RA05 Phase 1 - Report

Study title	Users' approaches and behaviors in accessing records and archives in the perspective of AI: a global user study Phase 1 – Italian study
Working group code	RA05
Status	Final
Version	1.1
Writers	Pierluigi Feliciati, University of Macerata, Italy, pierluigi.feliciati@unimc.it, study leader and Giorgia Di Marcantonio, University of Macerata, Italy, g.dimarcantonio@unimc.it
Date	April 2023

Document control

Version history			
Version	**Date**	**By**	**Notes**
1.0	March 2023	Pierluigi Feliciati & Giorgia Di Marcantonio	Final version for dissemination
1.1	April 2023	Corinne Rogers	Minor copy edits

Introduction

This study was organized within the International research project InterPARES Trust AI, whose P.I. is prof. Luciana Duranti, luciana.duranti@ubc.ca , with the Co-Direction of prof. Muhammad Abdul-Mageed, muhammad.mageed@ubc.ca, both at the University of British Columbia, Vancouver, Canada.

The study RA05 - *Users' approaches and behaviors in accessing records and archives in the perspective of AI: a global user study*, is coordinated by Pierluigi Feliciati, University of Macerata, Italy, and carried out by a team composed of: Jessica Bushey, San José State University, California, USA, Giorgia Di Marcantonio, University of Macerata, Italy, Darra Hoffmann, San José State University, California, USA, Lorette Jacobs, University of South Africa, Tshepho Mosweu, University of Botswana, Adele Torrance, Ingenium, Canada.

It aims to gather users' opinions of existing digital archival services and their experience of and expectations from, and possible misgivings about artificial intelligence-based solutions. The goal of the study is to analyze users' views to support the development of guidelines for the development of functional and reliable AI tools to improve the search and retrieval of archival documents.

In detail, the field research aims to produce more evidence about how the functions within reference and access in adopting AI support could be articulated and match those metrics with data from real users. Data is lacking internationally on the actual User Experience of accessing records and archives and no shared methodologies have been accepted worldwide. Without shared protocols and metrics, users' behavior and satisfaction (quality of access) studies are typically undertaken within specific services in their local context. How much do we know about how digital archival users perform their research? Do they use personal names? Places? Dates? Functions? Subjects? Are they comfortable with the language of interfaces and records? Are they willing to share their research data to improve archival services by adopting AI tools?
The data collected by involving a sample of final users could give an idea about their satisfaction with existing digital archival reference and access services and their actual awareness, expectations and concerns about adopting AI tools to reference and access archival records.
The study, started in the first weeks of 2022, was integrated with the collection of an updated bibliography to document the state-of-the-art about digital archival users' studies.

Methodology and analyzed sample

The initial study protocol, released on February 2022 was based on the direct collection of users' experience (UX) against archival digital reference services. In order to gather those UXs, focus groups and interviews were the methodologies adopted. The team discussed the protocol and the procedure and decided that the direct method would not be sustainable because it was too expensive in terms of time and travelling. The coordinator switched to the method of an online survey, an indirect methodology. Considering the limits due to the lack of direct interaction with real people and the loss of some nuances in the respondents' answers, the survey method may guarantee increased efficiency in the procedures and more data to analyze. The focus group protocol was adapted to the renewed perspective and shared again with the team (June 2022). Another issue considered was the possibility of launching several surveys in different parts of the world with a test of the first version of the questionnaire. After this discussion, there was a shift from a global approach to data collection to a local one. The Italian sub-team decided to translate the survey protocol into Italian and test its efficacy in a specific context.

The Italian online survey was launched on October 10th, 2022 in collaboration with four Italian State Archives: Turin, Milan, Venetia and Ascoli Piceno, and closed on December 11th. Together with the survey, to support Archives in engaging their selected users, we prepared a poster to be printed and posted in the reading rooms and a template letter to be sent to users' mailing lists, if existing.
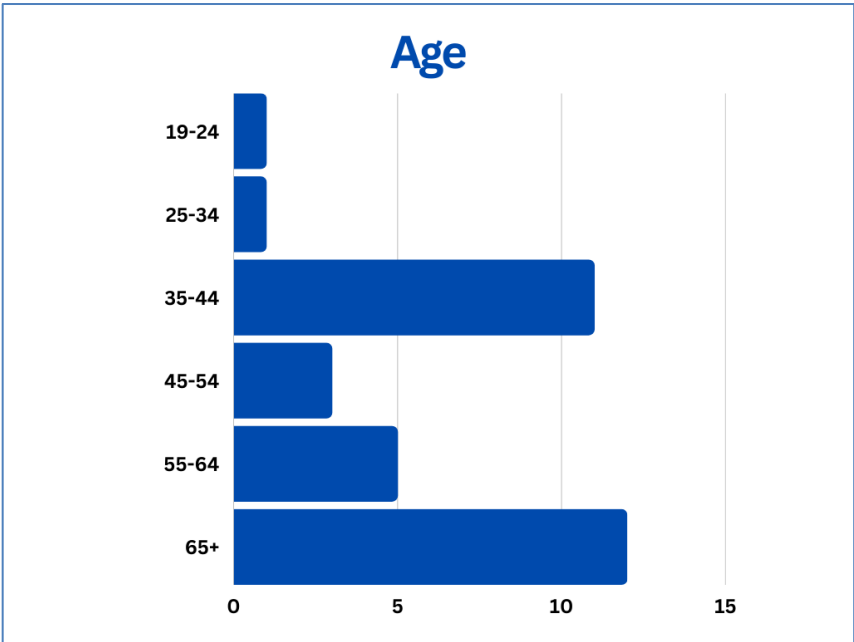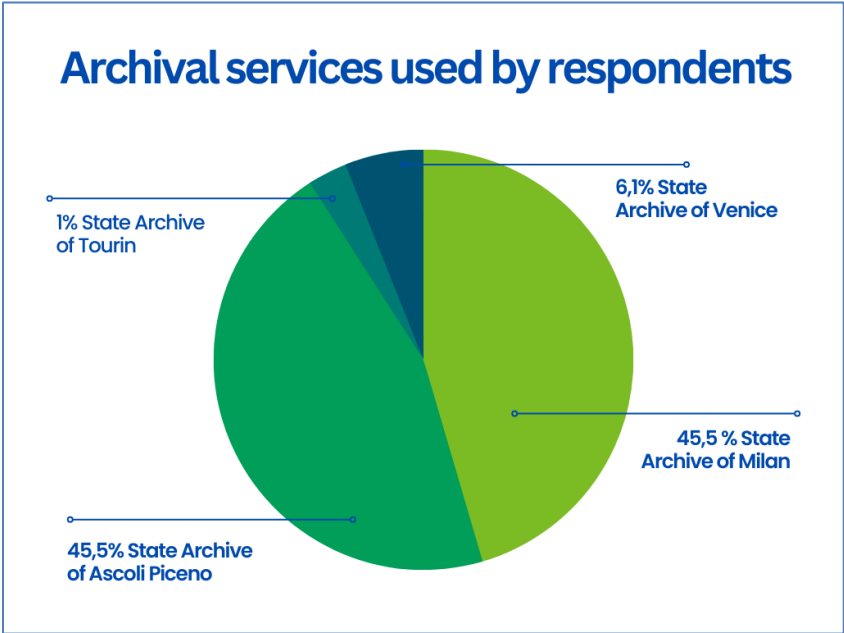
All the case studies offer digital reference and access services to their users, but we asked to focus on the following:
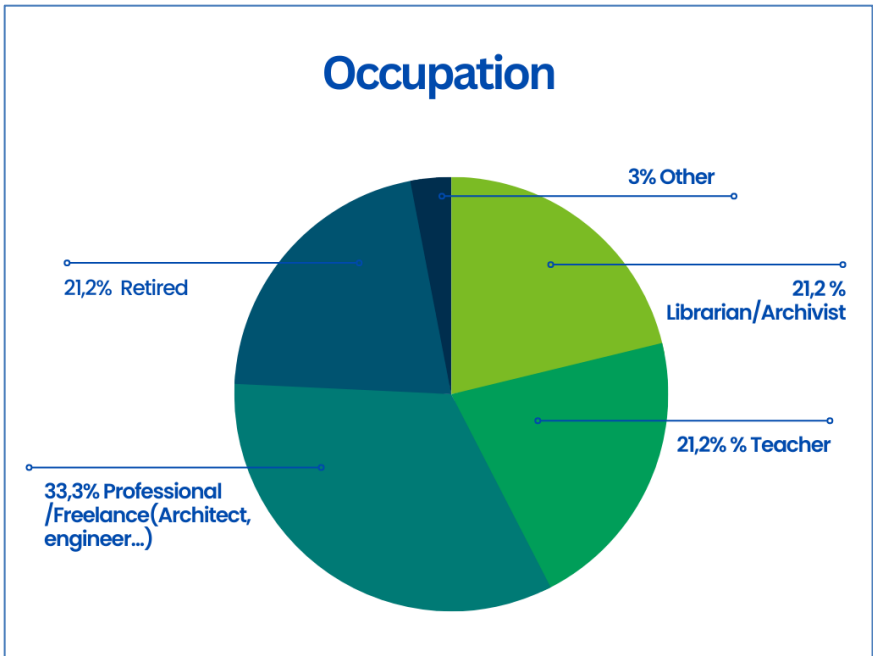
- Archivio di Stato di Ascoli Piceno - Cementi – Servizi Online (online descriptive database of 10.543 files (1961 - 1975) from Prefettura, *Cemento armato* series, regarding official building projects using reinforced concrete. After the 2016 earthquake, to obtain the re-building authorization, the request for such files increased; See: https://www.archiviodistatoap.it/cementi-servizi-online/;
- Archivio di Stato di Milano – Riproduzioni digitali in rete e Banche dati (database of ~50 archival funds including Prefettura, *Ufficio controllo opere in cemento armato*, 35677 files, 1947-1990); See: https://archiviodistatomilano.cultura.gov.it/patrimonio/banche-dati;
- Archivio di Stato di Torino - Sala di studio virtuale (virtual desk - database and digital library covering most of the archives held, including Prefettura, *Pratiche cemento armato*, 48.000 files, 1955-1971); See: https://archiviodistatotorino.beniculturali.it/sala-studio-virtuale/;
- Archivio di Stato di Venezia - moreveneto : il sistema informativo dell'Archivio di Stato di Venezia (general archival holdings portal); See: https://asve.arianna4.cloud/.

The survey was released through Google forms service and kept active till the first week of December.

Findings

The questionnaire received the highest number of responses from the State Archives of Milan (45.5%) and the State Archives of Ascoli Piceno (45.5%). A smaller proportion of responses were received from the State Archives of Venice (approximately 6.1%) and the State Archives of Turin (1%). The first section of the questionnaire aimed to gather personal information about the users, the majority of whom were archivists or librarians, teachers, and professionals from various fields, such as architecture and engineering. The survey participants were predominantly aged between 35-44 and over 65.

Archival services used by respondents

6,1% State Archive of Venice

1% State Archive of Tourin

45,5 % State Archive of Milan

45,5% State Archive of Ascoli Piceno

Age

- 19-24
- 25-34
- 35-44
- 45-54
- 55-64
- 65+

0   5   10   15

Occupation

3% Other
21,2% Librarian/Archivist
21,2% Teacher
33,3% Professional/Freelance(Architect, engineer...)
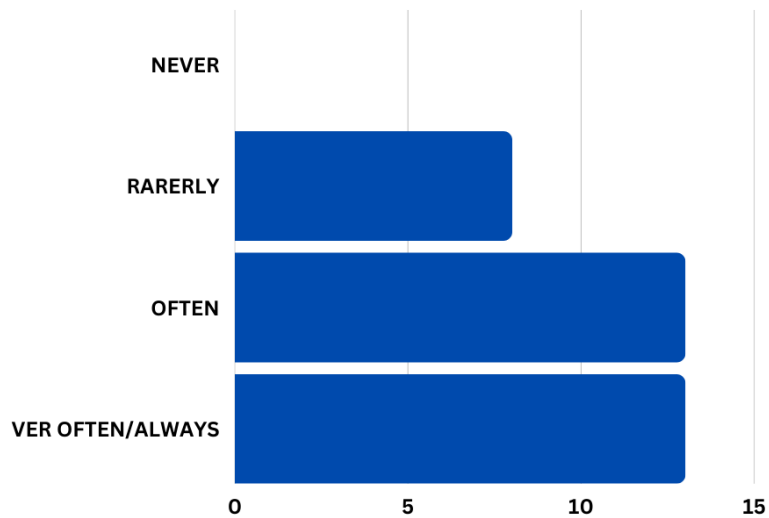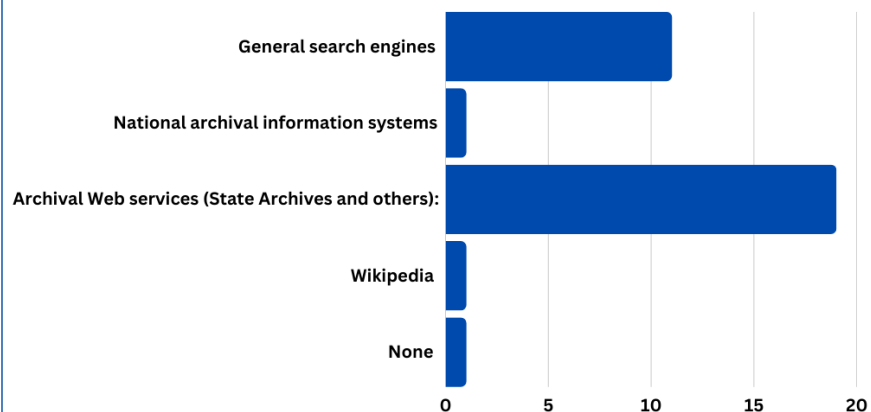21,2% Retired

The second section of the questionnaire was designed to investigate how users conduct online information searches. The responses revealed that most users rely on online search engines, with only a small proportion using advanced search tools. These findings are consistent with the results of the first responses of the survey, indicating that the majority of users adopt general search engines. Specifically, over 87% of respondents reported using keyword phrases (e.g., "Archives of financial office") when conducting online searches, while more than 81% of surveyed users make good use of Google Image search. This trend extends to searches for specific archival information, with most users opting for general search engines rather than National Archival Systems. However, more specialized users prefer having recourse to State Archives websites directly.


How often do you search information online?

NEVER
RARERLY
OFTEN
VER OFTEN/ALWAYS

0   5   10   15   20
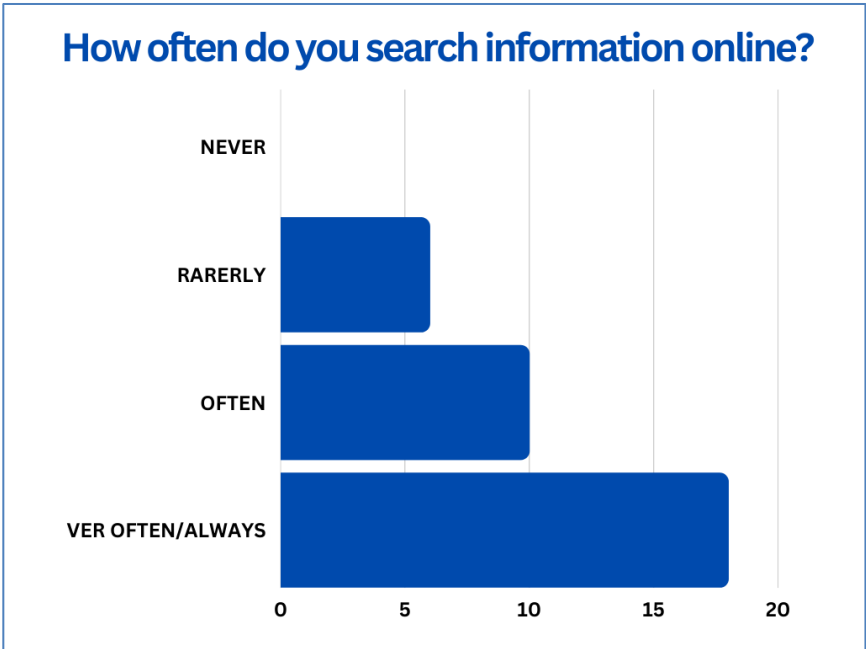
How often do you use advanced search features offered by search engine?

NEVER

RARERLY

OFTEN

VER OFTEN/ALWAYS

0　　　5　　　10　　　15

Which service do you use for online archival research?

General search engines

National archival information systems

Archival Web services (State Archives and others):
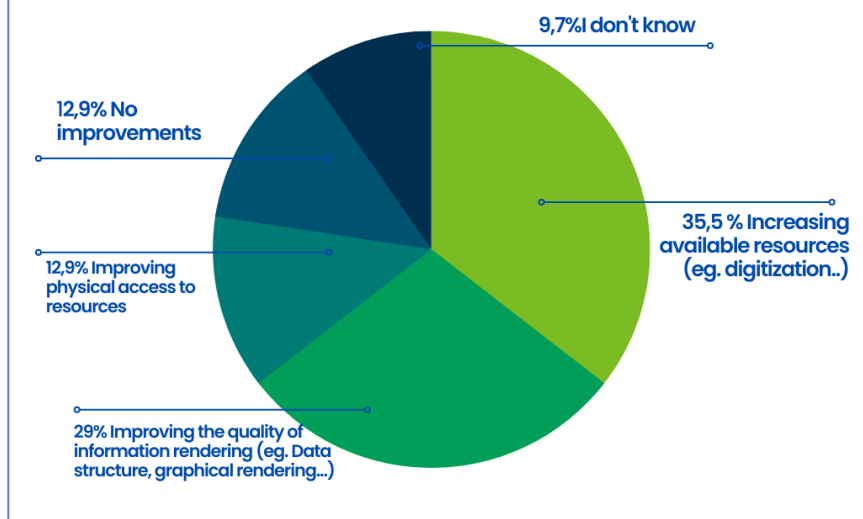
Wikipedia

None

0　　5　　10　　15　　20

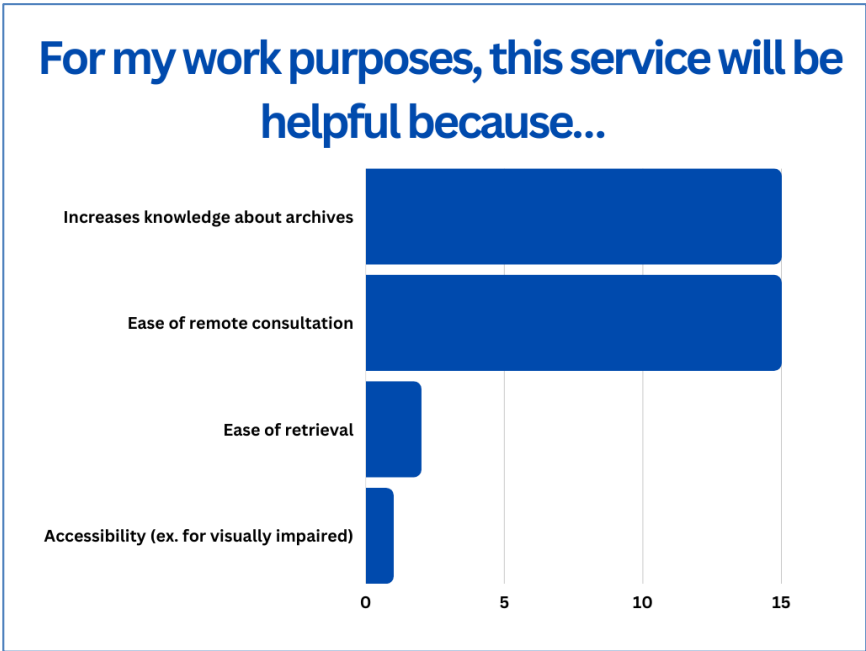**If you use advanced search features, which of these features are you confident in using?**
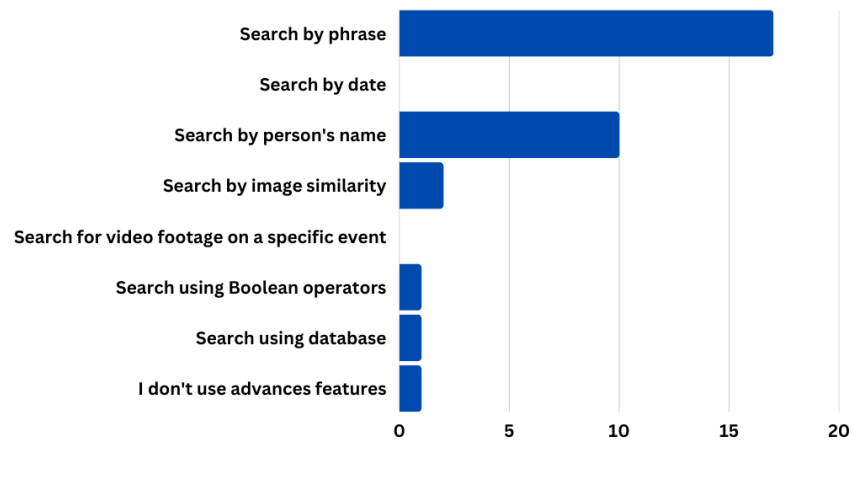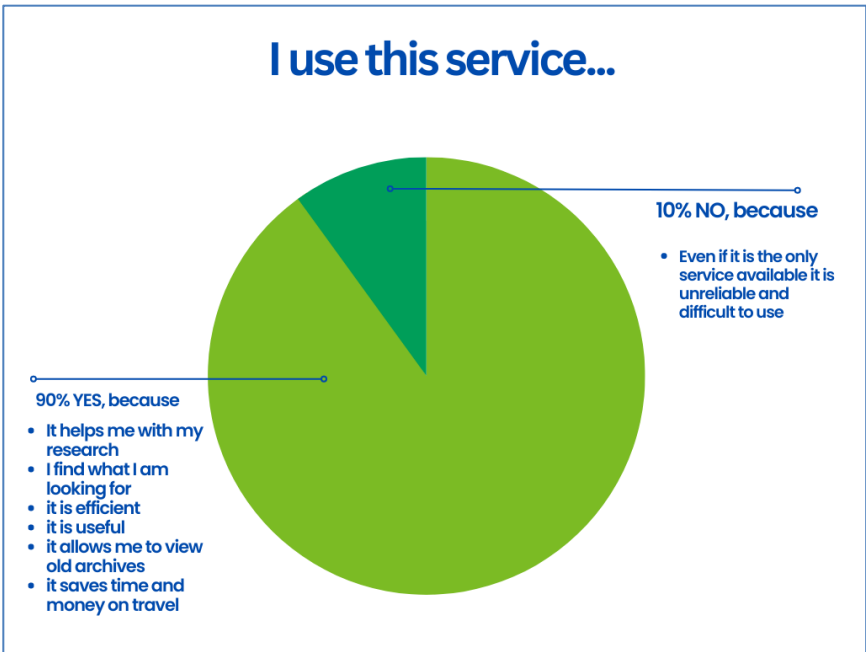
Regarding the specific archival service under investigation, participants expressed a desire for increased availability of digitized resources and an overall improvement in the quality of information retrieval. Additionally, they hope for greater freedom of access to physical resources held in the archive. Unfortunately, it is often not feasible for users to reserve online what they wish to view physically in the archives, and consultation is limited to only a few folders. Despite these limitations, more than 90 per cent of respondents indicated that they would continue to use the archival service surveyed, indicating a strong level of satisfaction.


I think this service could be improved by

9,7% I don't know

12,9% No improvements

35,5 % Increasing available resources (eg. digitization..)

12,9% Improving physical access to resources

29% Improving the quality of information rendering (eg. Data structure, graphical rendering...)

I use this service...

10% NO, because
- Even if it is the only service available it is unreliable and difficult to use

90% YES, because
- It helps me with my research
- I find what I am looking for
- it is efficient
- it is useful
- it allows me to view old archives
- it saves time and money on travel

For my work purposes, this service will be helpful because...

| Category	Value
Increases knowledge about archives	15
Ease of remote consultation	15
Ease of retrieval	2
Accessibility (ex. for visually impaired)	1

The final section of the questionnaire was dedicated to investigating the relationship between users and AI tools in archive research systems. It is noteworthy that most users surveyed were already familiar with the concept of artificial intelligence. Most respondents expressed a willingness to incorporate AI into search mechanisms. However, some of them acknowledged that they lack a deep understanding of AI processes and would therefore require more information regarding the advantages and disadvantages of utilizing AI tools in archive research systems.
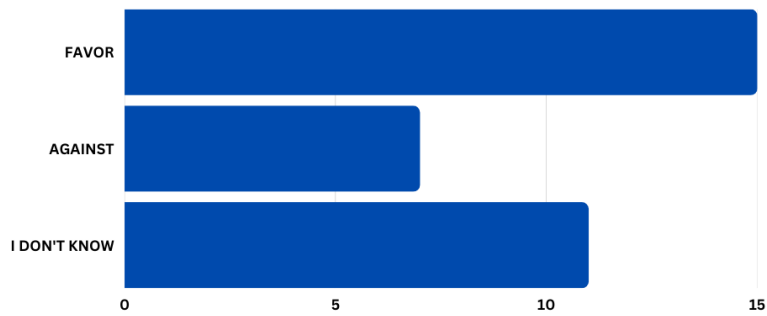
## Have you ever heard of Artificial Intelligence (AI)?

9,1% NO

90,9% YES
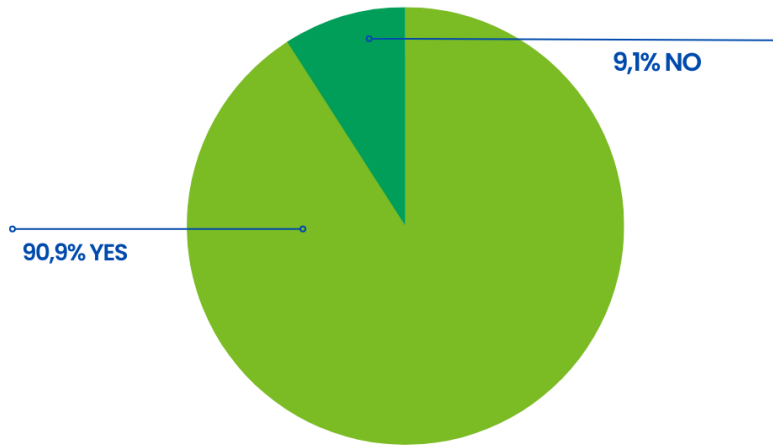
Are you aware that some web algorithms and AI tools already use your search data to improve your online experience?
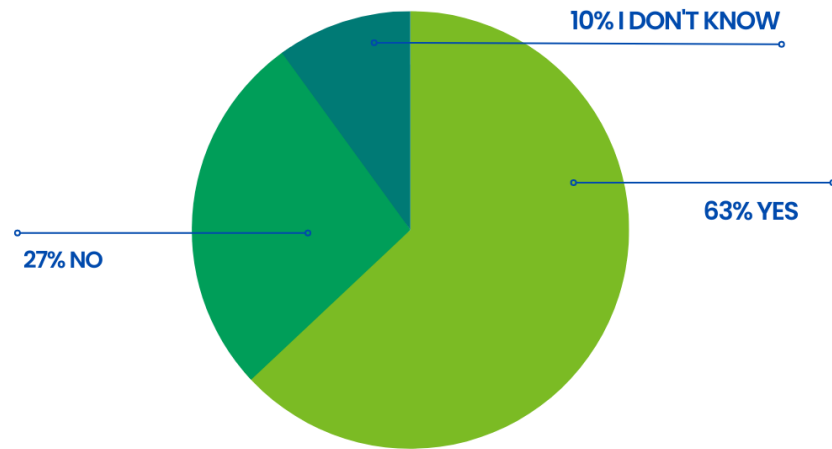
21,2% NO

78,8% YES
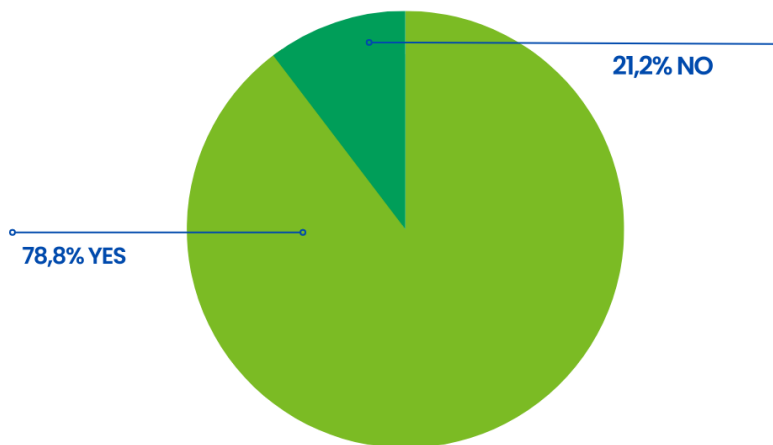
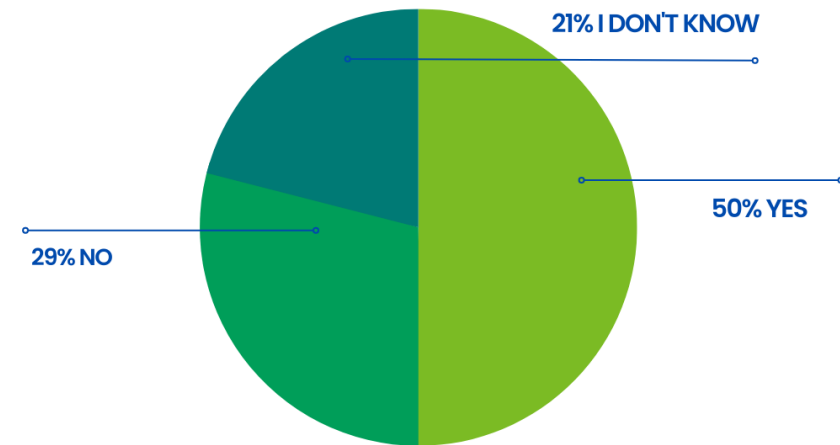**Are you for or against providing your research data to improve AI systems?**

Have you ever used AI tools to conduct research on records/archive materials?

9% I DON'T KNOW

24% YES

67% NO
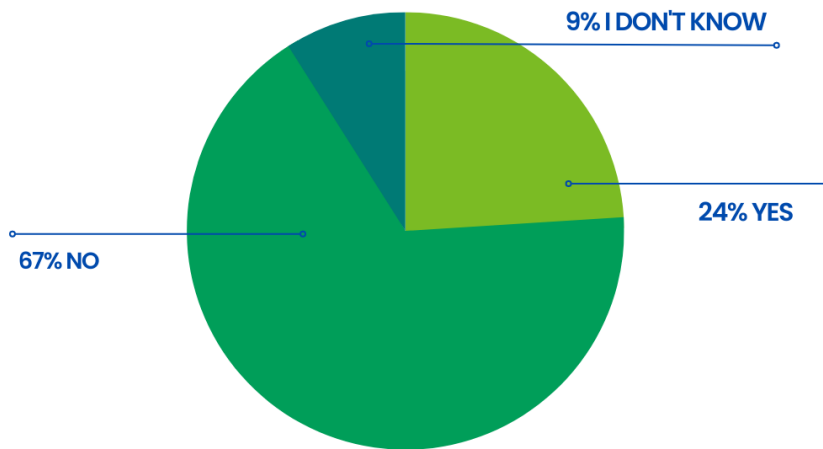
If you have never used them, do you think AI services could improve your research?

10% I DON'T KNOW

63% YES

27% NO

Do you have any doubts about how trustworthy are AI tools?

21% I DON'T KNOW

50% YES

29% NO

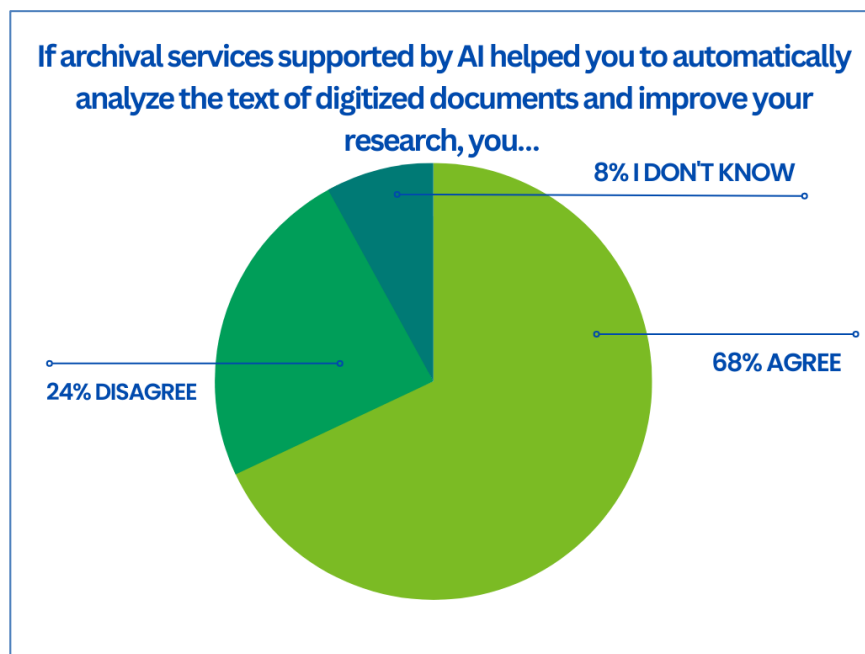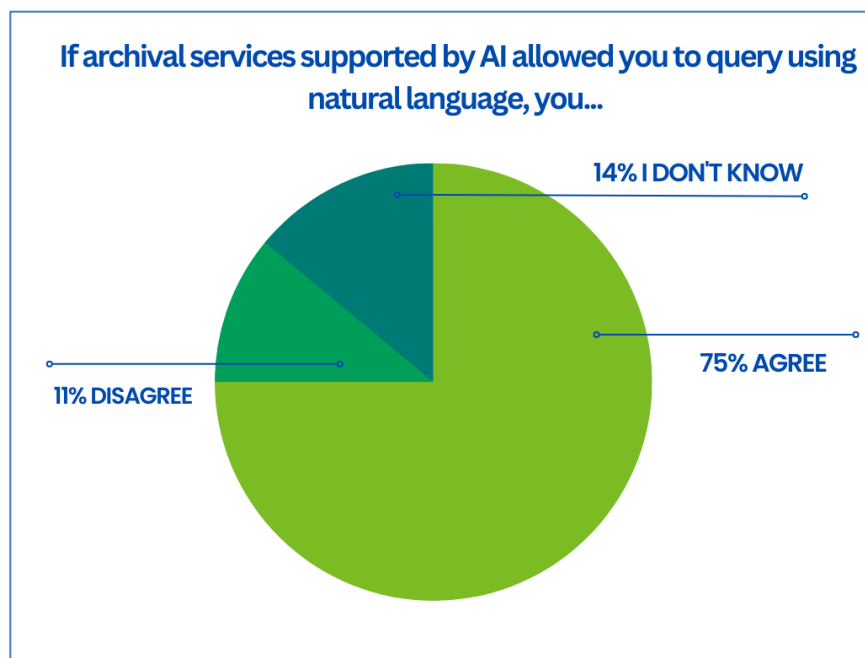## Would you be in favor or against records/archival service supported by AI for the preservation and processing of your research data?

11% I DON'T KNOW

17% NO

72% YES
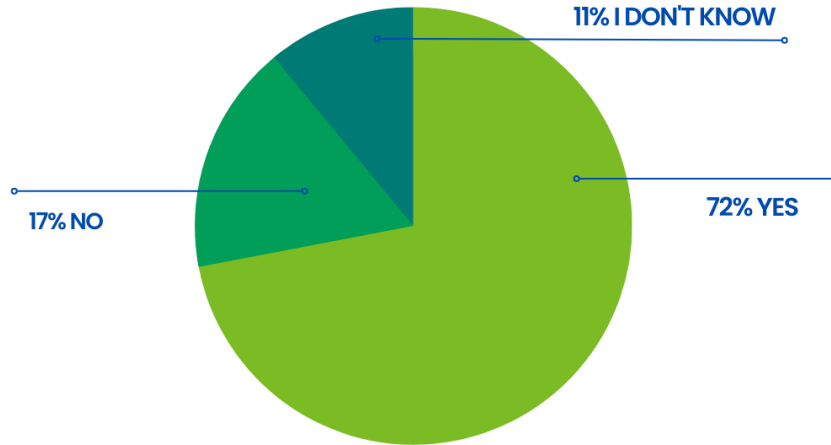
If archival services supported by AI provided search suggestions you...

21% I DON'T KNOW

9% DISAGREE

70% AGREE

If archival services supported by AI allowed you to query using natural language, you...

14% I DON'T KNOW

75% AGREE

11% DISAGREE

If archival services supported by AI helped you to automatically analyze the text of digitized documents and improve your research, you...

8% I DON'T KNOW

68% AGREE

24% DISAGREE
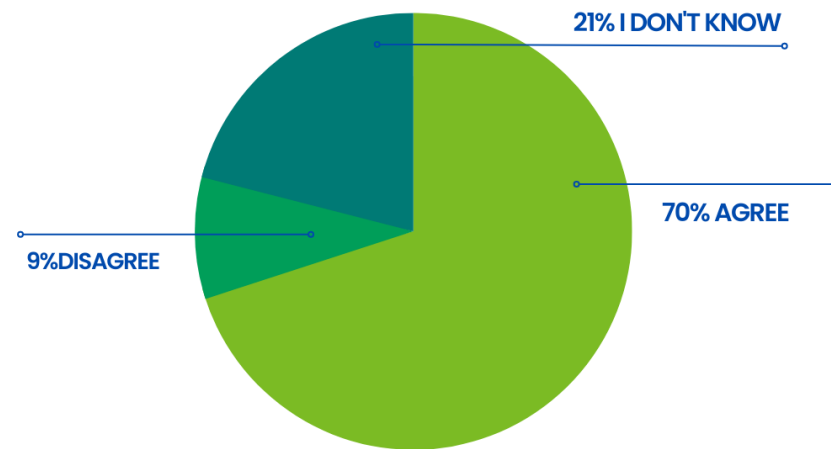
Conclusions

The results of this study provide valuable insights into the experiences and expectations of users in accessing digital archival services and their attitudes toward the adoption of AI-based solutions. The survey revealed that the sample of Italian respondents tend to rely on general search engines for online searches. They hope for increased availability of digitized resources and an overall improvement in the quality of information retrieval. Moreover, users expressed a willingness to incorporate AI into search mechanisms but they also indicated a need for more

information regarding the advantages and disadvantages of utilizing AI tools in archive research systems. Around 65% of respondents showed a positive attitude against adopting AI tools to improve archival reference and access services. This result could be read in light of the 2022 IPSOS survey: "78% of Chinese respondents (the highest proportion of surveyed countries) agreed with the statement that products and services using AI have more benefits than drawbacks. After Chinese respondents, those from Saudi Arabia (76%) and India (71%) felt the most positively about AI products. Only 35% of sampled Americans (among the lowest of surveyed countries) agreed that products and services using AI had more benefits than drawbacks". Italy, in this study, had just 50% of optimistic respondents. (Stanford HAI, 2023 AI Index Report, Chapter 8, Public Opinion, https://aiindex.stanford.edu/report/)

In the framework of ITrust AI, in our opinion, the study highlights the advice for further research to better understand users' behaviours and satisfaction with digital archival reference and access services, as well as the potential benefits and drawbacks of adopting AI tools to improve access to archival documents. The results, significantly when strengthened with data from other countries, could contribute to the development of more thoughtful consideration of functional and reliable artificial intelligence tools in archival services.

In conclusion, this study provides a starting point for further research and collaboration on the use of AI in archival research systems, and it is hoped that it will help to inform the development of more effective and user-friendly digital services for archives in the future.