



Archives search as a complex sensemaking task

Emi Ishita, Tokinori Suzuki, Douglas W. Oard, and Yoichi Tomiura

DOI: <https://doi.org/10.47989/ir31iConf64269>

Abstract

Introduction. This study investigated scholar's behavior in archival institutions to develop more effective search tools for searching paper-based materials in archival institutions.

Method. The scholar worked in the U.S. National Archives and was asked to keep a detailed diary of their actions and decisions. The principal analysis was conducted based on this diary.

Analysis. In an open coding process, the parts of the diary relevant to our research were identified and assigned an initial code. These codes were refined by repeating the process twice, *merging and splitting codes as necessary*.

Results. Six codes - 'Set a Goal', 'Choose a Strategy', 'Prioritising', 'Select Boxes', 'Examine Documents' and 'update mental model' - were identified. These represent an iterative, hierarchically structured sensemaking process in which changes to plans were based not only on what was found, but also on what was learned about what existed to be found.

Conclusion. This study shows an example of how complex actions and decisions occur when searching physical materials. It could help in understanding what kinds of tools could be developed to support such activities.

Introduction

Today, vast amounts of information are available in digital form, and most modern studies of search behavior thus target digital content. However, much valuable material remains undigitised. One notable example is the holdings of archival institutions. For example, the U.S. National Archives (NARA) preserves approximately 13 billion pages, of which only less than 3% has been digitised (National Archives, 2026). When searching for documents at archival institutions, researchers are typically required to request boxes that may contain the desired materials and review the contents to confirm their needs. Each of those steps is time-consuming. For instance, at NARA, obtaining boxes can easily take an hour, so ordering the right boxes is important. We ultimately aim to develop tools for searching paper materials in archival institutions. To obtain design inspiration, we investigated a scholar's search process in an archival institution.

We expected the search process to be complex, iterative, and multifaceted. That led us to adopt sensemaking as a conceptual frame. Although we expected learning what had actually happened (e.g., events, decisions, or policies) to be the scholar's ultimate goal, we also expected there would be things they would need to learn along the way (e.g., about how to find documents that could shed light on answers to their questions). Together, these perspectives led us to be curious about the structure of the process by which they would seek what they were looking for, and that in turn shaped our study design, our coding, and our analysis of the resulting data.

Related work

Search behavior of scholars in archives

The search problem in archives differs in two important ways from what we are used to in other settings: physicality and scale. The consequences of physicality are easily appreciated—materials in an archive often exist nowhere else in the world. Unlike a library, archival materials don't generally circulate. A library that holds the last existing copy of some book must take exceptional measures to preserve that book, and generally will not let it circulate. But in an archive, every item could be the only copy. So, to make extensive use of physical materials in an archive, you must physically go to that archive (Schmidt, 2016).

The second problem is scale, which precludes large-scale digitisation, and which severely limits the type and nature of descriptive metadata that can be created. Consider, for example, the difference in scale between the world's largest library (the U.S. Library of Congress; LoC), which has catalogued 35 million books, and the 13 billion undigitised pages at NARA. Given that a typical NARA document averages just a few pages, this is a two order of magnitude difference in number of documents. It is simply impractical for large archives to catalog much of their collection at the level of individual documents. For this reason, archival description typically focuses on higher levels of aggregation such as folders, boxes, or entire collections. Indeed, the situation is even more complex because the documents collected by an archive were originally created for some other purpose, and the way those documents are organised in the archive often reflects the way they were originally organised at the time they were in active use. So, there is no counterpart to the consistent topical classification or shelf ordering systems used in libraries, and searchers must learn how each set of documents they want to work with is organised. But there's one more factor to consider—most of the documents that were originally created have been destroyed—only a tiny fraction ever makes it to an archive. For example, a U.S. government document has about a 5% chance of being included in NARA's permanent collection (National Archives, 2018).

Searchers thus face several challenges. What documents were created? Which of those documents still exist? Where are they stored? Which might have information useful to the searcher? How can those be found? And they must do it all using the limited amount of description that is available. To answer these questions, scholars can turn to four broad classes of resources. The first is materials prepared by archivists to help searchers, which result from archival practices of

arrangement and description (Schellenberg, 2003). A second is the guidance of reference archivists, who know their collections well (Duff & Fox, 2006). A third is the work of prior scholars, in which references to specific archival materials can sometimes be found (Suzuki et al., 2023). And fourth is what can be learned from examining some of the other materials found in the archive (Duff & Johnson, 2002).

Sensemaking

Our interest is in what has been called exploratory search (Athukorala et al., 2015). In contrast to more narrowly scoped tasks such as question answering or known-item search, exploratory search is an iterative, multi-faceted, and open-ended process that supports synthesis of information, where the users' end goal requires some element of learning about things that are new to them. Exploratory search often surfaces new questions as the searcher learns more about their topic. Supporting exploratory search thus requires not just helping the searcher to find things; it also requires helping them learn about what is there to be found. This involves sensemaking, which Russell et al. (1993) described as consisting of four subprocesses: searching for suitable information, meaningfully organising it, iteratively adjusting that organisation as more is learned, and using what has been found to develop an understanding that enables achieving goals. Dervin (1998) has described sensemaking using *'the metaphor of human beings traveling through timespace, coming out of situations with history and partial instruction, arriving at new situations, facing gaps, building bridges across those gaps, evaluating outcomes and moving on.'* We are, thus, interested in not just what searchers do, but in why they do it, and in how their choices evolve over time.

Methodology

Data collection

We recruited a professor from a Japanese university, who flew to visit the NARA 'Archives 2' facility in Maryland and LoC for two weeks in September 2024 to learn about the treatment of Japanese Americans who the US government had classified as disloyal during the Second World War. They had been to NARA many times, studying other questions, for the past 30 years. The ways archives are structured was well known to them, as they teach archival science, and they have experience with the management of manuscript collections. We thus were studying expert search, observing someone with expertise in archives, history, and their research topic.

At the start of our study we met with the scholar and explained the purpose of our study, our procedure, and how we would handle resulting data, and we obtained their consent to participate. We paid half their travel costs from the research grant that supported our study. We submitted our study design to our Institutional Review Board, which determined it to be exempt from human subjects research review because we were studying only a single person, and that does not meet the board's definition of human subjects research.

We asked the scholar to keep a detailed diary of their actions and decisions, and to photograph some documents that would illustrate how they made decisions. We also silently observed and took notes. Each evening, they gave us their diary and all their photos (most were taken for their own purposes; in those they took specifically for us they sometimes indicated items of interest in a document by placing a physical marker). We reviewed the diary and photos each evening, and on most mornings, we conducted an audio recorded interview with them to enrich our understanding of what had happened the prior day.

The scholar worked at NARA (seven days) and LoC (two days) about 8 hours per day from September 12-23, 2024 (excluding three weekend days). At LoC, they worked in the manuscript library. Our principal focus is on the scholar's seven days at NARA because we have much more data from their time at NARA, but because the shift from NARA to LoC allowed us to observe the start of a search process at a new institution we have included their activity at LoC in our coding.

Coding and analysis

We focused principally on coding the diary. We used open coding to simultaneously evolve the design of our coding frame, code specific types of activities, and create a task model depicting sequential relations between activities. The first author, who also observed scholar's behavior, first identified parts of the diary related to our research and then assigned codes to actions and decisions of the scholar. These codes were then refined by repeating the process twice, merging and splitting codes as necessary. The second and third authors then conducted peer review of the coding, after which a few coding decisions were adjusted. Coding was done directly on the Japanese diary, although for the convenience of the research team the coding frame was developed using English labels. and an English Google translation of the diary was used by the third author (who cannot read Japanese) for peer review.

Drawing inspiration from Dervin (1998), we have used verb phrases to indicate acts, both physical and mental. Drawing inspiration from Russell et al. (1993), we have focused on transitions between different types of acts. The result is a sequence model, similar to those that result from contextual inquiry (Mirel, 1996), although our narrower focus on work tasks elides much of the richness that a broader approach to contextual inquiry might seek to include.

Result

Sample of searching behavior with codes

The diary is 23 A4 pages, including some personal notes unrelated to our study. Here is an example of our coding from the start of Day 5 (Example diary entries were translated from to English for presentation here, first using Google Translate, and then manually correcting the translation).

On that day, the scholar first decided which boxes to request by looking at the printed catalog and the finding aids at NARA (*Choose a Strategy*), as noted in their diary by 'I plan my strategy by looking at the catalog to choose which boxes should be looked at today. The first priority is to look for materials related to how the United States treated Japanese Americans who did not pledge allegiance to the United States. It is unclear whether they are available.' After that, they examined the contents of some boxes that had been requested earlier (*Examined Documents*), ultimately concluding that some of what they had found was 'Personal information? I thought it was an 'investigative report,' but it was most likely a personal record. It was pointless.' (*Updated Mental Model*). The scholar also took photos of some other documents by first looking at the contents of a folder and inserting paper slips to indicate records that they thought were important, and then later taking photos of those documents (*Examine Documents*). Because time was short, they wrote in their diary 'Tomorrow, I will take photos of the documents with slips in Box 4' (*Prioritisation*). Time pressure was also evident when they wrote, 'As soon as that is done, I will check the next box in hurry.' In the evening, they developed a strategy for how to search for the boxes the next day, writing to us in an email that evening that they wanted to 'search for records from the Department of the Army, focusing on those dealing with Japanese-Americans who have sworn allegiance to Japan and have not professed allegiance to the United States.' (*Choose a Strategy*).

Coding frame

Our final coding frame is shown here.

Set a Goal: Assigned when the scholar decided on, or changed their decision on, what they wanted to achieve. At both NARA and LoC, they had decided on a goal before first arriving, but in each case their goal changed.

Choose a Strategy: Assigned when they chose a strategy for finding something, including changing a strategy. For example, 'I decided to start by getting a general idea of the overall policy on the

treatment of Japanese Americans from the Department of Justice records.' (Day 1) Sometimes this involved things they planned to do after returning to Japan.

Select Boxes: Assigned to actions taken to determine which boxes to request from the closed stacks. This includes, for example, talking with archivists or other researchers, searching catalog metadata, and consulting finding aids (Wiedeman, 2019). There were also cases when the scholar sent an email to an archivist to arrange a meeting.

Examine Documents: Assigned to working in some way with documents. Documents in a box were normally organised into folders. At LoC some documents were available only on microfilm, organised by reel. Actions included skimming a folder, box or reel (usually their first step), reading documents, or photographing for later use.

Update Mental Model: Assigned to things that the scholar learned by examining documents. They learned in other ways too, such as conversations with archivists or by searching the NARA catalog, but to limit complexity we have coded learning from sources other than documents using our Choose a Strategy code, described above.

Prioritising: Although not a separate stage in our model, our scholar frequently chose how best to spend their limited time. This happened within every stage of our model, and also when deciding to continue one activity (e.g., Examine Documents) or shift to another (e.g., to Choose a Strategy). Examples include: *'I shot with a digital camera while keeping an eye on the time.'* (Day 12) and *'I did not have time to take photos of the list of Chinese forced labourers.'* (Day 12)

Task model

Figure 1 shows the task model that emerged from our analysis, a set of four nested loops; we observed each of the eight depicted transitions more than once.

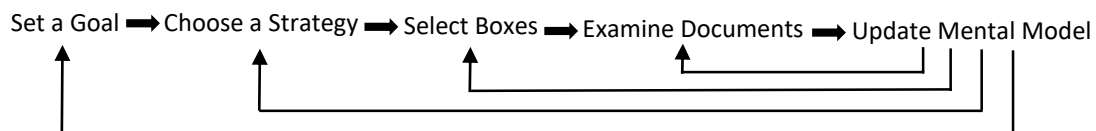


Figure 1. Task model for a scholar's searching behavior in archives.

Discussion

There was no single pattern for what would happen on a given day. We do, however, see a natural progression across days, with refinement of the goal occurring early on, examination of documents dominating the middle days, and an increasing need for prioritisation as the end of their time there approached. We frequently saw our scholar thinking strategically, and we saw many changes in their strategy, despite there being few changes in their goals.

We also saw two systematic effects from time pressure. First, the scholar photographed many documents they did not have time to read. While a useful expedient, this requires later remembering salient details of the context in which those documents had been found. It might therefore help to automatically capture some of that context and make use of it when later reviewing photos. A second time pressure effect was that our scholar chose not to do some things that could be done later. For example, while at LoC they didn't look at one set of materials because copies were available in Japan. We note that our scholar had considerable expertise, and less experienced scholars might not be as familiar with which things are available from other sources. It could thus be useful if tools in an archive could alert them to such cases.

Perhaps the most important thing that we learned is how complex this task is, as our task model illustrates. Although the scholar's goal was simple—to find documents that helped shed light on a specific topic—the search process was not at all simple. To know where to look, they had to understand the structure of the government that created the documents. To find those documents, they needed to learn how each part of the collection was organised, how it was described, and how it could be searched. They then faced the challenge of translating that to a request for specific boxes from specific locations. Only then were they faced with an overwhelming number of documents to deal with in limited amount of time. And, as noted above, every stage in the process was iterative, meaning that getting to their goal required doing those steps repeatedly.

Conclusion

What we have presented is a case study with a single scholar on a single trip in which they visited just two repositories. Of course, there is no way we could hope to capture the full diversity of scholarly work in archives by observing one scholar, many users of archives are not scholars, and quite a lot of what archives hold are not documents (there are also, for example, maps, photographs, and audio recordings). So, it would be a mistake to try to generalise too much from a single case study. Nonetheless, scholars are a group worthy of study, paper documents are an important target for their searches, NARA and LoC are large repositories, and this case study has generated useful insights into the structure of the search task. This study has thus helped us to enrich our thinking on what kinds of tools we might build that could be helpful for at least some parts of this task.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP23KK0005.

About the authors

Emi Ishita is a Professor in Department of Library Science, Kyushu University, Fukuoka, Japan. She received her PhD from Kyushu University, and her research interests include the searching behavior in memory institutions, information retrieval, and research data management. She can be contacted at ishita.emi.982@m.kyushu-u.ac.jp

Tokinori Suzuki is an Assistant Professor in the Institute of Systems and Information Engineering at the University of Tsukuba, Tsukuba, Japan. He received his PhD in Engineering from Kyushu University. His research interests include information access systems and digital libraries. He can be contacted at tokinori@cs.tsukuba.ac.jp

Douglas W. Oard is a Professor in the College of Information at the University of Maryland in College Park, Maryland, USA. He received his PhD in Electrical Engineering from the University of Maryland, and his research interests include the design and evaluation of information retrieval systems. He can be contacted at oard@umd.edu

Yoichi Tomiura is a Professor in Department of Informatics, Kyushu University, Fukuoka, Japan. He received her PhD from Kyushu University, and his research interests include Natural Language Processing, information retrieval, and research data management. He can be contacted at tomiura.yoichi.802@m.kyushu-u.ac.jp

References

- Athukorala, K., Glowacka, D., Jacucci, G., Oulasvirta, A., & Vreeken, J. (2015). Is Exploratory Search Different? A comparison of information search behavior for exploratory and lookup tasks. *Journal of the Association for Information Science and Technology*, 67(11), 2635–2651.
- Dervin, B. (1998). Sense-making theory and practice: An overview of user interests in knowledge seeking and use. *Journal of Knowledge Management*, 2(2), 36–46. <https://doi.org/10.1108/13673279810249369>
- Duff, W. & Fox, A. (2006). You're a guide rather than an expert: Archival reference from an archivist's point of view. *Journal of the Society of Archivists*, 27(2), 129–153.
- Duff, W. M. & Johnson, C.A. (2002). Accidentally Found on Purpose: Information-seeking behavior of historians in archives. *The Library Quarterly*, 72(4), 472–496.
- Mirel, B. (1996). Contextual inquiry and the representation of tasks. *ACM SIGDOC Asterisk Journal of Computer Documentation*, 20(1), 14–21. <https://doi.org/10.1145/227614.227616>
- National Archives. (2026). National Archives. <https://www.archives.gov/>
- National Archives. (2018). About the National Archives of the United States. National Archives. <https://www.archives.gov/publications/general-info-leaflets/1-about-archives.html>
- Russell, D. M., Stefik, M. J., Pirolli, P., & Card, S. K. (1993). The Cost Structure of Sensemaking. Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems, 269–276. <https://doi.org/10.1145/169059.169209>
- Schellenberg, T.R. (2003). *The Management of Archives*, Society of American Archivists.
- Schmidt, L. (2016), *Using Archives: A guide to effective research*. Society of American Archivists. <https://www2.archivists.org/usingarchives>
- Suzuki, T., Oard, D., Ishita, E., & Tomiura, Y. (2023). Automatically Detecting References from the Scholarly Literature to Records in Archives. *ICADL 2023. Lecture Notes in Computer Science*, 14458. Springer, Singapore. https://doi.org/10.1007/978-981-99-8088-8_9
- Wiedeman, G. (2019). The Historical Hazards of Finding Aids. *The American Archivist*, 82(2), 381–420. <https://doi.org/10.17723/aarc-82-02-20>

© [CC-BY-NC 4.0](https://creativecommons.org/licenses/by-nc/4.0/) The Author(s). For more information, see our [Open Access Policy](#).