Design and In-Situ Evaluation of a Mixed-Initiative Approach to Information Organization

Mona Haraty

Department of Computer Science, University of British Columbia

201-2366 Main Mall, Vancouver, BC, V6T 1Z4, Canada

1-778-994-2510
haraty@cs.ubc.ca

Zhongyuan Wang

Microsoft Research Asia, Beijing, China

Building 2, No. 5, Dan Ling Street, Haidian District, Beijing, P.R. China, 100080

+86-138-109-72076
Zhy.wang@microsoft.com

Helen Wang, Shamsi Iqbal, Jaime Teevan

Microsoft Research Redmond, Washington, United States

One Microsoft Way, Redmond, WA 98052, United States

425-707-4249, (425) 538-5568, 425-421-9299
{helenw, shamsi, teevan}@microsoft.com

Please send all correspondence to Mona Haraty at haraty@cs.ubc.ca

Abstract

Organizing personal information by folders or tags has proved to be effective for finding, remembering, and understanding information. However, past studies have shown that the cost of organization can be too high for some users to be worth the effort. Mixed-initiative approaches attempt to reduce the burden of manual organization by automatically identifying and suggesting organizational units such as folders to users. However, little is known about how such mixed-initiative approaches influence users' organizational experiences. In this paper, we explore a mixed-initiative approach that suggests high-level organizational units to users to facilitate email organization. In two in-situ experiments with 34 knowledge workers, we study how our mixed-initiative approach influenced users' experience with organization. We show that our approach made it easier to create organizational units without negatively affecting recall of those units, and led to the creation of units that otherwise would have not been created. Our findings suggest ways computers and people can most effectively work together to organize information.

Keywords: mixed-initiative; personal information management

Introduction

Organizing personal information is an important and onerous activity. Despite the difficult and time-consuming nature of organizing, people organize their personal information for many reasons, including but not limited to facilitating re-finding, supporting sensemaking, and managing tasks (Jones, 2007; Jones, Phuwanartnurak, Gill, & Bruce, 2005; Jones & Teevan, 2007; Whittaker, Matthews, Cerruti, Badenes, & Tang, 2011). The creation of new organizational units is especially difficult, incurring a significant cognitive burden (Cutrell, Robbins, Dumais, & Sarin, 2006; Whittaker et al., 2011; Whittaker & Sidner, 1996). To reduce the manual and cognitive burden of organizing, a large body of research focuses on using automatic approaches to support users (Bekkerman, 2004; Boone, 1998; Dredze, Lau, & Kushmerick, 2006; Gopsill, Payne, & Hicks, 2013; Kushmerick & Lau, 2005; Mock, 2001; Surendran, Platt, & Renshaw, 2005). Much of this work attempts to facilitate organization by classifying information into existing organizational units such as existing email folders, but a major challenge to filing is actually defining new units (Whittaker & Sidner, 1996). Automated clustering can be used to create new organizational units (Bekkerman, 2004; Surendran et al., 2005), but studies of such techniques have found that users are very intolerant of the inaccuracy of fully automated methods. Attempts to automate support are problematic for several reasons, including the fact that people are intolerant of errors and errors are more likely with the creation of new units. Effective naming of new units is especially difficult, as people can find it hard to remember and use automatically generated organizations (Lansdale, 1988).

In addressing the challenge of effective information organization, mixed-initiative methods—where the user and system share the responsibility of organizing (Horvitz, 1999)—have seen some success (Huang & Mitchell, 2009; Segal & Kephart, 1999; Whittaker et al., 2004). Mixed-initiative approaches may provide the right kind of balance — sufficiently automated to reduce the initial costs of organizing information while preserving an essential level of interaction between users and their information (Lansdale, 1988). However, previous mixed-initiative approaches did not

suggest new organizational units, and we know very little about how mixed-initiative approaches generally affect users' experience with organizing. Such approaches have been rarely studied from the perspective of the user and their evaluation has focused on assessing the quality of the suggested organizational units offered to people. In this paper, we aim to fill this gap by studying a mixed-initiative system to understand the effects of using such an approach on the information organizations people create, how they engage with the created organizational units, and how they later recall and recognize these organizational units for effective use. The mixed-initiative approach we study helps users organize their email into *activity* groupings intended to represent the activities that the user is involved in. As users identify activities, the system also clusters their emails and presents the results as *suggested activities*, an example of which shown in Figure 1. The user may then choose build on the suggested activity to actually create a new activity-based grouping of their email and other forms of information (e.g., files, web references).

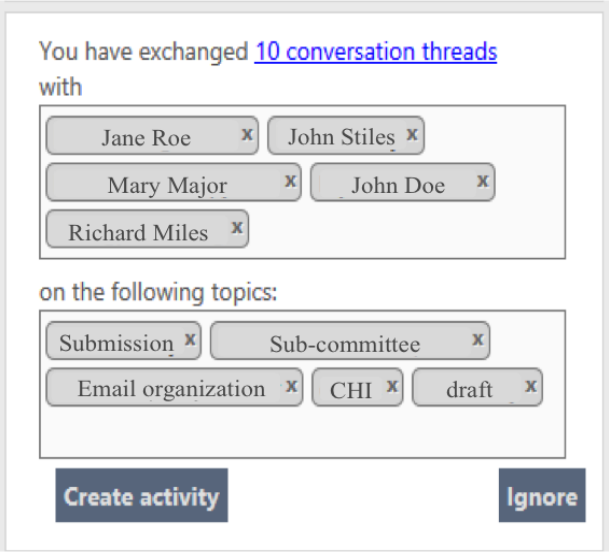

Figure 1: Suggesting high-level organizational unit (called *activity*) to users that helps to bootstrap the organization of email conversations.

In two in-situ studies with 34 knowledge workers, we explore whether our mixed-initiative approach to email organization can reduce the burden of organizing emails. We look at whether the use of suggestions influences the type of organization units (henceforth *activities*) created, how the activities are recalled and

recognized, and users' investment in them. We show that the mixed initiative approach made it easier to create activities than a manual method, and that activities suggested by the system helped participants create activities that they would have otherwise not created. Participants were able to recall the activities they created with suggestions as well as they were able to recall the ones created manually, indicating the suggested activities were salient. Overall, our results suggest that mixed-initiative approaches to organizing emails holds promise in reducing the burden of manual email organization for users without negatively affecting the importance of the organizational units created and how they are recalled.

Related Work

To address the problem of project fragmentation in PIM (Bergman, Beyth-Marom, & Nachmias, 2006), a large body of research has considered the notion of "activity" or "project" as new organization units for organizing personal information (Bannon & Bødker, 1997; Bellotti, Thornton, Chin, Schiano, & Good, 2007; Harrison, Cozzi, & Moran, 2005; Moran, Cozzi, & Farrell, 2005; W. Wang & Haake, 1997). Several approaches provide support for arriving at such organization units. For example, Bergman et al. proposed the single hierarchy solution to address the problem of project fragmentation. They suggested ProjectFolders as one possible implementation of such solution that allows users to organize all their project-related information in the same folder regardless of their form (e.g., email, files, web references, digital photos) (Bergman et al., 2006). The personal project planner supports creation of project plans as external representations that organize the tasks as well as various types of information needed to complete the tasks (Jones, Klasnja, Civan, & Adcock, 2008). UMEA creates project contexts through interaction histories; it monitors user activities and resources used when a user is working on a project and automatically organizes the resources to make them available when the user resumes working on a project (Kaptelinin, 2003). The Giornata system allows activity-based resource storage through tagging (Voida & Mynatt, 2009).

Somex, the system that we use in our study, has the same goal as the previous systems discussed above but it differs in its approach which relies on use of emails in a mixed-initiative clustering.

There is a substantial body of work on clustering *emails* into higher-level organizational units such as activity (Bekkerman, 2004; Boone, 1998; Dredze et al., 2006; Gopsill et al., 2013; Kushmerick & Lau, 2005; Mock, 2001; Segal & Kephart, 1999; Surendran et al., 2005). However, most of this work attempts to organize emails and other information items into *pre-existing* folders. For example, SwiftFile and BuzzTrack are two systems that have employed mixed-initiative approaches to help users organize their email messages. With SwiftFile the user selects one of three folders that are identified as most likely to be related to a message, or manually specifies another folder (Segal & Kephart, 2000). With BuzzTrack suggested topics and labels are overlaid on each conversation in the inbox to reduce the cost of tagging (Cselle, Albrecht, & Wattenhofer, 2007); users can accept the suggestions and only need to change the incorrect labels and topics. In another study, Dredze et al. have investigated use of several algorithms for classifying emails into existing activities (Dredze et al., 2006). Although this group of work has helped with reducing the burden of organizing emails into folders, they rarely address the burden of creating folders in the first place which is a major user filing problem (Whittaker & Sidner, 1996).

There have been some attempts toward helping users with creating new folders (Huang & Mitchell, 2009; Khoussainov & Kushmerick, 2005; Kushmerick & Lau, 2005; Mitchell, Wang, Huang, & Cheyer, 2006; Surendran et al., 2005). However, the initial organizational units are often too simple, containing only one conversation thread and thus not facilitating the collation of relevant messages (Bellotti, Ducheneaut, Howard, & Smith, 2002, 2003). In instances where the suggested organizational unit is more complex, detailed investigation of how a mixed initiative approach affects users' experience with information organization is missing. We know little about the nature of users' efforts, how willing they are to collaborate with a system in organizing their information, and what type of organizations they will create. Our work aims to fill this gap.

Using mixed-initiative approaches has clear benefits when compared to fully automatic approaches. For example, Huang and Mitchell's algorithm has shown significant improvements when user feedback has been incorporated in text and email clustering (Huang & Mitchell, 2006, 2009). To help users with organizing emails, Whittaker et al. experimented with a fully automated approach by clustering contacts based on frequency and reciprocity in email conversations (Whittaker et al., 2004). However, due to the users' dissatisfaction with the automatic technique (finding it neither intuitive nor useful for social communication tasks), they took a manual approach and supported users by showing them information about each of the contacts such as recency and frequency of communication, and users had to manually arrange contacts spatially to indicate their relationships.

A body of work explores the visual summary of emails by showing some email meta-data such as contacts, keywords, and attachments for understanding archives and better retrieval of emails (e.g., (Dumais et al., 2003; Hangal, Lam, & Heer, 2011; Ringel, Cutrell, Dumais, & Horvitz, 2003)). The goal of that work has often been to facilitate understanding of emails, reminiscing, and retrieval rather than bootstrapping the act of organizing emails which is the focus of our work.

We build on the body of work reported here and extend it by comparing a mixed-initiative approach with the manual approach of foldering to organizing emails to investigate the effect of organizing method on the type of organizations created, the importance of maintaining those organization units, and their recall as the effective use of organizational units relies on recalling them. Our work sheds light into users' preferences on different approaches to information organization.

Mixed-initiative approach to organizing

We implemented a mixed-initiative approach to organizing information in an activity-based system, called Somex (H. J. Wang et al., 2015), that enables organization of digital artifacts such as email, documents, contacts, webpages into organizing constructs called *activities*. Somex is designed to address the problem of information fragmentation (Jones, 2004) by enabling users to organize their information regardless of its type.

The Activity Creation Experience

Activities could be created manually, or as part of a mixed-initiative experience in response to automated suggestions. Once an activity is created, it is shown in the “My activities” Tab. Figure 2 shows how created activities, named “Fun,”, “Experiment,” “Somex Study,” are listed.

Manual Activity Creation

To manually create an activity into which users can organize their emails, users should specify the name of the activity, people who are involved in it, and keywords representing it (Figure 2); the latter two are provided to enable the system to suggest relevant email messages to an activity and thus making it easier for users to populate an activity once it is created. To add an information artifact—which can be an email, a file, a web page, a task, an appointment, or a contact—into an activity, users can drag and drop the artifact into the activity.

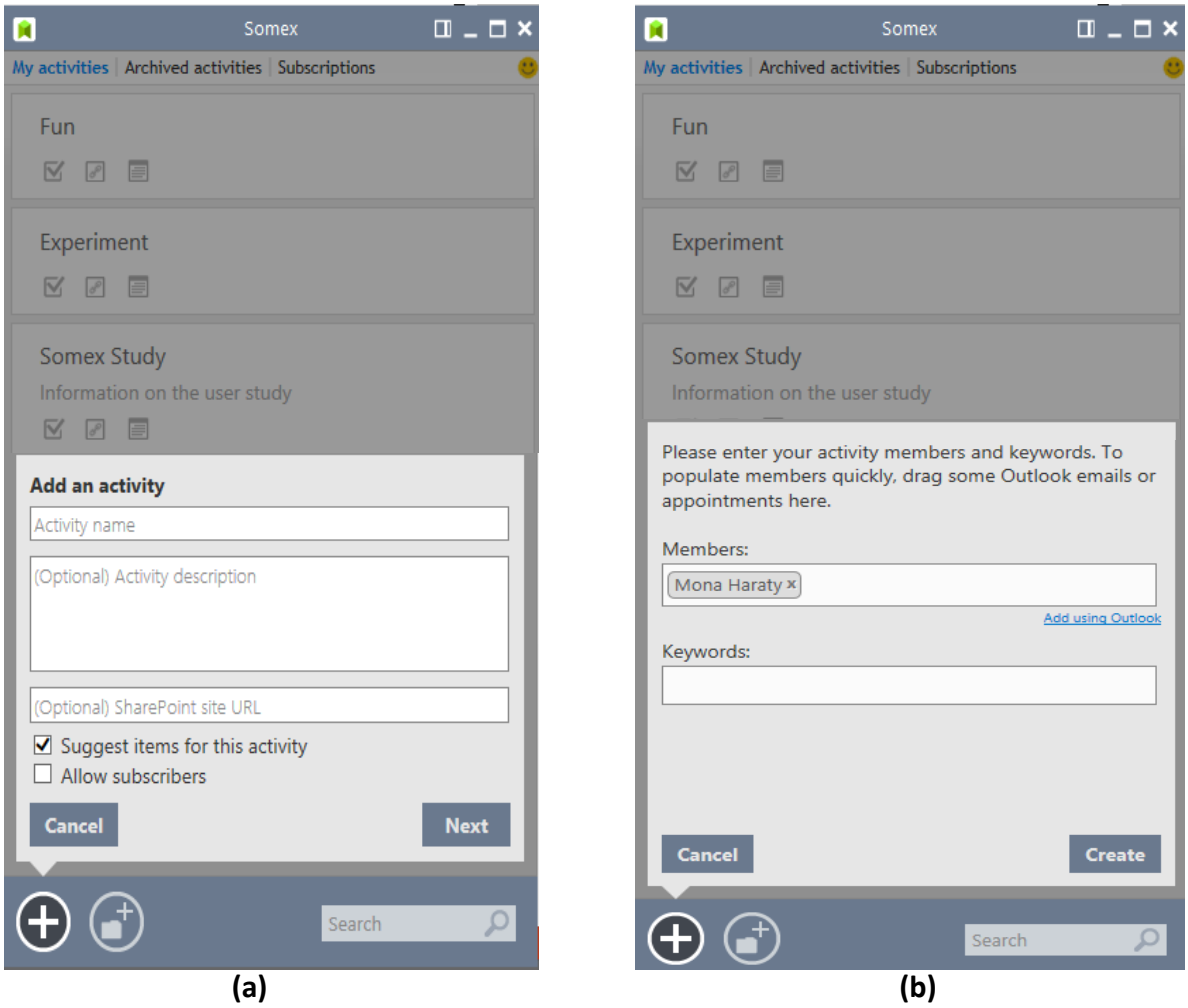

Figure 2: Creating activity manually. The user has clicked on the plus icon on the bottom left of the main view to create a new activity (a). In order for the system to suggest artifacts (emails) for this activity, the user has to enter the name of the people who are involved in this activity as well as the keywords related to the activity (b).

Mixed-Initiative Activity Creation

In our mixed-initiative approach, we facilitate creation of activities by clustering users' messages and identifying groups of relevant conversation threads. Our activity-suggestion mechanism presents its clustering results as a list of *suggested activities* to the user. Figure 3-b shows the first activity in the list of suggested activities. A suggested activity includes a group of relevant conversation threads, as well as extracted keywords and contacts from the emails; taking the form of "*You have exchanged 10 conversation threads with Bob, John, Sarah on the topics of Email organization, CHI, submission, draft*" (Figure 1). Users have the options of ignoring a suggested activity, which would remove it, or accepting it

to create an activity. All the elements of suggested activities are modifiable; users can remove any irrelevant keyword, member, or conversation that does not belong to the activity that they want to create and add more keywords or members to the suggestion. After creating an activity, the messages included in the activity were assigned a color-coded category with the activity name as its label in the email client, so people could distinguish which messages belong to which activities, in their email clients as well.

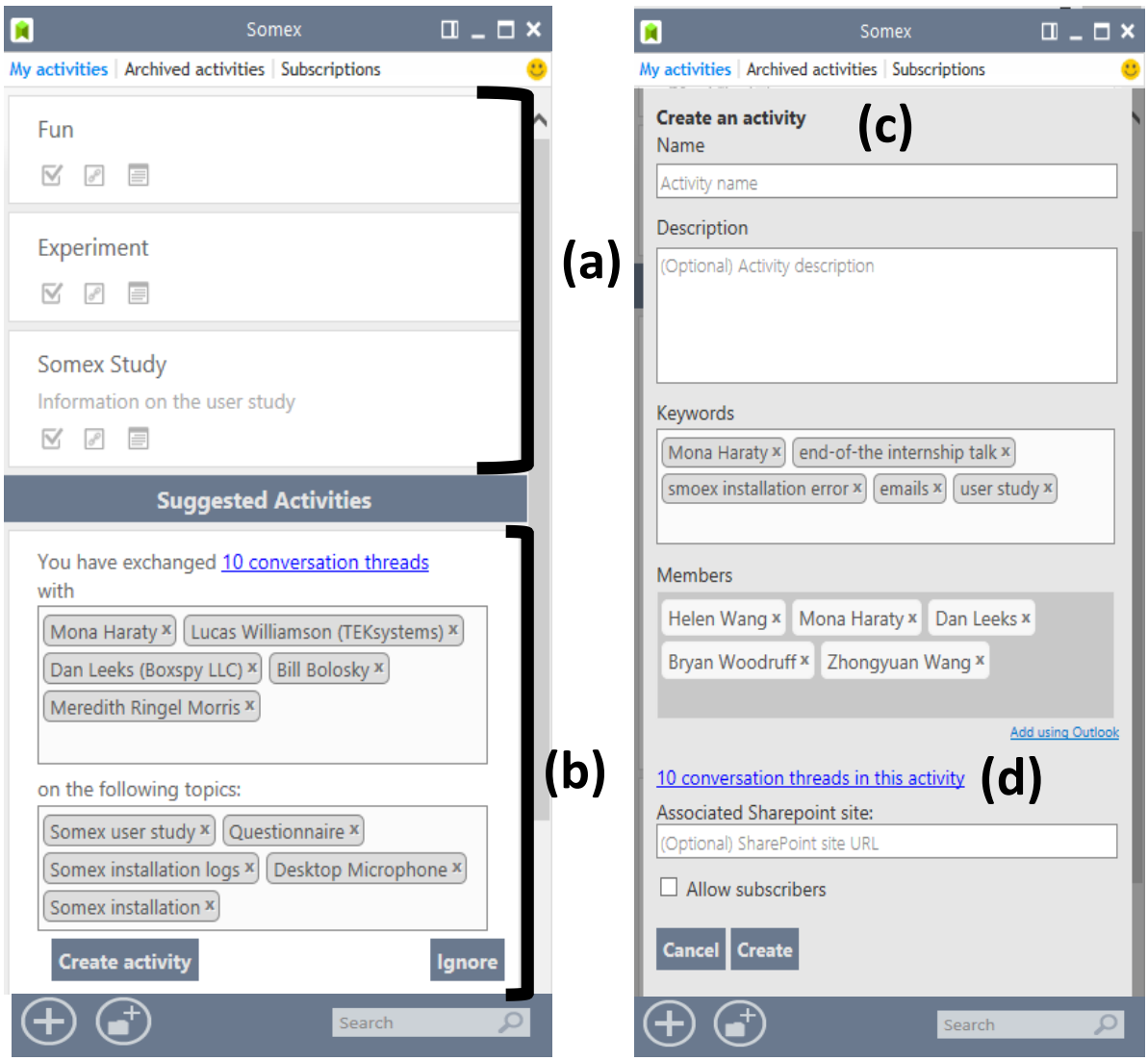

Figure 3: Creating activity in a mixed-initiative way using system suggested activities. (a) is the list of the activities (Fun, Experiment, Somex Study) created so far, (b) shows a system suggested activity, in the list of suggested activities, that includes 10 conversation threads with 5 people and 5 keywords. If the user decides to create an activity from the automated suggestion, she can click on the "create activity" button to go to (c) where she can name the activity (this is a different suggested activity than the one shown in b). The user can click on the conversation threads to see which threads are grouped together and to remove irrelevant ones (d). Users can add new or remove keywords and members.

Identifying Activity Suggestions from Email Messages

To cluster email messages into potential activities, we first extract the keywords in email messages (Grineva, Grinev, & Lizorkin, 2009). Taking a knowledge-based approach (Wang, Wang, & Wang, 2004), keywords are extracted from text without training data. Then, for each extracted keyword, the following features are constructed:

People feature: a vector of size *n*, where *n* is the total number of people (contact) in one's mailbox. The k^{th} dimension stands for how many times the k^{th} person is related with the keyword.

Date feature: a binary vector of size *n*, where *n* is the total time span (in date). The k^{th} dimension stands for whether the keyword appears in that day's mail.

Mail feature: a vector of size *n*, where *n* is the total number of mails. The k^{th} dimension stands for how many times the extracted keyword appears in the k^{th} email.

Forward/reply feature: a binary feature, if the mail is titled "FW:xxx" or "RE:xxx", the feature value is 1, otherwise it is 0.

Next, a refined *k*-Medoids clustering algorithm (Li, Wang, Zhu, Wang, & Wu, 2013) is run on the keywords, where the distance is based on cosine similarity between the feature vectors of two keywords. In this clustering algorithm, we use a threshold, which is tuned by experiments, to decide whether two keywords should be merged. If two keywords are merged, their feature vectors are also merged. To rank the clusters in terms of how likely they represent an activity, we use a liner model to combine the following signals:

Recency: the time distribution of the mails in a cluster, assigning decreasing weights to earlier timestamps of the mails.

*User engagemen*t: the number of mails sent by the user in a cluster.

Duration: last time stamp - first time stamp. If the messages of a cluster have a very short duration, then it is probably a very ephemeral or unimportant activity.

Burstiness: bursts of communications of an activity may represent an activity.

Regularity: regular communications at even intervals, such as recurring meetings, could represent an activity.

The input to the clustering algorithm was the most recent 500 email conversations of users from their Outlook inbox and any other email folders selected in the initial set-up of the system. The resulting clusters were shown to users as suggested activities in batches of five. There was a "show more suggested activities" button for users who wanted to see more. The number of contacts, keywords, and conversation threads included in a suggested activity were limited to the top five relevant ones, which were identified by the scoring signals described above. This was done to avoid showing too many potentially irrelevant entities in a single suggested activity, based on our pilot studies.

Methodology

The goal of our evaluation was to compare the two approaches of organizing email conversations into activities—manual vs. mixed-initiative. To do this comparison, we conducted two in-situ experiments with different goals. The goal of the first experiment was to do a within-subject comparison of the activities created manually with the ones created using system suggestions in terms of their importance, how well people recall them, and how willing people were to invest time to populate them. For example, we hypothesized people might not recall the activities they created using system suggestions or that they might end up in creating activities that are less important to them, just because they were suggested.

Additionally, we were interested to know if showing system-suggested activities to participants after they have created some activities on their own would cause them to create additional activities that they would have otherwise not created. To address the latter question, we conducted a second experiment, since the first experiment could not provide us with such understanding.

Both experiments were done at the participants' offices on their own computers using their own data, and the sessions were 35 to 70 minutes long. Below we describe the experiments' design and procedure, as well as the participants.

Experiments' Design and Procedure

Both experiments started with installing Somex on the participants' machines and providing a quick demo describing the relevant features of Somex. Participants were told that they could keep using the activities created by Somex after the study. The experiment's tasks and steps were then described. Both experiments comprised of five steps.

First, in both experiments participants were given time to create activities and populate each activity with at least one artifact such as an email, a task, an appointment, a bookmark, or a file — for maximum of 20 minutes in the first experiment and 10 minutes in the second. We gave participants a time limit to provide them with a sense of how long they should spend creating activities, but did not enforce the limit if they chose to spend more or less time. Activities were described to participants as "Organizational units that you can create to organize different types of information (email, files, bookmarks, to-dos) related to your activities in life such as a project, event organization, a trip, or a hobby." Participants were instructed to create as many activities as they could. The experiments differed in their first step:

In the first experiment, participants had both options of creating an activity manually and in a mixed-initiative way by using system suggestions for the whole duration of the experiment. This allowed us to do a within-subject comparison between the activities that were created manually and the ones created in a mixed-initiative way in terms of their importance, recall, and participants' investment in them. During the instructions, all the participants were first presented with the manual approach to creating an activity and then with the mixed-initiative approach. Participants were told that it is their choice which method to use to perform their task of creating activity and populating them with artifacts. Since both approaches to creating activity were available to the participants from the start, we refer to the first experiment as the *Suggestion+Manual* experiment.

In the second experiment, participants were initially exposed only to the manual approach to creating activities, and were introduced to the mixed-initiative approach once they were done creating activities on their own or after ten minutes. We refer to the second experiment as the *Manual-only-first*

experiment. Initially, they were told that they had ten minutes or until they were done to create as many activities as they were currently involved in, and to populate their activities with at least one artifact such as an email, a task, an appointment, a bookmark, or a file. When done, the experimenter then turned on the suggested activities and provided a quick demo of the second method of creating activity—which was using system suggested activities. Participants then were told that they have 10 more minutes or until they were done to have a look at the suggested activities to see if they wanted to create more activities or revise their existing activities. This experiment allowed us to understand if showing system-suggested activities to participants after they have created some activities on their own would cause them to create more activities that they would have otherwise not created.

In both experiments, participants were told to think aloud while performing their task of creating activities and populating them, and that for the purpose of the experiment they should prioritize creating as many activities as possible rather than populating each activity with a large number of artifacts. Both experiments shared the remaining following steps (second to fifth) after participants were done creating activities:

Second, we asked participants to talk about their current practices for organizing their emails, files, and bookmarks. The goal with this step was twofold: to distract participants from the activities they created since we wanted to assess how well they recall the activities, and to get a sense of their organizational practices.

Third, participants filled out a short online questionnaire on the usefulness of the suggested activities for creating activities, their preference, and whether suggested-activities remind them of any activities that they would have otherwise not created. See Appendix A for the list of questions participants were asked.

Fourth, they were asked to recall the name of the activities they created and to write them down on a piece of paper. Then, to score recall performance, we showed the participants the list of their activities so they could point to the activities that they did not remember. We were not concerned about the exact word-by-word match between the recalled activity name and the actual activity name.

Fifth, we asked them to provide meta-data about their activities; they rated the importance of each activity on a scale of 1-4 (1: unimportant, 2: slightly important, 3: important, 4: very important), they reported whether each activity was work-related or personal, and how likely they were on a scale of 1-5 to keep populating each of the created activities in the future. We did not define importance of an activity for the participants; we relied on their perception of importance.

Participants

Thirty-five participants (11 females) were recruited through snowball sampling. We excluded the data of one participant from our analysis, because she created some fake activities, such as 'test 1', despite the experimenter's instructions. Hence, we report data from 34 participants. Participants were randomly assigned to each experiment; 18/34 participated in the *first* experiment (Suggestion+Manual) and henceforth we refer to them as E1P01 to E1P18, and 16/34 participated in the *second* experiment (Manual-only-first) and henceforth we refer to them as E2P01 to E2P16. All were knowledge workers at Microsoft Research and were full time employees who have worked at the company for more than 4 months, with the exception of one participant (E2P10) who was an intern; 18 were researchers, 8 were developers/engineers, 5 were managers, 2 were admins. We included the criteria of having worked for more than 4 months in our call for participation to attract participants whose email inbox was old enough to represent a set of diverse activities that they have been involved in. Participants' age ranged from 24 to 54.

While reporting their existing email organization practices, most of the participants (25/34) stated that they organized emails into folders mainly for archiving important information or keeping their inbox clean. A common practice was to create a rule based folder for automatic filtering of messages of low priority (such as emails from mailing lists). Participants reported manually archiving messages that include important information or attachments (such as flight itineraries) that might be needed in the future. Another common behavior reported was manually organizing messages into folders, to keep the inbox clean so that the inbox can act as a to-do list. Overall, the participants' email organization practices seem to be consistent with the past studies of email use (Boardman & Sasse, 2004; Whittaker & Sidner,

1996; Whittaker et al., 2011). Even though most people had folders, they reported rarely opening their folders for *finding* emails, and instead relied on searching or a combination of searching and browsing their folders. This is consistent with the prior studies reporting that people rely on search for finding email messages as the email clients' search gets more powerful (Jones, Wenning, & Bruce, 2014; Whittaker et al., 2011).

Results

We collected usage log data while users were creating the activities, post-questionnaire responses, meta-data about the created activities, and notes of the participants' actions and comments made while performing their tasks, taken by the experimenter. We first describe how participants created and populated their activities in both experiments. Then we report descriptive statistics on the number of activities created, the amount of time spent on creating activities in each experiment, and characteristics of the activities such as their size (the number of artifacts such as email in them), importance (the importance as defined by the participants), and type (the nature of the activity, such as a collaborative project). Next, we present the results of investigating the effect of approach to organizing (manual vs. mixed-initiative) on the importance of the created activities, recalling them, and how much they were populated with artefacts; these analyses were done on the Suggestion+Manual experiment. Finally, we report the results of the post-questionnaire, showing how participants compared the two approaches to organization in both experiments.

Characteristics of the Activities Created

The type of activities created typically included collaborative projects—which were sometimes named as "collaboration with X" when the activity involved only another person—events to organize (e.g., conferences, workshops, visits, talks, social gatherings), trips, and personal activities (e.g., sports-related). While the participants were browsing or searching through their inbox to find relevant emails to add to the activities they had created, sometimes, they encountered emails that led them to create new activities, especially in cases where an email represented an outstanding task. As mentioned before, we used the notion of activity as an organization unit for organizing digital artifacts such as email. While the

notion of 'activity' helped participants in thinking of how to organize their emails, not all the messages that participants wanted to organize belonged to an activity. For example, E2P03 mentioned: "*my default is to map the folders that I have to activity but it's not right because some are not "activities"* ", therefore he did not create activities for organizing such emails from mailing lists. This suggests that notions used as organizational units need to be open to different interpretations to address different purposes for organizing.

To manually create an activity in the Manual-only-first experiment, participants started by performing a variety of actions: browsing their inbox, scanning their calendar, glancing at their email folders, or just starting based on what they had in mind. However, almost all the participants in the Suggestion+Manual experiment started with a quick scan of the suggested activities. To populate their activities, most participants searched their inbox for the people who were involved in their activities, and dragged and dropped the messages from the search results into their activities. Three of the 13 participants who already had folders corresponding to their activities dragged and dropped all or a subset of the messages in some of their folders to activities.

Most participants chose to see more suggested activities by clicking on "show more" button, because—as mentioned by many—they "*were curious to see what else the system has detected*." Many of the suggested activities in the Manual-only-first experiment, as mentioned by the participants while thinking aloud, mapped to an existing activity that had already been created in the first phase of the experiment and thus were ignored. When suggested activities represented activities that were completed, most participants ignored the suggestions, but two participants used them to create activities so that they can archive them. For example, E1P08 who created 2 activities from suggestions for this purpose said: "*I wouldn't remember when they are not recent activities. But I'd like to keep track of things that are ancient.*"

Number of Activities Created and Time Spent in both experiments

Participants created a comparable number of activities in both experiments: an average of six activities per participants. Participants in the Suggestion+Manual experiment, created a total of 108

activities, 52 using the suggestions and 56 manually, and participants in the Manual-only-first experiment, created a total of 106 activities, 42 using the suggestions and 64 manually.

Most of the suggested activities that were accepted by the participants were modified (78/94); only 17% (16/94) of the suggested activities were accepted without any modification. In all the modified suggested activities, participants either added more contacts or removed some irrelevant ones. But the conversation threads included in the suggested activities were modified in only 8 (10%) of them; participants could modify the conversations by removing the irrelevant ones. However, because the process of going through the list of conversations for finding the irrelevant ones was relatively time-consuming—compared to scanning the included contacts and keywords—often when a suggested activity had included a large set of irrelevant conversations, it was ignored altogether. The keywords were modified in 46 (59%) of the modified suggested activities. The large number of modifications made to suggested activities highlights importance of using a mixed-initiative approach as opposed to a fully automated one.

Table 1: summary of the number activities created (left of the table) and the time spent on creating them (right of the table) across the two experiments, using the two methods of manual and mixed-initiative. (middle) percentage of the activities that were not recalled.

Experiment	Method	**Number of activities created (count)**				**Activities not recalled (%)**	**Time spent on creating all the activities (min)**		
		Min	Max	Mean	Total/method		Min	Max	Mean
Suggestion+Manual	M	1	9	3.7	56	16% (9/56)	9.5	35	20.37
	S	1	8	3.3	52	17% (9/52)			
Manual-only-first	M	1	11	4	64	14% (9/64)	7	23	14.18
	S	1	8	2.8	42	43% (18/42)	3	12	6.5
Total					214	21% (45/214)			

In the Suggestion+Manual experiment, participants spent 9.5 to 35 minutes (mean = 20.37) on creating 4 to 10 activities (mean = 6) and on populating them: on average 3.7 activities were created manually and 3.3 using the suggestions. See Table 1 for a summary. Overall, 86% (93/108) of the

activities were work-related and the remaining were personal/social. The activity creation of the 18 participants in this experiment seemed to fall into four distinct patterns (see Figure 4):

(1) 8/18 participants started with using suggested activities and created all or all but one of their activities using the suggestions [E1P04, E1P03, E1P02, E1P12, E1P09, E1P08, E1P06, E1P15].

(2) 4/18 participants started creating activities on their own and none or only one of their activities was created using the suggestions [E1P16, E1P14, E1P10, E1P01]. When asked E1P10 about what he thought about the suggested activities at the end of experiment, he mentioned that "*Frankly, I didn't even notice them*." For the other three participants in this group, either the suggested activities did not match the ones they had in mind, or the suggested activities included a large number of conversation threads thus making it very difficult for them to go through all of them to remove the irrelevant ones. Although E1P01 attempted to remove some conversations from a suggested activity, he gave up in the middle and created that activity manually instead.

(3) 3/18 participants created their first activity using the suggestions, but the rest on their own [E1P07, E1P11, E1P17]. This group had similar reasons for not using the suggestions as the previous group.

(4) 2/18 started with creating activities on their own but alternated between the two methods or continued with using the suggestions [E1P05, E1P13, E1P18]. The quality of the suggested activities

were mixed for these participants.

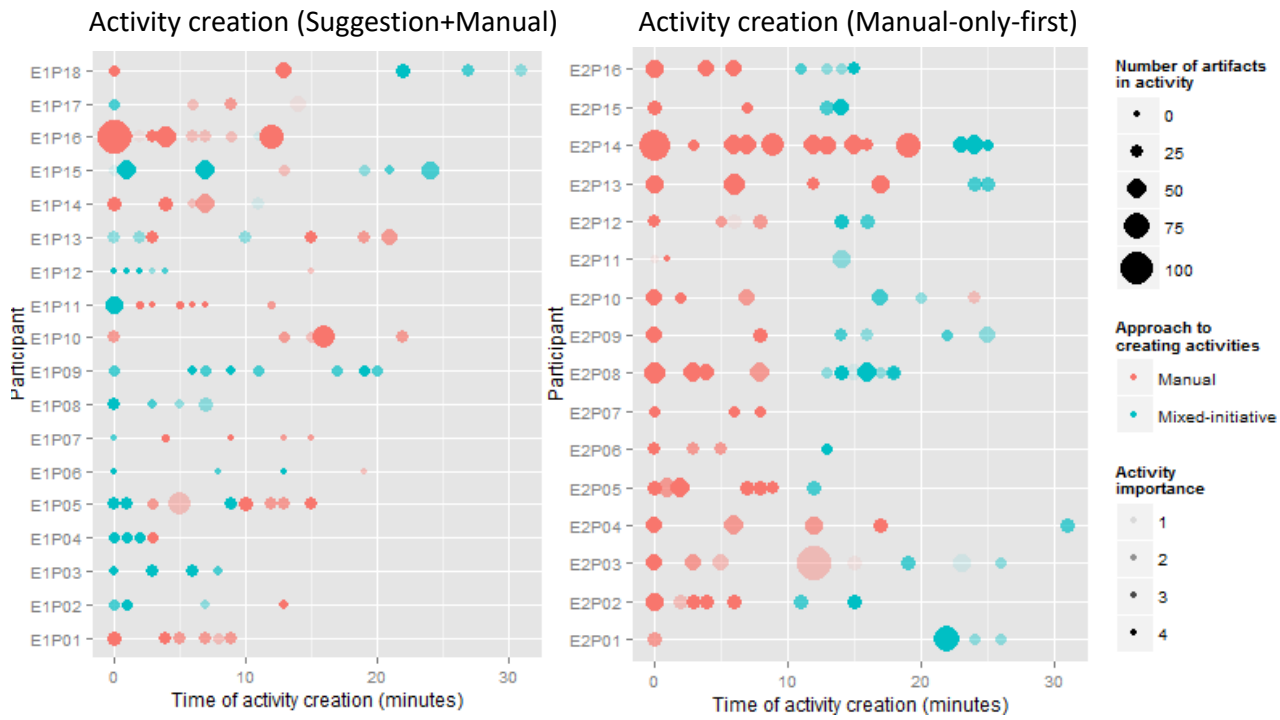

Figure 4: Activity creation over time: each dot represents an activity created; blue dots are activities created using suggestions and red dots are activities created manually; size of the dots represents the number of artifacts that were added to an activity during the experiment; color opacity represents importance of the activities to participants (the higher opacity represents higher importance); the horizontal axis is the time from when the first activity was created. The chart on the left shows the activities created by the participants of the first experiment (S+M). The chart on the right shows the activities created by the participants in the second experiment (Manual-only-first).

In the first phase of the Manual-only-first experiment, participants spent 7 to 23 minutes (mean = 14.18 minutes) to manually create 1 to 11 activities (mean = 4). After they were done with creating their activities and populating them, they were presented with suggested activities where they spent another 3 to 12 minutes (mean = 6.5) to create 1 to 8 activities (mean = 2.8) using the suggestions. See Table 1 for a summary. Overall, 94% (100/106) of the activities created in this experiment were work related. We conducted this experiment to see if participants will create activities using suggestions once they are done manually creating activities. As mentioned earlier, 42/106 activities in this experiment were created using suggestions. In addition, when we ran a regression analysis, we found no significant association between the importance of activities in this experiment and whether they were created manually or in a mixed-initiative way (p=0.84).

Mixed-Initiative vs. Manual Approach to Organizing in the Suggestion+Manual experiment

We ran various types of mixed-model regression analyses on the Suggestion+Manual data to investigate the effect of approach to organizing (manual vs. mixed-initiative) on the importance of the created activities, recalling them, and how much they were populated with artefacts as will be described below. In all the models, we included the fixed effects of activity creation approach (manual vs. using suggestions), activity importance, activity type (work-related vs. personal), time of activity creation (minutes from the start), and likelihood of populating activity in the future (rated by participants on scale of 1-5), as well as the random effect of participant. In this section, we report the results of these analyses.

No evidence of recalling activities being affected by the approach to creating them

Overall, 16% (18/108) of the created activities were not recalled; half (9/18) were created using suggested activities and half were created manually; 13/18 were rated as important; and 15/18 were work-related. To understand if recalling an activity was associated with whether or not an activity was created manually or using the suggestions, we ran a logistic mixed-model regression analysis. We did not identify any significant association between recalling an activity and the approach to creating it ($p=0.61$). None of the other included factors were significant: time of activity creation ($p=0.31$), importance ($p=0.48$), likelihood of populating activity ($p=0.06$), and activity type ($p=0.37$). In the distractor task, we asked participants to talk about their current practices for organizing their emails, files, and bookmarks, which might have affected the recall of the activities they had created. However, that should have affected recall of all activities regardless of the approach to their creation.

No evidence of any association between the importance of the created activities and the approach to creating them

To understand the effect of approach to creating activities on the importance of the created activities—an ordinal variable from 1 to 4—we ran an ordered mixed-model regression analysis. Approach to creating activities (manual vs. mixed initiative) ($p=0.84$) was not a significant factor. However, we found significant associations between importance and the following two factors: the time

of activity creation (p=0.005, coef = -0.08), and activity type (p<0.001, coef = 2.62); activities that were created earlier and work-related activities were all rated as more important.

Activities created using suggestions in the Suggestion+Manual experiment included more artifacts than the ones created manually

Participants in the Suggestion+Manual experiment populated their activities with 0 to 149 artifacts with a mean of 27.6. The artifacts were mostly email conversations. To understand if approach to organizing affects people's investment in organizing—as measured by the number of artifacts they added to their activities—we ran a Poisson regression analysis and found a significant association between approach to organizing and the number of artifacts in activities (p=0.035, coef=0.12); suggesting that activities that were created using suggestions were populated with 12% more artifacts than the activities created manually. However, this is not surprising given that most of the system-suggested activities already included some artifacts such as email conversations. We also found significant associations between the number of artifacts in activities and the following factors: activity importance (p < 0.001, coef=0.13), likelihood of populating activity (p<0.001, coef = 0.08), and activity type (p=0.004, coef = 0.23). Activities that were rated as more important were populated with 13% more artifacts compared to the less important activities. Not surprisingly, activities that were rated as more likely to be populated in the future were populated with 8% more artifacts than the ones rated less likely. Finally, work-related activities were populated with 26% more artifacts compared to personal activities.

Subjective Experience of Activity Creation

Ease of creating activities

Activity creation appeared to be easy when the suggestions were good. 73% (25/34) of the participants agreed or strongly agreed that using suggested activities made it easier for them to create activities; only 9% found it more difficult. For some participants, our clustering algorithm performed very well; for example, E1P09 who created all his 8 activities using suggestions said: "*I really liked the suggestions […] I already use OneNote for organizing my activities. This made me wish OneNote had suggestions."* After seeing the suggested activities, E2P09 said: "*Everything that I keep in my head, this*

tool is making them explicit...there is an overhead to make all these explicit but it's nice to get a bird eye view of all the pending things that I'm involved." However, when the clustering did not perform well, modifying a suggested activity was time-consuming and thus making it more difficult to create an activity from an inaccurate suggestion. E2P16 who disagreed with suggestions facilitating activity creation commented: *"The suggestions were mostly not helpful, so seemed like a bit of a waste of time to weed through them."*

Not all the suggested activities seemed worthy of becoming an activity; for example E2P03 said: *"Some of the suggestions were very helpful, many others identified communication threads that I didn't want (or need) to represent by activities."* To summarize, the factors that seemed to make it difficult for a few participants to create activities using suggestions were inaccuracy of the clustering algorithm, suggested activities that combined multiple activities of users, or multiple suggestions for a single activity of users.

Ease of populating activities

Similar to activity creation, whether or not using suggestions facilitated activity population depended on how relevant the conversations in a suggested activity were to each other. 20/34 agreed or strongly agreed with suggested activities facilitating populating activities. For some participants who already had many folders to organize their emails, a set of relevant conversations in a suggested activity had little effect on effort needed to populate an activity. For example, E2P03 who identified himself as an aggressive filer populated two of his activities with going to his corresponding email folders and selecting and dragging all the conversations to his activity. He commented: "*Some of the suggestions did track down messages spread across e-mail folders. However, I do typically aggressively file my e-mail away in folders, so the benefit of its suggestions were somewhat diminished*."

Preference about when to see suggested activities

Seeing the suggested activities *before* manually creating activities was generally preferred to seeing them only after manual creation. To see whether participants had any preference as to when to see suggested activities, we asked them if they preferred to see them before creating activities manually;

which 70% (24/34) of the participants agreed to. Almost all the participants in the Suggestion+Manual experiment (16/18) preferred seeing the suggested activities before creating activities manually, except for E1P16 and E1P18 (who was neutral); E1P16 said: "*I prefer to first make my own set of activities then have the suggestions maybe help populate those activities or find something I missed, but I think I'd rather not have suggestions till I've made more activities as a base set. Though I can see someone wanting to use suggestions as the bootstrapping for the entire thing*." Some participants preferred seeing the suggestions first, because they could avoid the manual work of creating an activity: "*I had a good idea of which activities I wanted to create, but I wanted to find the matching suggestion for them, so I didn't have to do the manual work of creating it. So, I would prefer to see the suggested activities first. Also, one of the suggestions was for an activity I wouldn't have necessarily thought to create on my own*" [E1P13]—who alternated between creating activities with and without suggestions.

Among the participants in the Manual-only-first experiment, 4/16 preferred their own experience of creating their activities first manually to seeing the suggestions from the beginning; for example, E2P03 and E2P13—similar to E1P16—said: *"I probably would have preferred to take a first pass myself at creating activities, but would then want to adjust/augment them with the system suggestions"* [E2P03], *"I would prefer to create activities on my own, to stay consistent with my normal tracking process"* [E2P13]. However, half of the participants in this experiment (8/16) preferred seeing the suggested activities from the beginning. The main reason as mentioned by some was that several of the suggested activities were overlapping with the ones they had already created manually. These participants thought that if they had seen the suggested activities first, "*it would help getting started*" [E2P04] or "*it saves time*" [E2P08]. This was consistent with some of the comments made by some of participants in the Suggestion+Manual experiment (see above).

Role of suggested activities in reminding of forgotten activities

The keywords and the contacts in the suggested activities reminded 68% (23/34) of the participants of some of their activities: "*There were some transient activities I was previously engaged in or am in engaged in now that were suggested -- I had forgotten all about them!"*[E2P03], *"it did show me*

many activities I otherwise wouldn't have remembered, but I didn't see some suggestions that I was expecting to see" [E1P13]. While looking at the suggested activities, E2P10 mentioned: *"It's finding some stuff here that I kinda lost track of."* The reminding feature of suggested activities, however, can be related to several factors such as the number of activities that an individual is currently involved in and whether a suggested activity represented an old, recent, or an ongoing activity. For example, E2P08 did not find any of the suggestions reminding him of any activity that he had forgotten: "*The suggested activities were all close in time, so I remembered them*." E2P12 did not find the keywords helpful: "*Most of the keywords were not something I would have associated with an activity*." Similarly, E1P02 did not find any value in the keywords, but found the contacts useful.

Discussion

Here, we discuss our findings and suggest ways to improve mixed-initiative approaches to email organization.

Using the mixed-initiative approach reduced the burden of organizing

Using mixed-initiative approach reduced the manual burden of creating activity for most of our participants: 73% found it easier to create activities based on suggestions, while only 9% found it more difficult. The fact that 83% of the suggested activities were turned into activities only after manual modifications indicates that the success of the activity-suggestion system was tied to keeping the user in the loop. While inaccuracy of the clustering algorithm in some cases (e.g., unrepresentative keywords, unrelated conversations, unrelated people) made it difficult for some to use suggested activities, several participants in the Manual-only-first experiment commented that suggested activities were similar to the activities they had already created manually, indicating the high quality of the suggested activities.

Improvements to the mixed-initiative approach

Other than improving the accuracy of the algorithm, we think that the following features could greatly improve the usefulness of a mixed-initiative approach: 1) allow merging multiple suggested activities into a new or an existing activity, 2) allow splitting a suggested activity into several activities, 3) suggest new modifications based on the modifications made by user—for example modifying contacts or

keywords in a suggested activity can often be linked to adding or removing conversations accordingly. In addition, since it was easy to create an activity using system suggestions, a few participants forgot that they had already created the activity and they created duplicate activities. One of the participants who did so said: *"I should be careful not to duplicate activities that already exist"* [E1P11]. This is similar to creating duplicate folders; a well-known problem with foldering (Whittaker & Sidner, 1996). To prevent users from creating duplicate organizations, when users are creating new activities, we suggest that the system should provide them with a list of similar activities that they have created by detecting the overlap between them.

No evidence of any difference in recalling the activities created based on suggestions and the ones created manually

We did not identify any negative effect of creating activities based on suggestions on participants' ability to recall the activities. This suggests that using a mixed-initiative approach to creating activity does not hurt users' ability to recall the activities. Another possibility is that we might have missed detecting an effect due to the way we assessed recall: asking the participants to recall the name of the activities they created after a 10-minute distractor task. Although the 10-minute distraction was short, it seemed to be long enough for participants to forget some of the activities they created: 45/214 activities were not recalled. One thing that might have confounded recall though was that some participants had email folders that directly matched the activities they created, so it was easier for them to recall the activities they created.

No evidence of difference in the importance of the activities created based on suggestions and the ones created manually

We had hypothesized that users might end up creating less important activities just because they were suggested. However, we did not identify any significant difference between the importance of the activities that were created manually and that of the activities created using the suggestions in the Suggestion+Manual experiment, where participants had the opportunity of using suggestions to create activities at any point during the experiment. Similarly, we did not find any association between activity

importance and the approach to creating it in the second experiment. This latter finding was somewhat surprising given that participants in that experiment were done creating their activities when they were presented with the system suggestions. Not rejecting the null hypothesis does not confirm that there is no difference in the importance of the activities, and calls for further research in order to draw some conclusions.

Combine mixed-initiative approach to organizing with rule-based approaches

Creating rules for automatically organizing emails into folders was another approach that our participants reported as one of their organizational practices, but it was mostly used for avoiding the messages sent to mailing lists from cluttering their inbox. We think that a mixed-initiative approach can benefit from combining with rule-based approaches to organizing, where users can define their own rules as to what might represent an activity. This is useful, especially given the diversity of what an activity means to different individuals. For example, we found that for some of our participants, all the conversations exchanged with a single person belonged to a single activity, but for others conversations with a common group of people belonged to several different activities.

Limitations

While it is important to assess whether people would remember the organizations they create when using the system-suggested organizations, we acknowledge that the recall task we used is not a common task that users would perform. To assess impact of the activity as organization units on re-finding or recall, re-finding tasks similar to those used in prior work can be used (Bergman, Whittaker, Sanderson, Nachmias, & Ramamoorthy, 2010; Capra III & Pérez-Quiñones, 2005; Jones et al., 2014)**.**

While our exploratory study was a first step toward addressing our questions about a mixed initiative approach to information organization, a longitudinal study would address the questions more adequately.

Conclusion

We designed and implemented a mixed-initiative approach to organizing emails where high-level organizational units (activities), identified by clustering messages in the user's inbox, were suggested to facilitate creation of new activities. Our study is a step toward evaluating the effect of mixed-initiative organizational approaches on personal information management behaviours such as organizing, recalling, and maintaining (Jones & Teevan, 2007). Our study suggests there is promise for such approaches when organizing emails. Most of our participants (73%) found it easier to create activities using the mixed-initiative approach than with the manual method, and suggested activities reminded 68% of the participants of some activities they would not have otherwise created. In addition, we did not identify any negative effect of creating activity using a mixed-initiative approach on either the participants' ability to recall their activities, or the importance of the activities created.

We discussed several ways to improve our mixed-initiative approach that we hope can inform the design of future mixed-initiative approaches to organizing information. Longitudinal studies are needed to understand longer-term effects of mixed-initiative approaches on people's information organizational behaviours. For example, it is important to understand how use of these high-level organizational units affects re-finding practices. Previous research showed that people who receive highly threaded emails are more successful at retrieving messages (Whittaker et al., 2011). Displaying conversation threads grouped by *activities* in one's inbox could potentially have similar benefits as threads have in re-finding emails.

This work has broader implications, beyond email organization, for systems that support people to organize their information of different types. In our mixed-initiative approach, we used email to suggest activity-based grouping of information to users, however the created activities can also be used to organize other types of information as our participants did in this study. This approach could be of great benefit to activity-based systems like (Muller, Geyer, Brownholtz, Wilcox, & Millen, 2004; Voida & Mynatt, 2009) where users have often reported difficulty in creating new activities.

References

Bannon, L., & Bødker, S. (1997). Constructing common information spaces. In *Proceedings of the fifth conference on European Conference on Computer-Supported Cooperative Work* (pp. 81–96). Lancaster, UK: Kluwer Academic Publishers. Retrieved from http://portal.acm.org/citation.cfm?id=1241986

Bekkerman, R. (2004). Automatic categorization of email into folders: Benchmark experiments on Enron and SRI corpora. Retrieved from http://scholarworks.umass.edu/cs_faculty_pubs/218/?utm_source=scholarworks.umass.edu%2Fcs_faculty_pubs%2F218&utm_medium=PDF&utm_campaign=PDFCoverPages

Bellotti, V., Ducheneaut, N., Howard, M., & Smith, I. (2002). Taskmaster: recasting email as task management. *PARC, CSCW*, *2*. Retrieved from http://www.researchgate.net/publication/228912071_Taskmaster_recasting_email_as_task_management/file/72e7e51700e6071e30.pdf

Bellotti, V., Ducheneaut, N., Howard, M., & Smith, I. (2003). Taking email to task: the design and evaluation of a task management centered email tool. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 345–352). New York, NY, USA: ACM. http://doi.org/http://doi.acm.org.proxy.lib.sfu.ca/10.1145/642611.642672

Bellotti, V., Thornton, J., Chin, A., Schiano, D., & Good, N. (2007). TV-ACTA: embedding an activity-centered interface for task management in email. *Work*, *5*, 11.

Bergman, O., Beyth-Marom, R., & Nachmias, R. (2006). The project fragmentation problem in personal information management. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 271–274). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1124813

Bergman, O., Whittaker, S., Sanderson, M., Nachmias, R., & Ramamoorthy, A. (2010). The effect of folder structure on personal file navigation. *Journal of the American Society for Information Science and Technology*, *61*(12), 2426–2441.

Boardman, R., & Sasse, M. A. (2004). “Stuff goes into the computer and doesn’t come out”: a cross-tool study of personal information management. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 583–590). Vienna, Austria: ACM. http://doi.org/10.1145/985692.985766

Boone, G. (1998). Concept Features in Re:Agent, an Intelligent Email Agent. In *Proceedings of the Second International Conference on Autonomous Agents* (pp. 141–148). New York, NY, USA: ACM. http://doi.org/10.1145/280765.280791

Capra III, R. G., & Pérez-Quiñones, M. A. (2005). Using web search engines to find and refind information. *Computer*, *38*(10), 36–42.

Cselle, G., Albrecht, K., & Wattenhofer, R. (2007). BuzzTrack: Topic Detection and Tracking in Email. In *Proceedings of the 12th International Conference on Intelligent User Interfaces* (pp. 190–197). New York, NY, USA: ACM. http://doi.org/10.1145/1216295.1216331

Cutrell, E., Robbins, D., Dumais, S., & Sarin, R. (2006). Fast, Flexible Filtering with Phlat. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 261–270). New York, NY, USA: ACM. http://doi.org/10.1145/1124772.1124812

Dredze, M., Lau, T., & Kushmerick, N. (2006). Automatically Classifying Emails into Activities. In *Proceedings of the 11th International Conference on Intelligent User Interfaces* (pp. 70–77). New York, NY, USA: ACM. http://doi.org/10.1145/1111449.1111471

Dumais, S., Cutrell, E., Cadiz, J., Jancke, G., Sarin, R., & Robbins, D. C. (2003). Stuff I’Ve Seen: A System for Personal Information Retrieval and Re-use. In *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval* (pp. 72–79). New York, NY, USA: ACM. http://doi.org/10.1145/860435.860451

Gopsill, J. ., Payne, S. J., & Hicks, B. J. (2013). An Exploratory Study into Automated Real-Time Categorisation of Engineering E-Mail. In *2013 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 4806–4811). http://doi.org/10.1109/SMC.2013.818

Grineva, M., Grinev, M., & Lizorkin, D. (2009). Extracting Key Terms from Noisy and Multitheme Documents. In *Proceedings of the 18th International Conference on World Wide Web* (pp. 661–670). New York, NY, USA: ACM. http://doi.org/10.1145/1526709.1526798

Hangal, S., Lam, M. S., & Heer, J. (2011). MUSE: Reviving Memories Using Email Archives. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology* (pp. 75–84). New York, NY, USA: ACM. http://doi.org/10.1145/2047196.2047206

Harrison, B. L., Cozzi, A., & Moran, T. P. (2005). Roles and relationships for unified activity management. In *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work* (pp. 236–245). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1099245

Horvitz, E. (1999). Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 159–166). Pittsburgh, Pennsylvania, United States: ACM. http://doi.org/10.1145/302979.303030

Huang, Y., & Mitchell, T. M. (2006). Text Clustering with Extended User Feedback. In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 413–420). New York, NY, USA: ACM. http://doi.org/10.1145/1148170.1148242

Huang, Y., & Mitchell, T. M. (2009). Toward Mixed-Initiative Email Clustering. In *AAAI Spring Symposium: Agents that Learn from Human Teachers* (pp. 71–78). Retrieved from http://www.aaai.org/Papers/Symposia/Spring/2009/SS-09-01/SS09-01-011.pdf?origin=publication_detail

Jones, W. (2004). Finders, keepers? The present and future perfect in support of personal information management. *First Monday*, *9*(3). Retrieved from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/1123

Jones, W. (2007). Keeping Found Things Found: The Study and Practice of Personal Information Management (Interactive Technologies)(Interactive Technologies).

Jones, W., Klasnja, P., Civan, A., & Adcock, M. L. (2008). The Personal Project Planner: Planning to Organize Personal Information. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 681–684). New York, NY, USA: ACM. http://doi.org/10.1145/1357054.1357162

Jones, W., Phuwanartnurak, A. J., Gill, R., & Bruce, H. (2005). Don't take my folders away!: organizing personal information to get ghings done. In *CHI'05 extended abstracts on Human factors in computing systems* (pp. 1505–1508). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1056952

Jones, W., & Teevan, J. (2007). *Personal information management*. University of Washington Press.

Jones, W., Wenning, A., & Bruce, H. (2014). How Do People Re-find Files, Emails and Web Pages? Retrieved from https://www.ideals.illinois.edu/handle/2142/47300

Kaptelinin, V. (2003). UMEA: translating interaction histories into project contexts. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 353–360). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=642673

Khoussainov, R., & Kushmerick, N. (2005). Email Task Management: An Iterative Relational Learning Approach. In *CEAS*. Retrieved from http://egov.ufsc.br/portal/sites/default/files/anexos/5859-5851-1-PB.pdf

Kushmerick, N., & Lau, T. (2005). Automated email activity management: an unsupervised learning approach. In *Proceedings of the 10th international conference on Intelligent user interfaces* (pp. 67–74). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1040854

Lansdale, M. W. (1988). The psychology of personal information management. *Applied Ergonomics*, *19*(1), 55–66.

Li, P., Wang, H., Zhu, K. Q., Wang, Z., & Wu, X. (2013). Computing Term Similarity by Large Probabilistic isA Knowledge. In *Proceedings of the 22Nd ACM International Conference on Conference on Information & Knowledge Management* (pp. 1401–1410). New York, NY, USA: ACM. http://doi.org/10.1145/2505515.2505567

Mitchell, T. M., Wang, S. H., Huang, Y., & Cheyer, A. (2006). Extracting Knowledge about Users' Activities from Raw Workstation. In *AAAI 2006* (Vol. 1, p. 181). AAAI Press. Retrieved from http://www.aaai.org/Papers/AAAI/2006/AAAI06-029.pdf

Mock, K. (2001). An experimental framework for email categorization and management. In *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 392–393). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=384033

Moran, T. P., Cozzi, A., & Farrell, S. P. (2005). Unified activity management: supporting people in e-business. *Communications of the ACM*, *48*(12), 67–70.

Muller, M. J., Geyer, W., Brownholtz, B., Wilcox, E., & Millen, D. R. (2004). One-hundred Days in an Activity-centric Collaboration Environment Based on Shared Objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 375–382). New York, NY, USA: ACM. http://doi.org/10.1145/985692.985740

Ringel, M., Cutrell, E., Dumais, S., & Horvitz, E. (2003). Milestones in time: The value of landmarks in retrieving information from personal stores. In *Proc. Interact* (Vol. 2003, pp. 184–191). Retrieved from http://books.google.ca/books?hl=en&lr=&id=PTg0fVYqgCcC&oi=fnd&pg=PA184&dq=stuff+I%27ve+seen+merrie+ringle&ots=O9KIDIfyz0&sig=n0xFOZdva20Z_NlDvVndDn88rd0

Segal, R. B., & Kephart, J. O. (1999). MailCat: An Intelligent Assistant for Organizing e-Mail. In *Proceedings of the Third Annual Conference on Autonomous Agents* (pp. 276–282). New York, NY, USA: ACM. http://doi.org/10.1145/301136.301209

Segal, R. B., & Kephart, J. O. (2000). Swiftfile: An intelligent assistant for organizing e-mail. In *AAAI 2000 Spring Symposium on Adaptive User Interfaces, Stanford, CA*. Retrieved from http://www.aaai.org/Papers/Symposia/Spring/2000/SS-00-01/SS00-01-023.pdf

Surendran, A. C., Platt, J. C., & Renshaw, E. (2005). Automatic Discovery of Personal Topics to Organize Email. In *CEAS*. Retrieved from http://131.107.65.14/pubs/69181/personaltopics.pdf

Voida, S., & Mynatt, E. D. (2009). It Feels Better Than Filing: Everyday Work Experiences in an Activity-based Computing System. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 259–268). New York, NY, USA: ACM. http://doi.org/10.1145/1518701.1518744

Wang, F., Wang, Z., & Wang, S. (2004). Exploiting Description Knowledge for Keyphrase Extraction. Presented at the PRICAI.

Wang, H. J., Moshchuk, A., Gamon, M., Haraty, M., Iqbal, S. T., Brown, E. T., … Dumais, S. T. (2015). The Activity Platform. In G. Candea (Ed.), *15th Workshop on Hot Topics in Operating Systems, HotOS XV, Kartause Ittingen, Switzerland, May 18-20, 2015*. USENIX Association. Retrieved from https://www.usenix.org/conference/hotos15/workshop-program/presentation/wang

Wang, W., & Haake, J. (1997). Supporting user-defined activity spaces. In *Proceedings of the eighth ACM conference on Hypertext* (pp. 112–123). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=267450

Whittaker, S., Jones, Q., Nardi, B., Creech, M., Terveen, L., Isaacs, E., & Hainsworth, J. (2004). ContactMap: Organizing Communication in a Social Desktop. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *11*(4), 445–471. http://doi.org/10.1145/1035575.1035580

Whittaker, S., Matthews, T., Cerruti, J., Badenes, H., & Tang, J. (2011). Am I Wasting My Time Organizing Email?: A Study of Email Refinding. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 3449–3458). New York, NY, USA: ACM. http://doi.org/10.1145/1978942.1979457

Whittaker, S., & Sidner, C. (1996). Email overload: exploring personal information management of email. In *Proceedings of the SIGCHI conference on Human factors in computing systems: common ground* (pp. 276–283).

Appendix A questionnaire

1. Using activity suggestions made it easier to create activities.

   (Strongly Agree…Strongly Disagree)

   Comments…

2. Using activity suggestions made it easier to populate new activities with artifacts such as emails.

   (Strongly Agree…Strongly Disagree)

   Comments…

3. I prefer to see the suggested activities before creating activities on my own.

   (Strongly Agree…Strongly Disagree)

   Comments….

4. The keywords and the contacts in the suggested activities reminded me of some activities that I would have otherwise not remembered.

   (Strongly Agree…Strongly Disagree)

   Comments….

5. Overall, I found the activity suggestions helpful.

   (Strongly Agree…Strongly Disagree)

   Comments….