Effects of Past Interactions on User Experience with Recommended Documents

Farnaz Jahanbakhsh
Massachusetts Institute of Technology
farnazj@mit.edu

Ahmed Hassan Awadallah
Susan T. Dumais
[hassanam,sdumais]@microsoft.com
Microsoft Research

Xuhai Xu
University of Washington
xuhaixu@uw.edu

ABSTRACT

Recommender systems are commonly used in entertainment, news, e-commerce, and social media. Document recommendation is a new and under-explored application area, in which both re-finding and discovery of documents need to be supported. In this paper we provide an initial exploration of users' experience with recommended documents, with a focus on how prior interactions influence recognition and interest. Through a field study of more than 100 users, we investigate the effects of past interactions with recommended documents on users' recognition of, prior intent to open, and interest in the documents. We examined different presentations of interaction history, and the recency and richness of prior interaction. We found that presentation only influenced recognition time. Our findings also indicate that people are more likely to recognize documents they had accessed recently and to do so more quickly. Similarly, documents that people had interacted with more deeply were also more frequently and quickly recognized. However, people were more interested in older documents or those with which they had less involved interactions. This finding suggests that in addition to helping users quickly access documents they intend to re-find, document recommendation can add value in helping users discover other documents. Our results offer implications for designing document recommendation systems that help users fulfil different needs.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**.

KEYWORDS

Document Recommendation, User Behavior, Presentation of Recommended Items, User-Document Interactions, Document Re-finding, Document Discovery

ACM Reference Format:
Farnaz Jahanbakhsh, Ahmed Hassan Awadallah, Susan T. Dumais, and Xuhai Xu. 2020. Effects of Past Interactions on User Experience with Recommended Documents. In *2020 Conference on Human Information Interaction and Retrieval (CHIIR '20), March 14–18, 2020, Vancouver, BC, Canada.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3343413.3377977

CHIIR '20, March 14–18, 2020, Vancouver, BC, Canada

ACM ISBN 978-1-4503-6892-6/20/03…$15.00
https://doi.org/10.1145/3343413.3377977

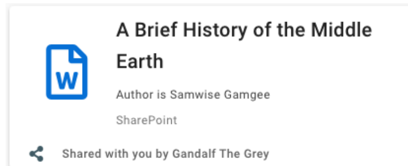

Figure 1: Example of a document in the Basic presentation condition where only basic properties are displayed. These properties include an icon denoting the file type, title, author, location, and who shared the document with the user.

1 INTRODUCTION

Recommender systems have been well integrated into many aspects of our lives [6, 7, 25]. In many domains such as e-commerce, entertainment, news feeds, hiring platforms, and social networks, these systems are primarily used to help users discover new items that might be of interest to them [11, 16, 20, 23, 33]. Document recommendation however, is a unique domain in that its chief concern is to facilitate re-finding of user's items [26].

In comparison with other recommendation domains, document recommendation has been less examined with only few studies in this area focusing on improving the accuracy of the recommender algorithm behind the scenes [15, 26]. While the algorithm is an important aspect of the system, knowing about the effects of other aspects, for instance, presentation, explanations, and users' interaction with the recommended items on user experience helps with designing more effective recommender systems.

In this work, we focus on user's experience with recommended documents. Because individuals and their collaborators are likely to have had previous interactions with the document, the extent to which they recognize the recommended documents or how useful they find the documents may depend on these past interactions. Inspired by previous work in personal information management and refinding [2, 13], our study seeks to investigate the effects of three important dimensions of users' past interactions with documents recommended to them on their recognition of, prior intent to open, and interest in the documents. The three dimensions are: recency of access, richness of prior interaction, and the presentation of interactions in document summaries.

We developed an experimental platform in which participants were presented with a set of recommended documents, one at a time. The recommended documents were selected from a person's own cloud-based document repository, and included both documents they had created and documents that others had shared with them. Studying user's interaction with the recommended documents enabled us to investigate the following research questions:

- **RQ1:** How does users' recency of access to documents recommended to them affect their recognition, interest in the documents, and their intent to open the documents in the near future?
- **RQ2:** How does users' richness of prior interaction with documents recommended to them affect their recognition, interest in the documents, and their intent to open the documents in the near future?
- **RQ3:** How does displaying interactions of the user and their collaborators on the summary of recommended documents affect their recognition and interest in the documents?

Our results indicate that the more recent the user's last access to a document is, the more likely they are to recognize it and more quickly. In addition, they are more likely to have the intent to open recent documents, compared to older ones. We observed a similar behavior for the documents with which the users have had more involved interactions. However, users are more interested in older documents or those with which they have had shallower or no interactions. These findings suggest that focusing document recommender systems only on documents that the user intends to open would result in missing documents that the user would otherwise find interesting and deter us from exploring a "discovery" space in this domain.

We did not observe a significant effect of how previous interaction history was summarized on outcomes. However, free-text responses suggested that viewing the history of interactions can be useful in recognition and in understanding the status of ongoing editing. Future work is needed to understand under what circumstances displaying interaction history can improve user experience.

The main contribution of our work is an empirical understanding of how three aspects of users' past interactions - recency of last access, richness of interaction, and presentation - impact the effectiveness of the documents recommended to them. We study these aspects in an environment where participants interact with their own documents while having interactions that are not limited to click through data obtained from usage logs only.

2 RELATED WORK

We situate our work in prior work on personal information management, information refinding, and recommendation systems.

2.1 Personal Information Management

Personal information management (PIM) seeks to understand and support the activities people perform to acquire, organize, maintain, retrieve, and use personal information [18]. Personal information can come in many forms including documents, web pages, email messages, notes, etc. People use personal information to complete tasks and fulfil different roles. Jones [18] outlines six ways in which information can be personal: Owned by, about, directed toward, sent by, experienced by, or relevant to "me".

Several studies have looked into characterizing and supporting personal information management strategies. One line of research focuses on studying the role of search in personal information management. For example, Cutrell et al. [12] argue that search systems can alleviate the need to organize personal information by helping us find it no matter where we encountered it, what we remember about it, and even if we forget it exists. Bergman et al. [4] study whether improvements in search have changed this fundamental aspect of PIM. They also offer theoretical explanations for differences between PIM and Internet retrieval, and suggest alternative design directions for PIM systems. A detailed survey of research on information seeking, information needs, and user behavior is presented in [22].

Another thread of work focuses on personal information management in more specific domains. One domain that has received significant attention is email. For example, Venolia et al. [31] focus on understanding activities and workflows surrounding how people use email. Siu et al. [24] study email use in the context of everyday work practices. They examine how users interlace email with their day-to-day, ongoing work processes. Other studies focus on documents. Folder navigation to retrieve documents is studied in detail in [5]. They argue that people dedicate considerable time to creating systematic structures to facilitate such retrieval. They also use a predictive model to formulate the effect of folder depth and folder size on retrieval time. An empirical study to compare two methods of organizing documents – placing them into folders or tagging them with labels – is described in [5]. Study results point to the importance of designing tools that combine strengths of folders and labels while avoiding their weaknesses.

2.2 Information Recall and Refinding

Another body of work investigates how people recall and refind previously seen information. Research has shown that a significant portion of an individual's web accesses tends to be revisits [10, 27, 29]. Jones et al. [17] study how people retain web information they have found for future use. In a user study on search, Teevan [28] reports that what makes a search result memorable is the rank and whether it was clicked on.

In addition to studying refinding in the context of Web search, researchers have studied re-finding of personal information, especially with an emphasis on email. Grbovic et al. [14] show that, with the increase of email messages over time, users tend to rely on search for refinding emails as opposed to using human-generated folders and tags. Dumais et al. [13] describe the design of a system that facilitates information re-use. The system provides a unified index of information that a person has seen, regardless of whether it was seen as email, web page, document, etc. and uses rich contextual cues in the search interface. They found that that email was the most commonly retrieved source of personal information (e.g. files, web history, emails, etc.). More recently, researchers have also studied re-visitation patterns [1] and refinding strategies [21] employed by users to go back to previously seen email messages.

Other studies have been concerned with finding presentations that help with better recognition. Cockburn et al. [9] design thumbnails that show a person's interaction with webpages for instance, by marking pages they have visited close together in time and those they have frequently accessed. In another work, Teevan et al. [30] examine how different representations of web pages affect refinding and finding new information that the user has never seen before. They find for example, that thumbnails help with refinding, but only when the page's thumbnail has been seen before.

2.3 Recommendation

Personalized recommendations are increasingly employed in a variety of areas, most commonly in entertainment to for instance recommend music or videos [11, 19, 32, 35], product recommendation in online shopping platforms [25, 34], and social media platforms [16, 23]. Document recommendation, on the other hand, has not received as much attention. Guan et al. [15] studied document recommendation in the context of social tagging. They argue that annotating documents with freely chosen keywords (tags) can provide meaningful collaborative semantic data which can potentially be exploited by recommender systems. In more recent work, a document recommendation system to provide quick access to documents on the Google Drive platform was described in [26]. The system aims to surface the most relevant documents when a user visits the home screen. The paper reports significant productivity gains, in terms of time to locate documents, compared to other approaches that rely on search or browsing.

Although document recommendation is not a heavily studied area, it can benefit from several insights in studying recommendations in other domains. For example, several papers have attempted to model repeated consumption behavior and its impact on recommender systems. Several important aspects such as item popularity, recency of access [2], user reconsumption patterns [8] and inter-consumption frequency [3] were highlighted.

Document recommendation is related to personal information management, refinding and reconsumption of information and recommender systems in many ways. Most studies on personal information management have focused on search, foldering and tagging as a means for facilitating information access. On the other hand, refinding and recognition of previously seen information have been studied in the context of Web information. Less is known about the domain of personal documents and how users' previous interactions with their documents affect their recognition. Although a vast amount of work has looked at personalized recommendation, most of the attention went to domains like shopping, entertainment, etc. with recommendations for productivity in general, and document recommendation in particular, receiving little attention.

Our work attempts to bridge this gap, while also leveraging and drawing insights from previous work, by studying how richness of interaction with a document, recency of access and presentation styles affect user's recognition of, their prior intent to open, and their interest in the document. In this work, we incorporate contextual cues - user's and their collaborators' past interactions with a document and their timestamps - into a document recommender system to study if presenting such cues helps with productivity. Furthermore, we study how presenting such context along with the documents affects people's interest in them.

3 METHOD

To investigate our research questions, we developed a platform where study participants could interact with documents in their personal cloud repository. This includes both documents that they had created and those to which they had access (e.g., because they were shared by a collaborator). The documents that we retrieved resided in a document repository on cloud and were the output of an earlier stage of an existing document recommender system that

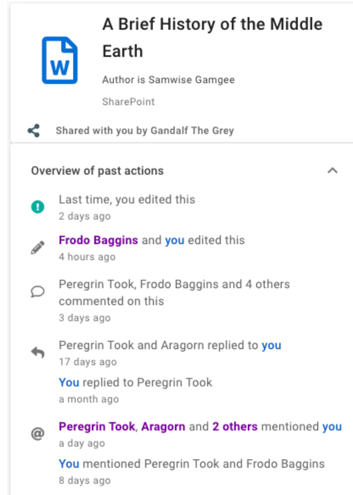

Figure 2: Example of a document in the Advanced presentation condition where interactions are displayed in addition to the document's basic properties. User's actions are differentiated with a blue color. Purple shows actions that have taken place since the user last interacted with the document.

uses features of the user history, context, and document properties to identify candidate documents for recommendation to the user. Therefore, the subset we selected for the experiment from this document was a realistic sample of recommended documents for the user. Our study was approved by our Institutional Review Board.

3.1 Aspects of Interaction

In this paper we investigate how people's previous history of interactions with documents to which they have access affects their recognition of, intent to open, and interest in the documents. Based on previous work in PIM and refinding, we consider three aspects of interaction - recency of prior interaction, richness of prior interaction, and presentation of interaction history.

As described in more detail below, we experimentally varied how the history of prior interactions was presented. Since we were working with people's own documents, we could not manipulate how they had previously interacted with them, so we systematically selected documents that spanned different time horizons of interaction and richness of interaction.

3.1.1 Recency of Interaction. To study the effect of last access time, we compare outcome measures for documents that the user has accessed recently or in the more distant past.

Specifically we consider:

- the last 24 hours (*Today*)
- 1-7 days ago (*Last Week*)
- more than 7 days ago (*Older*)
- never accessed (*Never*)

3.1.2 Richness of Interaction. To examine richness of interaction, we consider how people and their collaborators previously interacted with the documents. We compare active engagement (e.g., editing/commenting), only reading, and no prior interaction.

Specifically we consider:

- contributed to the document at some point in the past by editing, commenting, replying to a comment, or mentioning someone in a comment (*Contribute*)
- only opened the document (*Open Only*)
- never opened the document (*No Interaction*)

3.1.3 Presentation. To investigate how presenting past interactions with the document affect users' recognition, prior intent to open, and interest in the document, we created two presentation conditions: *Basic* and *Advanced*. In the *Basic* presentation (see Figure 1), for each document we showed only the basic properties in the document tile: an icon denoting the document type, title, author, location, and who, if anyone, has shared the document with the user. In the *Advanced* presentation (see Figure 2), we displayed a summary of the past interactions of the user and their collaborators in addition to the basic properties. These past interactions were broken into 5 segments: the last interaction the participant has had with the document, names of the people including the participant who have edited the document, names of those who have commented on the document, names of people who have replied to a comment by the participant or those to whom the participant has replied, and names of people who have mentioned the participant in the document or those the participant has mentioned. Next to each set of actions, we also displayed a timestamp for the most recent action in that set. The participant's actions were highlighted by a blue color. Those actions that had taken place since the participant last interacted with the document were indicated with a purple color. Not all documents in the *Advanced* condition had every action taken place on them, in which case the segment for the missing action was simply not displayed. If the participant had never interacted with the document, the *last interaction* segment would show the message *"You have never opened the document"*.

3.2 Document Selection

By logging into our platform, study participants enabled the client to access, on their behalf, the metadata of up to 100 documents in their cloud repository as well as their history of interaction with each document and that of their collaborators. The 100 documents were the output of an earlier stage of an existing document recommender system that from all the documents the user has access to, selects a few to recommend to the user. This list is compiled by ranking all documents, to which the user has access, by the time they were touched by the user or other collaborators. Note that some of these 100 documents were created by the participant and others were owned by others but shared with the participant.

After a participant consented to the study, our platform would request the participant's 100 most recently touched documents and the history of past interactions for each. From this set, we then selected a subset to present to the users to test recognition, intent to open, and interest. Because we were interested in how recognition varies by the richness of prior interaction with the document and the recency of access, we wanted to sample documents that spanned these dimensions. To do so, we separated participant's documents into 4 recency of access partitions (see section 3.1.1), and 3 richness of interaction partitions (see section 3.1.2). From each partition, we randomly sampled 4 documents and added them to the set of the documents that were displayed to the user during the experiment. If a partition for a participant contained fewer than 4 documents, we selected all the documents within that partition. Participants who had 4 or more documents in each partition would be presented with 28 recommended documents to judge. Table 1 shows the distribution of participants' documents across all recency of access and richness of prior interaction conditions in our study.

Table 1: Distributions of participants' documents across conditions. The columns show the recency of last access to the document. The rows represent the richness of interaction. The greyed-out cells are not applicable.

	Never	Older	Last Week	Today
No Interaction	335			
Open Only		436	164	77
Contribute		373	160	92

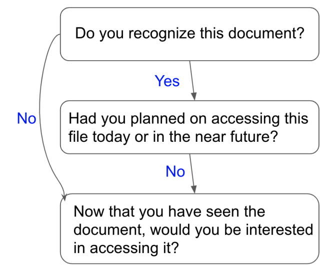

Figure 3: The flow of questions for each document in the study. The blue text represents participant's response.

3.3 Study Design and Procedure

We expected the distribution of users' documents along the two axes of recency and richness of interaction to be different across users. Therefore, to have a finer control and a more balanced dataset across the 3 factors of our study, we used a within-subjects design where for each user, we randomly split the documents from each of the 7 interaction and recency conditions in half and assigned half from each to the *Basic* presentation and the other half to the *Advanced* presentation. Documents were presented one at a time. Participants interacted with a block of documents in one presentation followed by documents in the other presentation. The order of presentation blocks was determined randomly for each participant.

After consenting to the study, the participants read the task instructions and were shown an example document along with explanations for the different parts of the interface. Participants were informed that they could open each document by clicking on its title if they needed to, although they were not encouraged to do so. The instructions specified that they would be shown a series of documents one at a time and asked a few questions for each through a form. Figure 3 depicts the flow of questions. For each document, the form asked whether the user recognized the document, to which they clicked either "Yes" or "No". Participants were informed that we were also interested in how long it took them to recognize the document. To encourage more normal interactions, the instructions specified that the purpose was not to test the participant but rather the system and that they should not feel pressured into selecting an answer before they had made a decision. If the participant stated that they did recognize the documents, the form would follow up with another question asking whether they had planned on accessing the file that day or the near future. If the participants specified that they did not recognize the file or that they had not intended to access the document, the form would ask whether they would be interested in accessing the document in the future now that they had seen it. In addition to the yes/no responses, the form asked for elaboration via free-text fields.

Throughout the experiment, we logged participants' responses, de-identified file metadata, and their interaction with the website including how long they took to select an answer for the recognition questions and whether they opened any file by clicking on its title.

Following the experiment, participants completed a demographics survey. At the study's conclusion, the platform asked whether the participant wished to revoke their consent for participating in the study. Because the study data was de-identified from the beginning, we could not link study data to an individual once the session for that participant was over.

3.4 Participants

We recruited participants by sending an invitation to a random set of employees within a large technology company. A total of 108 participants completed the study of which we discarded the responses from 2 participants because they had very few documents. 29% of the remaining participants were female. They were distributed across a wide age range with a median of 35-44 years. The median of the highest education achieved was a Bachelor's degree. Participants came from a diverse set of roles within the company including: software development, program management, sales, marketing, administrative assistance, IT support, finance, retail, content writing, etc.

4 RESULTS AND DISCUSSION

We first report on the main effects of 3 aspects of interaction - users' recency of last access to the document, richness of their past interactions with it, and the presentation of such interactions - on outcome measures. We then proceed to discuss the interaction effects of the factors that have significant main effects on the outcomes.

We consider four outcome measures. When we examine recognition, we discuss our results with relation to two outcome measures: *recognition ratio* – among all documents presented to the participants, what proportion they recognized and *recognition time* – how long it took participants to decide whether they recognize the documents presented to them. The two other outcome measures that we examine are: the users' prior *intent to open* documents that they recognized, and their *interest* in documents that they did not recognize or did not have prior intent to open. The *intent to open* measure is a way to examine documents that the participant wanted to open and for quickly accessing which they might use a recommender system. The *interest* measure seeks to identify documents that participants were not explicitly looking for but may be interested to discover.

Our dataset contained responses for 1637 files of which we included 1613 in our analyses. We discarded 34 responses for which

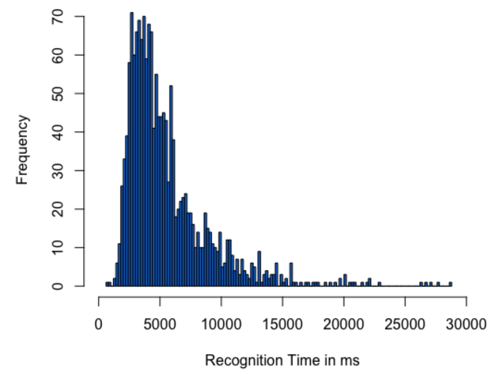

Figure 4: Distribution of recognition time – how long participants took to decide whether they recognize the documents – across all recency of access, richness of interaction, and presentation conditions.

the recognition time was more than 2 standard deviations above the mean as outliers [μ = 6576.63, σ = 11240.86 milliseconds]. The distribution of recognition time for the filtered dataset across all conditions is shown in Figure 4. To analyze the measure of intent to open and interest, we looked at that portion of the dataset for which these questions were answered (1168 documents for intent to open, 1163 for interest).

To investigate how different factors affect the outcome measures, we developed several (generalized) linear mixed effects models with the factor as the independent variable. In all our models, we considered participant as a random factor to account for variation in the outcome caused by an individual rather than the experimental factor. Each user would therefore have its own random intercept in the model, the parameters for which are estimated in addition to the coefficient for the experimental (fixed) factor. When testing recognition, intent, or interest the outcome measure, we used the function "glmer" from the R package "lme4" to define the model and fit it to our data. Because recognition, interest, and intent are Boolean-valued outcomes, we used the family function "Binomial" with the link "logit" to define models with these measures as dependent variables. The logit of the response variable, rather than the response variable itself, was used to accommodate the assumption of linear models that the residuals are normally distributed. When testing time which is a continuous-valued outcome, we used the function "lmer" from "lme4" to define the model.

Throughout this section, we present figures showing outcome means across different conditions. The error bars in these figures are standard errors around the mean. When we include participants' free-text responses, we identify each participant with an identifier of the form "p" + number.

4.1 Recency of Access

Figure 5 shows the recognition ratio and recognition time for each of the recency conditions. Figure 6 displays interest and intent to open the documents per recency condition. A Wald Chi-Square test on the fitted model to recognition as the dependent variable and recency of last access as the independent variable indicated that

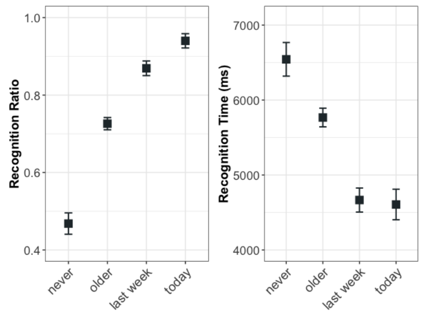

Figure 5: Recognition ratio and time per recency conditions.

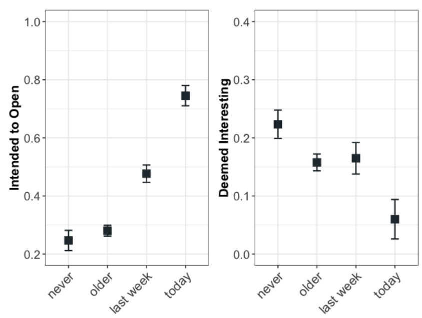

Figure 6: Ratio of documents that participants had a prior intent to open in the near future and those they deemed interesting across recency of access conditions.

Table 2: Pairwise comparison of recognition, recognition time, interest, and intent to open between adjacent recency conditions

Outcome	Recency conditions	z values
Recognition	Never vs Older	z=-8.18, p<0.01
	Older vs Last Week	z=-5.25, p<0.01
	Last Week vs Today	z=-2.07, p=0.11
Recognition Time	Never vs Older	z=3.80, p<0.01
	Older vs Last Week	z=4.30, p<0.01
	Last Week vs Today	z=0.65, p=0.87
Interest	Never vs Older	z=2.33, p=0.06
	Older vs Last Week	z=-0.66, p=0.87
	Last Week vs Today	z=1.74, p=0.22
Intent to Open	Never vs Older	z=-1.02, p=0.65
	Older vs Last Week	z=-6.49, p<0.01
	Last Week vs Today	z=-5.32, p<0.01

recency of last access has a significant effect on recognition [$\chi^2(3) = 154.61$, $p < 0.01$]. We performed similar tests for recognition time, interest, and intent [$\chi^2(3) = 62.30$, $p < 0.01$ for recognition time, $\chi^2(3) = 8.74$, $p = 0.03$ for interest, $\chi^2(3) = 125.79$, $p < 0.01$ for intent], all of which were also statistically significant. For significant effects, we then performed post hoc multiple comparisons tests using the function "glht" from the R package "multcomp" with Tukey adjustment. Table 2 shows the pairwise comparisons.

The results suggest that the more recent participants' last access to the document is, the more likely they are to recognize the document and take less time to do so. Interestingly, users also recognize those documents they have never opened more than 45% of the time. To gain more insight into why users recognize documents they have never interacted with, we analyzed participants' free-text responses. Some responses indicated that users in fact know about such documents from outside channels: *"[I recognized the document from] title and author, the file name is the same as a topic that was mentioned to me in a hallway conversation"* (p127 - *Basic* presentation), or that users have seen the documents but deferred opening them to a later time or have simply ignored them because they were not considered relevant at the time: *"The title and the author [who] had sent this document around via email...It has reminded me that I should probably review this."*) (p72 - *Advanced* presentation).

Participants are also more likely to have the intent to open recent documents compared to older ones. This could be because the older documents are no longer relevant to their currently active projects or those documents had not been of interest to them in the first place. However, participants seemed to be more interested in older documents and the ones they had never opened. In some cases, people knew about the document, but needed to be reminded: *"Need to refresh my memory on why the author shared it with me."* (p77 – *Basic* presentation), lost the document in the cloud: *"Author left the company and I was looking for it to save in my own OneDrive."* (p22 - *Basic* presentation), wanted to know about the recent changes to the document: *"Curious to know when this document was last edited - does it have the latest results?"* (p25 – *Basic* presentation), or encountering the document reminded them of a task: *"The doc reminded me that I need to start to prepare for the October [PROJECT NAME] meeting. Some of the material from the current doc will go into the Oct doc."* (p86 – *Basic* presentation). In other cases they were not aware of the document but the information on the document tile intrigued them. This interest sometimes arose from information related to the content: *"I didn't know there was a functional spec for this particular feature."* (p72 – *Advanced* presentation), and sometimes referred to other people who collaborated on the document: *"Author is a peer of mine, last edits has someone in my org chain."* (p70 – *Advanced* presentation). Interestingly, 38% of the documents that users assessed as interesting to access, were documents that they did not recognize. The rest were documents that they did recognize but had not planned to access.

4.2 Richness of Prior Interaction

Figure 7 displays recognition ratio and time across richness of interaction conditions. Figure 8 presents prior intent to open and interest for the same conditions. Similar to the tests for recency of last access, we performed Wald Chi-Square tests on the fitted model to our data to explain our outcome measures with richness of interaction as the independent variable. The tests showed that richness of prior interaction indeed has a significant effect on how likely participants are to recognize their documents [$\chi^2(2) = 182.36$, $p < 0.01$], how long it takes them to recognize the document [$\chi^2(2) = 60.97$, $p < 0.01$], whether they intend to open it [$\chi^2(2) =$

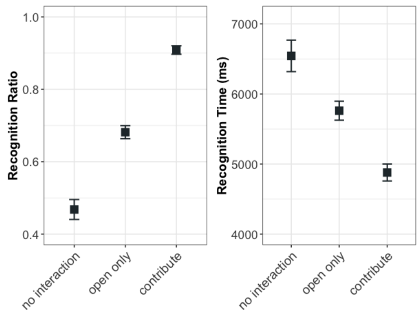

Figure 7: Recognition ratio and time per interaction conditions.

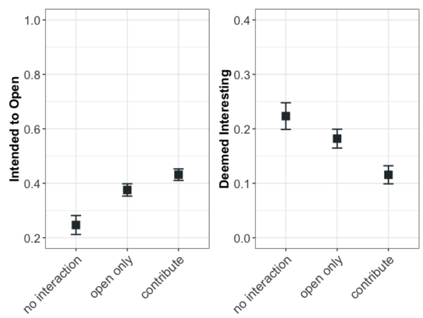

Figure 8: Ratio of documents that participants had a prior intent to open in the near future and those they deemed interesting across richness of interaction conditions.

Table 3: Pairwise comparison of recognition, recognition time, interest, and intent to open between adjacent interaction conditions

Outcome	Interaction Conditions	z values
Recognition	None vs Open Only	z=-6.66, p<0.01
	Open Only vs Contribute	z=-9.26, p<0.01
Recognition Time	None vs Open Only	z=3.39, p<0.01
	Open Only vs Contribute	z=5.17, p<0.01
Interest	None vs Open Only	z=1.25, p=0.36
	Open Only vs Contribute	z=2.54, p=0.02
Intent to Open	None vs Open Only	z=-3.22, p<0.01
	Open Only vs Contribute	z=-2.43, p=0.03

23.79, $p < 0.01$], and whether they find the document interesting [$\chi^2(2) = 12.00, p < 0.01$]. Table 3 displays the pairwise comparisons for these measures across interaction conditions.

The results indicate that the more involved the participants' interaction with the document has been, the more likely they are to recognize the document and take less time to do so. Interestingly, users recognize the documents they have only opened less than 70% of the time. Furthermore, participants are more likely to intend to open the documents with which they have had more involved interactions. However, it is the documents with which participants have had shallower interactions that they find interesting. Reasons for this interest were similar to those enumerated for the older documents in the previous section.

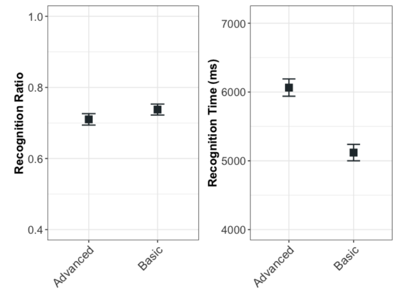

Figure 9: Recognition ratio and time across Advanced and Basic presentation conditions.

One alternative explanation for the high recognition ratio and the low recognition time for the documents in the *contribute* condition is that contribution can be an ongoing continuous activity. Therefore, the length of interaction and not the nature of that interaction may be responsible for the high recognition rate. To test this hypothesis, we separated out the files to which the user had contributed only once (169 documents). We then fit the model to the data from these files and the files the user had only opened and included recognition as a dependent variable. The interaction condition (*open only* or *contribute once*) served as the independent variable. A Wald Chi-Square test performed on the model revealed that effect of these conditions on recognition is indeed significant [$\chi^2(1) = 15.39, p < 0.01$]. A post hoc pairwise comparison indicated that users are more likely to recognize the documents they have contributed to even only once, compared to those documents they have ever only opened [$z = 3.92, p < 0.01$].

While contribution by itself can boost recognition, it is possible that it is correlated with more frequent Open actions which may serve as another reason why the documents the user has contributed to have such high recognition rates. To test this hypothesis, we constructed a model from the *open only* and *contribute once* conditions. We considered document opening frequency as the dependent variable and the interaction condition as the independent variable. A Wald Chi-Square test showed that interaction condition does not have a significant effect on how frequently the users open the documents [$\chi^2(1) = 2.96, p = 0.09$].

4.3 Presentation

Figure 9 shows recognition ratio and time for documents in the presentation conditions *Advanced* vs *Basic* across all recency and interaction conditions. Figure 10 shows users' prior intent to open documents and their interest in them across the two presentations. We performed Wald Chi-Square tests on the fitted model to the data to explain the outcomes with presentation as the independent variable. The tests showed that presentation has a significant effect on recognition time [$\chi^2(1) = 35.80, p < 0.01$]. As can be seen in Figure 9, the recognition time is higher for the *Advanced* presentation. This observation is expected because in the *Advanced* condition, there is more text to read and process on the document tile.

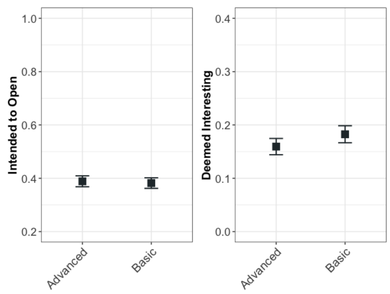

Figure 10: Ratio of documents that participants had a prior intent to open in the near future and those they deemed interesting across presentation conditions.

However, the effects of presentation on recognition, prior intent to open, and interest were not significant [$\chi^2(1) = 1.40, p = 0.24$ for recognition, $\chi^2(1) = 0.17, p = 0.68$ for intent, $\chi^2(1) = 1.23, p = 0.27$ for interest]. Because a user's prior intent to open a document should not depend on the presentation, we did not expect an effect of presentation for that measure.

We were surprised to not see an effect for *Basic* vs. *Advanced* presentations on recognition ratio. To gain a deeper understanding of how viewing actions of other collaborators influences recognition, we narrowed our dataset down to those documents on which any of the Edit, Comment, Reply, or Mention actions had been performed by the participant's collaborators because in the *Advanced* condition, these actions would be displayed on the document tile. The filtered dataset consists of 1022 documents. Similar to the results for the entire dataset, there was no significant difference in recognition across the two presentation conditions in this partition of the data [$\chi^2(1) = 2.09, p = 0.15$].

We further analyzed participants' free-text responses to gain a deeper insight into presentation elements that help with recognition. We found that document title and author were the most cited cues for recognition in either presentation condition: *"The file name is a common pattern that we use for team meetings, and the author [NAME] is a colleague."* (p127 - *Basic* presentation), *"Name is pretty self-explanatory. The author's name is familiar and confirmed the larger context in my recall for the document. The information that I last accessed it 7 months ago is additional information for the context/recall."* (p163 - *Advanced* presentation). Some participant responses however, indicated that viewing interaction-related contexts along with the document can be useful: *"I see others have been editing this long after I left my comments. I'd be curious to see what state the document is in now."* (p166), *"unlikely to be related to my immediate tasks. not recently edited by people related to my immediate tasks or workgroup."* (p161), *"Even though I don't recognize the document, I see that I opened it in the past, so I am curious about the content and the context."* (p25) all in the *Advanced* presentation.

4.4 Interaction Effects

We observed that both recency of last access and richness of interaction affect recognition ratio and recognition time. In Figures 11 and 12 we present how these measures are affected by the interaction of the two factors. One interesting observation from the figures is that users can recognize documents to which they contributed more than a week ago at a proportion as high as those that they have opened within the last 24 hours. In addition, users recognize documents to which they contributed longer than a week ago as fast as those they opened within the past week.

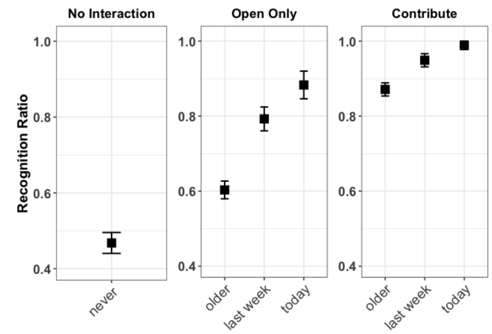

Figure 11: Recognition ratio across recency and interaction conditions.

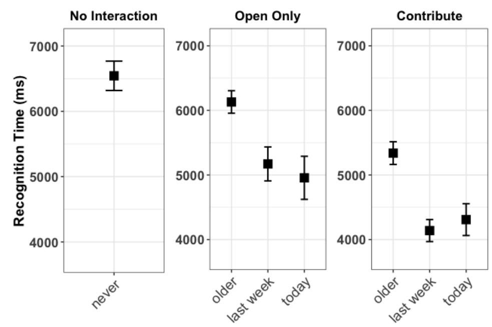

Figure 12: Recognition time across recency and interaction conditions.

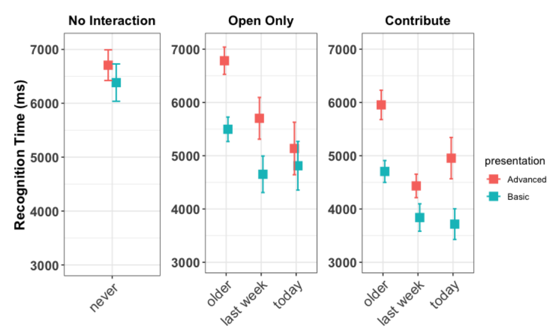

Figure 13: Recognition time across all presentation, recency, and interaction conditions.

Because all factors of last access, richness of interaction, and presentation impacted recognition time, we present how recognition time varies by the interaction of these 3 factors (see Figure 13).

4.5 Document Type

Although the selection criteria for users' documents did not include document type, it is possible that file type has an effect on our

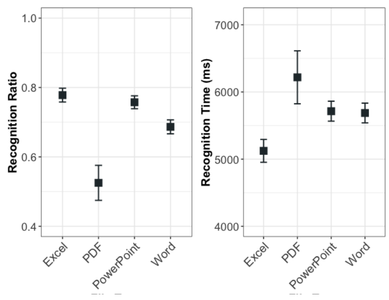

Figure 14: Recognition ratio and time per different file types.

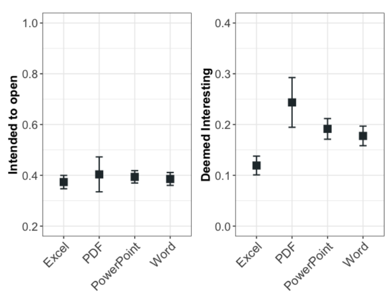

Figure 15: Intent to open and interest per different file types.

outcome measures. Figure 14 presents how recognition rate and time vary by document type. Figure 15 shows users' intent to open and interest per document type. Wald Chi-Square tests suggest that file type does in fact impact recognition and interest [$\chi^2(3) = 31.20$, $p < 0.01$ for recognition, $\chi^2(3) = 7.29$, $p = 0.06$ for recognition time, $\chi^2(3) = 9.11$, $p = 0.03$ for interest]. However, it does not appear to affect intent to open [$\chi^2(3) = 3.10$, $p = 0.38$].

A possible explanation for the lower recognition rate for PDFs could be that PDFs may serve as short-lived references. In contrast, other types of documents tend to evolve and therefore, may have a longer lifespan and relevance to the user's immediate work. The evolving of files surfaces especially in participants' responses about Excel files, which tend to be used for long-term book keeping: *"Using this for current do do list items, it's a live tracking doc"* (p61), *"will continue to use this file throughout the year for tracking purposes"* (p146). Because our participants' documents do not contain enough PDFs for which they answered the questions related to intent and interest, we cannot test this hypothesis. Future work is needed to understand consumption patterns for different types of recommended documents.

Although the difference in outcomes across file types is suggestive, our dataset does not contain equal number of documents of different types, and does not have a uniform distribution of file types across our experimental factors. Future work can further test this factor through a controlled experiment.

5 CONCLUSION AND FUTURE WORK

Our study investigated the effect of three dimensions of users' past interactions with documents recommended to them on their recognition of, prior intent to open, and interest in the documents.

Our results indicate that the more recent users' last access to a document is, they are more likely to recognize the document and in a shorter time. Users are also more likely to intend to open recent documents compared to older ones, potentially because recent documents may be more relevant to their ongoing projects. However, they find older documents, that they did not intend to open, more interesting when recommended to them. A similar pattern existed for the richness of user's past interactions with documents. If the user's past interactions with the document have been more involved (e.g., they edited the document in the past rather than having only opened it), they recognize the document with a higher probability, in a shorter time, and are more likely to intend to open it. However, those documents the user has had less involved interactions with are rated as more interesting for future access or discovery.

Our experiments on displaying past actions of users and their collaborators on the document tile did not yield a significant difference in recognition across conditions. However, it could be that our design of presenting such interactions or their granularity led to presentations that were not effective. Because we did not have access to the content of the documents, we could not reveal more context around the interactions for instance, by displaying that a certain collaborator has added or removed a section. Nevertheless, analyzing participants' free-text responses suggested that at times they find such context useful. Future research can further study under what circumstances displaying users' past interactions can be useful, for instance by investigating cases where there is close and constant collaboration on a document.

One of the most interesting findings of this work is characterizing the different scenarios a document recommender system needs to support. The study showed that users were interested in two types of recommendations. In some cases, they knew about a document and needed to get back to it. In this case, the recommendations helped them revisit known items more effectively. In other cases, they were not aware of the document but the information on the document tile intrigued them to access it. This observation suggests that document recommendation systems should support both re-finding known items and discovery of new ones and has implications for designing ranking algorithms that can support both scenarios. It also has implications for choosing the best way of presenting and explaining the recommendation to users.

Finally, our study had some limitations that could be addressed in future work. Although we used users' own documents, we did not capture their natural reasons for consuming recommended documents. For instance, we did not simulate whether people were turning to the document recommender to satisfy an acute information need (e.g., looking for a specific document to work on), or to discover new content. Such a setup is difficult to simulate in a field or lab study, and would require a large scale randomized deployment of different treatments. To alleviate this, we asked the user whether they had an intent to access the document to establish whether the intent to access the file had existed at the time of the study. Additionally, even though our study had over 100 participants who were able to access their own documents from their own machines, all participants were recruited from a single organization which could limit the generalizability of the findings. To partially address this, we included participants from various job roles (e.g., sales, software development, administrative, content writing, etc.).

REFERENCES

[1] Tarfah Alrashed, Ahmed Hassan Awadallah, and Susan Dumais. 2018. The Lifetime of Email Messages: A Large-Scale Analysis of Email Revisitation. In *Proc. CHIIR.*

[2] Ashton Anderson, Ravi Kumar, Andrew Tomkins, and Sergei Vassilvitskii. 2014. The dynamics of repeat consumption. In *Proceedings of the 23rd international conference on World wide web.* ACM, 419–430.

[3] Austin R Benson, Ravi Kumar, and Andrew Tomkins. 2016. Modeling user consumption sequences. In *Proceedings of the 25th International Conference on World Wide Web.* International World Wide Web Conferences Steering Committee, 519–529.

[4] Ofer Bergman, Ruth Beyth-Marom, Rafi Nachmias, Noa Gradovitch, and Steve Whittaker. 2008. Improved search engines and navigation preference in personal information management. *ACM Transactions on Information Systems (TOIS)* 26, 4 (2008), 20.

[5] Ofer Bergman, Steve Whittaker, Mark Sanderson, Rafi Nachmias, and Anand Ramamoorthy. 2010. The effect of folder structure on personal file navigation. *Journal of the American Society for Information Science and Technology* 61, 12 (2010), 2426–2441. https://doi.org/10.1002/asi.21415 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/asi.21415

[6] Hugo L Borges and Ana C Lorena. 2010. A survey on recommender systems for news data. In *Smart Information and Knowledge Management.* Springer, 129–151.

[7] Lukas Brozovsky and Vaclav Petricek. 2007. Recommender system for online dating service. *arXiv preprint cs/0703042* (2007).

[8] Jun Chen, Chaokun Wang, and Jianmin Wang. 2015. Will you" reconsume" the near past? fast prediction on short-term reconsumption behaviors. In *Twenty-Ninth AAAI Conference on Artificial Intelligence.*

[9] Andy Cockburn, Saul Greenberg, Bruce McKenzie, Michael Jasonsmith, and Shaun Kaasten. 1999. WebView: A graphical aid for revisiting Web pages. In *Proceedings of the OZCHI*, Vol. 99. 15–22.

[10] Andy Cockburn and Bruce McKenzie. 2001. What do web users do? An empirical analysis of web use. *International Journal of human-computer studies* 54, 6 (2001), 903–922.

[11] Paul Covington, Jay Adams, and Emre Sargin. 2016. Deep neural networks for youtube recommendations. In *Proceedings of the 10th ACM conference on recommender systems.* ACM, 191–198.

[12] Edward Cutrell, Susan T Dumais, and Jaime Teevan. 2006. Searching to eliminate personal information management. *Commun. ACM* 49, 1 (2006), 58–64.

[13] Susan Dumais, Edward Cutrell, Jonathan Cadiz, Gavin Jancke, Raman Sarin, and Daniel Robbins. 2003. Stuff I've Seen: A System for Personal Information Retrieval and Re-Use. *SIGIR Forum (ACM Special Interest Group on Information Retrieval)*, 72–79. https://doi.org/10.1145/860435.860451

[14] M. Grbovic, G. Halawi, Z. Karnin, and Y. Maarek. 2014. How many folders do you really need?: Classifying email into a handful of categories. In *Proc. CIKM.*

[15] Ziyu Guan, Can Wang, Jiajun Bu, Chun Chen, Kun Yang, Deng Cai, and Xiaofei He. 2010. Document Recommendation in Social Tagging Services. In *Proceedings of the 19th International Conference on World Wide Web (WWW '10).* ACM, New York, NY, USA, 391–400. https://doi.org/10.1145/1772690.1772731

[16] Ido Guy, Naama Zwerdling, David Carmel, Inbal Ronen, Erel Uziel, Sivan Yogev, and Shila Ofek-Koifman. 2009. Personalized recommendation of social software items based on social relations. In *Proceedings of the third ACM conference on Recommender systems.* ACM, 53–60.

[17] William Jones, Susan Dumais, and Harry Bruce. 2002. Once found, what then? a study of "keeping" behaviors in the personal use of web information. *Proceedings of the American Society for Information Science and Technology* 39, 1 (2002), 391–402.

[18] William Jones and Michele Tepper. 2008. RESOURCES-Reviews-Keeping Found Things Found: The Study and Practice of Personal Information. *netWorker: The Craft of Network Computing* 12, 1 (2008), 33.

[19] SeoYoung Lee and Junho Choi. 2017. Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity. *International Journal of Human-Computer Studies* 103 (2017), 95–105.

[20] Jiahui Liu, Peter Dolan, and Elin Rønby Pedersen. 2010. Personalized news recommendation based on click behavior. In *Proceedings of the 15th international conference on Intelligent user interfaces.* ACM, 31–40.

[21] Joel Mackenzie, Kshitiz Gupta, Fang Qiao, Ahmed Hassan Awadallah, and Milad Shokouhi. 2019. Exploring User Behavior in Email Re-Finding Tasks. In *The World Wide Web Conference (WWW '19).* ACM, New York, NY, USA, 1245–1255. https://doi.org/10.1145/3308558.3313450

[22] Jens-Erik Mai. 2016. *Looking for information: A survey of research on information seeking, needs, and behavior.* Emerald Group Publishing.

[23] Xueming Qian, He Feng, Guoshuai Zhao, and Tao Mei. 2013. Personalized recommendation combining user interest and social circle. *IEEE transactions on knowledge and data engineering* 26, 7 (2013), 1763–1777.

[24] Nelson Siu, Lee Iverson, and Anthony Tang. 2006. Going with the Flow: Email Awareness and Task Management. In *Proc. CSCW.* 441–450.

[25] Brent Smith and Greg Linden. 2017. Two decades of recommender systems at Amazon. com. *Ieee internet computing* 21, 3 (2017), 12–18.

[26] Sandeep Tata, Alexandrin Popescul, Marc Najork, Mike Colagrosso, Julian Gibbons, Alan Green, Alexandre Mah, Michael Smith, Divanshu Garg, Cayden Meyer, et al. 2017. Quick access: building a smart experience for Google drive. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* ACM, 1643–1651.

[27] Linda Tauscher and Saul Greenberg. 1997. How people revisit web pages: empirical findings and implications for the design of history systems. *International Journal of Human-Computer Studies* 47, 1 (1997), 97–137.

[28] Jaime Teevan. 2008. How people recall, recognize, and reuse search results. *ACM Transactions on Information Systems (TOIS)* 26, 4 (2008), 19.

[29] Jaime Teevan, Eytan Adar, Rosie Jones, and Michael AS Potts. 2007. Information re-retrieval: repeat queries in Yahoo's logs. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval.* ACM, 151–158.

[30] Jaime Teevan, Edward Cutrell, Danyel Fisher, Steven M Drucker, Gonzalo Ramos, Paul André, and Chang Hu. 2009. Visual snippets: summarizing web pages for search and revisitation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, 2023–2032.

[31] L. Cadiz J. Venolia, G. Dabbish and A. Gupta. 2001. *Supporting Email Workflow.* Vol. 2088.

[32] Sergey Volokhin and Eugene Agichtein. 2018. Understanding music listening intents during daily activities with implications for contextual music recommendation. In *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval.* ACM, 313–316.

[33] Kangning Wei, Jinghua Huang, and Shaohong Fu. 2007. A survey of e-commerce recommender systems. In *2007 international conference on service systems and service management.* IEEE, 1–5.

[34] Bo Xiao and Izak Benbasat. 2007. E-commerce product recommendation agents: use, characteristics, and impact. *MIS quarterly* 31, 1 (2007), 137–209.

[35] Lili Zhao, Zhongqi Lu, Sinno Jialin Pan, and Qiang Yang. 2016. Matrix factorization+ for movie recommendation.. In *In Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI'16).* 3945–3951.