels. By developing and constructing archive webpages for such institutions.

Most European and American national archives have taken advantage of the Internet's benefits and begun the creation of their websites on the Internet in order to reach the greatest number of beneficiaries worldwide.

These nations have also digitized their paper documentary history and made it accessible to the public and scholars on the web, resulting in online access to archives.

With the advent of new applications, such as the usage of social networking services in archives such as (Facebook - Twitter - YouTube...) to boost contact among scholars and therefore utilized these services to obtain additional advantages for researchers in the area of archives.

On the other hand, given the risks and dangers related to information security, many businesses are looking for strategies, preventative measures, and processes to deal with security threats in order to execute the role of information security and protection.

The importance of protecting it is growing so as to guarantee the continuity of work and the quality of the information provided, which would improve the stability of the information provided, which would also improve the stability of organizations to play their role of offering services, most of which are now provided automatically, among the most significant of which are: protection of information at the level of systems and programs.

- Information protection at the individual and system worker levels, in addition to technological techniques of information protection, the most essential of which are as follows:

• Software fire walls.
• Electronic and biological detectors
• IC card.
• Antivirus software.
• Encryption and concealment.
• Intrusion detection tools.
• Network security.
• Provide a backup system.

Margins :

1) Al Sharif, A. A (2015), Web Archives, Dar Al-Gawhara House for Publishing and Distribution, Cairo.
2) Al Sharif, A. A. (2015), Web archives in foreign countries and Gulf countries, Dar Al-Gawhara House for Publishing and Distribution, Cairo.
3) Al-Sharif, A. A. (s.d.), The previous reference, sur https://web-beta.archive.org, Consulté le 09 01, 2019
4) Archives., T. N. (s.d.), Web archiving guidance .
5) Chabin, M.-A. (2020), Archive management Paris: Hermesscience publications.
6) Chaimbault, T. (s.d.), Web archiving, on Ensib: http://www.enssib.fr/bibliotheque-numerique/documents/1730-l-archivage-du-web.pdf, Accessed 12 28, 2019.
7) Farraj, A. R. (2002), Libraries and digital collections in directories and papers available on the World Wide Web. Recent trends in libraries and information (17), pp. 293-299.
8) Farraj, A. R. (2020), The spider application of the archives: an evaluation study of the Arab national archives on the Internet in the light of the method of spider measurements. International Journal of Library and Information Sciences, 7 (2).
9) https://archive.org. (s.d.). Consulté le 01 09, 2019.
10) Internet archives. (2016), sur youtube:https://www.youtube.com/watch?v=5eMqwgFEavA, Consulté le 12 28, 2019.
11) Kamel, O. (2017), Archiving systems, Dar Al-Wafaa for Printing and Publishing, Alexandria.
12) Khalifa, M. A.-S. (2005), Internet Archive: A Study of Journals of Scientific Use. Journal of Arab Libraries and Information, 25 (03), p. 107.
13) Projects and Activities, I. a. (2017), Internet Archive, sur Bibliotheca Alexandrina website: http://www.bibalex.org/ar/Project/Details?DocumentID=283&Keywords, Consulté le 12 28, 2019.
14) Sameh Zeinhom, A. g. (2012), Digital Libraries and Archives. Tass Printing Company, Cairo.
15) Shaheen, S. K. (2002), Metadata as a resource for cataloging standard records for Arabic online electronic information sources: an exploratory pilot study. Recent trends in libraries and information (18), pp. 79-146.
16) The cybrarians website has issued a specialized electronic journal in the field of libraries and information called. (s.d.)., sur cybrarians journal: http://www.cybrarians.info/journal/, Consulté le 12 29, 2019.
17) Yusriya, A. H. (1999). Electronic Documents on the Internet: An International Attempt to Codify Bibliography. Recent trends in libraries and information (12), pp. 69-85.