Distributed Sensemaking: Improving Sensemaking by Leveraging the Efforts of Previous Users

Kristie Fisher[1,2], Scott Counts[1], Aniket Kittur[3]

[1]Microsoft, [2]Microsoft Research
1 Microsoft Way
Redmond, WA 98052 USA
{kfisher, counts}@microsoft.com

[3]Carnegie Mellon University
5000 Forbes Ave
Pittsburgh, PA 15213
nkittur@cs.cmu.edu

ABSTRACT

We examine the possibility of *distributed* sensemaking: improving a user's sensemaking by leveraging previous users' work without those users directly collaborating or even knowing one another. We asked users to engage in sensemaking by organizing and annotating web search results into "knowledge maps," either with or without previous users' maps to work from. We also recorded gaze patterns as users examined others' knowledge maps. Our findings show the conditions under which distributed sensemaking can improve sensemaking quality; that a user's sensemaking process is readily apparent to a subsequent user via a knowledge map; and that the organization of content was more useful to subsequent users than the content itself, especially when those users had differing goals. We discuss the role distributed sensemaking can play in schema induction by helping users make a mental model of an information space and make recommendations for new tool and system development.

Author Keywords

Sensemaking, search, collaboration, knowledge mapping.

ACM Classification Keywords

H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous;

INTRODUCTION

A large majority of people engage in information seeking behaviors online: 81% of online adults research products; 68% make travel reservations; 89% use search engines; 75% get health information; 71% buy something online[1]. Beyond simply seeking information, people engage in *sensemaking:* constructing a mental representation of interrelated pieces of information relevant to accomplishing a goal, such as planning a vacation, planting a garden, or understanding a current news event [26, 30].

In most cases, after each sensemaking episode in which an individual develops a useful mental representation of an information space for herself, her work is essentially lost, benefiting no one else[2]. Furthermore, if asked today, she may have forgotten much of what she spent so much time learning in the first place. With users spending at least a third of their time online finding and gathering information [15], usefully capturing such efforts represents a significant opportunity. In this paper we explore this possibility by examining the viability of *distributed* sensemaking in the context of web search. We use the term distributed sensemaking here to refer to an iterative process in which users save and organize their own sensemaking efforts, which are then available to subsequent users with whom they are neither collaborating nor communicating, and may not even know.

At its best, distributed sensemaking could help convey the information, evidence, judgments, relations and perspectives that individuals process during sensemaking, increasing the depth and/or speed for each subsequent information seeker. On the other hand, the costs of integrating information from others may be prohibitive, or may lead to biases or gaps otherwise avoided, especially if the goals or expertise of the initial and subsequent users differ. The focus of this paper is on addressing the questions of whether, and in what ways, this type of sensemaking actually helps users.

Potential Advantages

Potential benefits of distributed sensemaking include: 1) saving the user time and/or effort during sensemaking, 2) higher quality sensemaking within a given amount of time, and 3) compensating for a user's lack of knowledge on a given sensemaking topic.

These benefits could be accomplished in two ways. First, having access to the output of other users' sensemaking efforts could lead to more content and/or more relevant content. Second, the user could leverage the way another

---

[2] Reviews do incorporate some of this, but only for the things people actually buy, and also is highly biased towards those with either very positive or negative experiences [12].

user has *organized* content for a particular topic. That is, even if the content itself is not ideally tailored to the user, perhaps the content categories or the notes the previous user included could help the user to abstract a "schema" or mental model of the information space she is trying to make sense of. Studies from cognitive psychology suggest that people form such models through the induction of relational schemas which explicate the relationships between items [8, 13]. These theories suggest an important distinction between the structure of the content and the content itself, with structure being especially important for learning and generalization. For example, novice problem solvers are likely to focus on surface similarities between problems, while experts focus on deep structural similarities [4].

Thus helping users gain expertise in an area may require not only providing them with appropriate content, but also supporting their development of an organizational schema or mental model. This schema could facilitate the user's search for content that is relevant by helping prioritize and focus information gathering. In the context of distributed sensemaking, users would leverage schemas created, iterated, and passed on by previous users, generating a "virtuous cycle" of increasing ease and quality of sensemaking.

Potential Disadvantages
Distributed sensemaking might fail to facilitate, or even hinder, sensemaking. First, for a given sensemaking topic two users are likely to have different preferences, different levels of expertise, unique needs, and so on. This could render the content and its organization irrelevant or confusing to subsequent users. Starting from the output of someone else's sensemaking would make the task *more* effortful and more time-consuming for the user, because she would first have to make sense of the previous person's efforts before engaging in her own sensemaking. This raises a number of questions regarding the cost of distributed sensemaking. How quickly and accurately can a user parse the output of someone else's sensemaking? How does the overhead of figuring out that user's sensemaking process balance again the benefit of leveraging their effort? As subsequent users continue to iterate a sensemaking task, how is this cost-benefit ratio changed?

RELATED WORK
Understanding and improving sensemaking has become a prominent topic of HCI research. Computational models have been developed that characterize how people seek out and determine the relevance of information they find on the web (i.e., SNIF-ACT & ACT-IF; see [23] for a review). As described below, several prototype systems have also been developed to assist individual users with sensemaking by providing an organizational space for them to save and annotate information and by incorporating machine learning algorithms to help with information classification.

Tools for Sensemaking
An early example of such a tool was SenseMaker [2], which emphasized the need for interfaces that could support the organization and exploration of diverse sources of information at different "levels of granularity." More recently, SSIGS [25] allowed a user to create a visual representation of online information in the form of a "knowledge tree" in a web browser side-bar. ScratchPad [11] allowed users to take and arrange "snapshots" of web content akin to visual bookmarks. Entity Workspace [3] was created to support sensemaking in the medical profession by integrating search engine capabilities with document reading and note taking.

Collaborative Search & Sensemaking
Research and tool development in *collaborative* sensemaking aims to support groups of two or more people explicitly working together on web-based sensemaking tasks. Tools such as SearchTogether help users better accomplish goals such as shopping and travel planning [18]. Social interactions and collaboration can also shape an individual's internet search behaviors even when the users are not engaging in synchronous, joint web searching [6].

This collaborative or social component of search and sensemaking is likely to become even more relevant in the future with the explosion in popularity of online social networking, and the availability of the web on a growing number of hand-held devices. Accordingly, several efforts have been made to develop tools and systems to support collaborative search and sensemaking, both when people are working together in time and space [1, 19, 21], and when they are collaborating remotely [14].

Distributed and Iterated Information Organization
The previous work on collaborative search and sensemaking typically has supported groups of users explicitly collaborating towards a common goal, at least at some level [6, 14, 22]. The group of internet users at large is a vastly bigger set of people to draw on for sensemaking assistance. Can individuals benefit from the sensemaking of previous users, even those they don't know? Or, conversely, would exposure to someone else's efforts induce more cost than benefit? Systems like Wikipedia and delicious provide proofs of concept that distributed effort can generate high quality content [9] and useful information organizations [10]. These systems often involve coordination among contributors, as in Wikipedia, or are less task-focused, in the case of social bookmarking. Research on Mechanical Turk [17] shows that iterated effort by a group of people unknown to one another can solve fairly difficult tasks such as decoding text in a blurred image. Together, the systems and research around Wikipedia, social bookmarking, and crowd sourcing suggest that users engaged in sensemaking tasks on the web may benefit from the sensemaking efforts of other users previously engaged in a similar sensemaking task.

RESEARCH STUDY

To study the iterative process of distributed sensemaking, we needed a common artifact that participants could jointly author. For this we turned to digital knowledge maps. Digital knowledge maps are one way to create an "external representation" of sensemaking, and having such a representation has been recommended by several other researchers [14, 28, 29]. We will use the term "knowledge map" to refer to a visual representation of the output of a sensemaking task. That is, knowledge maps consist of information items and a representation of how those items are related.

We observed the process and outcome of sensemaking when users created knowledge maps with and without previous users' knowledge maps to build from. We also asked users to complete detailed surveys about their experience, and analyzed looking patterns as users examined knowledge maps created by other users. Together these measures bear on the "value proposition" of distributed sensemaking: that the process and outcome of sensemaking for any given user is improved by leveraging the efforts of previous users.

Knowledge Map Pilot

We first conducted a pilot study to test the viability and usefulness of this approach using a low-fidelity prototype. A low-fidelity approach helps to ensure that the results would not be constrained by existing technological factors, such as how specific tools influence the ease of annotation or organization of information. It also minimizes training and more closely captures participants' thought processes, though it has drawbacks such as making it difficult to quantify and compare participants' work and to scale up to a larger sample. As such our goal with this pilot was to determine the kind of functionality a tool would need to support for a larger study on distributed sensemaking, and to ascertain early on the likely viability of such a study.

In the pilot study, participants were asked to engage in the sensemaking task of selecting an appropriate venue for a party from a variety of local venues. To generate the source materials, we collected and printed out approximately 200 pieces of evidence from the web, including reviews, menus, maps, photos, and general information, which were pasted onto card stock. We then provided participants with a large whiteboard placed horizontally on a table, annotation tools such as post-it notes and markers, and a folder with all of the evidence. Additional participants, who were unknown to the initial users, were brought in for each initial map, and completed the same task but starting with the end state of the first participant's workspace. During the experiment we employed a think-aloud protocol (shown to provide information about participants' working memory contents without affecting them; [5, 20]), and also captured images of their workspace every few minutes. Results of the pilot were promising, as subsequent participants reported finding the maps and annotations of the first participants useful. One example of the iterative construction process between participants is shown in Figure 1. All participants used structured spatial layouts, often constructing grids with competitors on one axis and features to compare them on another axis. Participants also added annotations, using sticky notes on specific evidence or writing comments on the whiteboard relating to dimensions or competitors. Some interesting emergent themes are discussed below:

Provenance. We were initially concerned that users would either find the organization and annotations of others useless and simply start over, or would base their decisions on the results of the previous participant without examining the evidence for themselves, but neither turned out to be true. Although most participants used others' organizational schemas, all also explored provenance data: *why* someone made an annotation, such as adding stickies indicating items as particularly helpful or organizing items in certain ways. This highlights the importance of enabling users to make informed risk/reward decisions about using the work of others (for example by enabling users to directly view underlying evidence). Showing such provenance information has been shown to have significant impact on trust in other user-generated content systems [7, 16, 24, 27].

Collaborative dimension discovery. Participants learned not only about the features of venues that others found important but also about the dimensions that others found useful for evaluating those features. For example, one participant created a column for the dimension "type of music" which was later found to be useful and adopted by future participants who had not initially thought of it.

Externalization of prior knowledge. We also found evidence that capturing participants' prior knowledge could be extremely helpful. For example, one participant who was native to the area annotated maps with sticky notes that indicated whether or not the venue was on the more desirable side of a river in town. This information was important because while the farther side of the river was much less desirable than the closer side, it appeared essentially equivalent in terms of distance on the map.

Overall, the results of the pilot study were promising in suggesting that spatial landscapes of information developed by one person could be useful for others and that simple tools supporting spatial layouts could be used to further investigate distributed sensemaking. We also found evidence that people not only found the content of others' maps useful, but potentially more useful were the organization and relations between competitors and dimensions. This suggests that "knowledge maps" could provide an important function in promoting the development of rich mental schemas. We further examine these issues through a larger study using digital knowledge maps.

User Study

Building from the results of the pilot study, a more in-depth user study was conducted in order to better quantify the

	This participant (P2) started with the organization and annotations of the previous participant (P1). P1 had organized the information with the rows corresponding to the top four venues, with summaries (orange text) such as "good food" or "good snacks, good service, decent music", but with idiosyncratic columns.
	P2 in the process of reorganizing the information landscape according to his own mental model. The landscape resulted in clearly structured columns, differentiating benefits for each venue listed on the left, and annotations on post-it notes for some of the evidence.
	After P2 was finished, P3 was brought in to do the same task. This figure shows the final results of P3's landscape, which has changed very little from P2's landscape, suggesting the utility of P2's landscape (and confirmed by post-experiment questioning).

Figure 1. Pictures from the pilot experiment. Participants iteratively constructed a "knowledge map" for the sensemaking task of selecting a venue for a party.

effects of distributed sensemaking. The user study consisted of two phases. The primary purpose of the first phase was to assemble a collection of knowledge maps that would be used by people in the second phase. In the second phase, we analyzed users' sensemaking output under different conditions of distributed sensemaking and we tracked their eye movements while they examined other users' knowledge maps.

Phase 1: Knowledge Map Creation & Iteration
To create knowledge maps in this phase of the study, participants were first asked to use a search engine to find resources on the web that were related to one of the following six sensemaking topics:

1. Planning a multi-day trip to a local national park.
2. Exploring options for starting a vegetable garden at your home.
3. Planning a small gala with food and entertainment for your work team (approx. 30 people).
4. Entertaining friends from out-of-town for the weekend with tourist activities, etc.
5. Exploring options for attending a local professional sporting event.
6. Finding resources for do-it-yourself kitchen remodeling.

After finding useful web content, participants used a screen capture tool to capture images of relevant search results, including image, video and any other type of search result. Thus the items making up our sensemaking knowledge maps were snips of web search results. Participants pasted the screen capture images onto a PowerPoint slide and organized them in whatever way most the "made sense" to them.

This method enabled us to build on our pilot by incorporating a greater variety of tasks, as well as by using a digital environment that supported flexibility for organizing and annotating content. Starting with a blank PowerPoint slide meant that participants were able to organize information in a way that was largely free of system constraints. PowerPoint allowed participants to spatially position items in any 2-dimensional layout, size items, connect them with lines and arrows, add labels and notes, color-code, and so on. Thus this approach provided a low-effort way for participants to externalize their mental model of how relevant pieces of information were related to one another and to the sensemaking goal. As subsequent participants iterated on these knowledge maps, they were able to modify them with relative ease.

Twelve participants from our company (2 female), with an average age of 35.58 (*SD* = 8.16) were each asked to create knowledge maps using the method described above. Participants were recruited via internal e-mail solicitation, and they participated in the study by performing the task at their own workstation and then e-mailing us their resulting knowledge maps.

Three participants each created two knowledge maps, with the result being a set of six knowledge maps, one for each of the six sensemaking topics. Participants were asked to keep track of how much time they spent creating each map, taking no more than 20 minutes. They also completed a survey about the knowledge map they created and their experience in the task. These maps were then distributed to three subsequent users who were told specifically to start with the provided map and add, delete, change, or move as much or as little of the content from the previous user as they desired.

This process was then repeated for a second iteration group and then a final iteration group, resulting in a knowledge map for each topic that had been created with the input of four users. The knowledge maps participants created in each round of iteration were evaluated and compared to the initial maps they were given in order to track what had and had not changed during each round.

One notable limitation is that the search results snips used to build knowledge maps were not live, which severely impairs the ability of a subsequent user to evaluate and use them. To reduce this problem, we told users to only include search result snips and not snips of pieces of a web page, so that subsequent participants could see both a preview of the content of the website visited by the other user, as well as the web address in case they wanted to visit the site themselves.

Phase 2: Distributed Sensemaking & Eye-tracking
The next phase of the study was the primary investigation. 21 participants (5 female), all members of our company and with an average age of 37.19 (*SD* = 10.06) were recruited via internal e-mail.

Value of Iteration for Sensemaking Participants first created knowledge maps for three different topics, one in each of three conditions: 1) Starting from a "blank slate" with no access to another user's map (Solo Condition), 2) Building from one of the maps created in Phase 1 by a single other user (Other Condition), and 3) Building from one of the maps that had been iterated in Phase 1 by four other users (Iterated Condition). This allowed us to test for benefits of previous sensemaking: how did the input of a single user or of multiple previous users impact sensemaking in comparison to sensemaking "from scratch"?

The order in which they completed the three tasks was counter-balanced across participants. When completing the task in the Other and Iterated conditions the participants were told simply that the knowledge maps had been "created by one other user," and "had been created with the input of several users over time," respectively. In each case, they were told to work from the knowledge map they had been given, changing as little or as much as they would like in order to create their own knowledge map for the given topic. Participants were given a time limit of 20 minutes to complete each task, but most did not choose to use all of the allotted time. Participants completed a detailed survey to evaluate their experience in each of the three conditions, and the knowledge maps they created were evaluated and compared to the knowledge maps that they had been given.

Parsing the Sensemaking of Others Participants were then shown two additional knowledge maps, each on a different topic. They were not told which sensemaking topic had been given to the user who created the map, and they were not told that one map had only been created by one user and that the other map had been iterated on by several users. Participants' task was to examine the map and try to figure out what it was about in order to answer some questions about it. They were told to look at the map for as much or as little time as necessary. While participants examined each map, we recorded their eye movements. To do this we used a Tobii X50 eye tracker and the accompanying Tobii Studio software. The Tobii system records eye position at 40ms intervals and defines a fixation as he maintaining of eye position for 100ms. For our analyses, total fixation time was summed for pre-defined regions of the screen containing different types of content. This process allowed us to examine how much total time participants needed to accurately assess what the knowledge map was about and to determine what aspects of the knowledge map participants paid the most attention to when trying to understand it. Recall that the effort required to make sense of another person's map represents a potential downside of distributed sensemaking. Comparing this process across maps made by a single other person versus those that were iterated provided a more nuanced picture, with the possibility that iteration mitigates the risk of the organization of any single other user being difficult to understand.

Knowledge maps and sensemaking topics were counter-balanced across participants. At the end of the study, we conducted informal interviews with each participant, and they were permitted to give additional open-ended feedback about the viability of a distributed sensemaking system and their ideas and preferences for specific tools and features.

RESULTS

Properties of Created & Iterated Knowledge Maps
The knowledge maps created by participants in the first phase of the study had some interesting consistencies, most notably in how they were organized. The most popular way of organizing content was to create small groupings of content and give a text label to each group (83% of all maps created across all four rounds of creation and iteration), as shown in Figure 2. Furthermore, the way a map was

organized tended to be quite stable over the rounds of iteration. For the first round of iteration, four out of the six topic maps maintained their original organizational structure. The only two maps whose organizational structure was changed started out as being a disorganized "collage" or a hybrid structure (groups and networks) and ended up being transformed into a group and label structure. In the second and third rounds of iteration, five out of the six maps maintained their group structure. We find this result surprising given that participants were free to choose any 2-dimensional organizational structure. Also, in sensemaking tools developed by previous researchers [11, 26], a network or hierarchy was the only type of organizational structure available to users, but neither of those was spontaneously used by our participants.

The typical look and content of the knowledge maps were as follows. The maps on average, across creation and iterations, contained about 13 search result snips ($M =$ 12.67, $SD = 5.00$) with a range from six to 25 items. When these items were organized into groups, there tended to be between four and five on average ($M = 4.67$, $SD = 2.73$). When iterating on others' knowledge maps, participants added about six new items on average ($M = 6.04$, $SD = 5.62$), with a range of zero to 24 items. Similarly, participants deleted about six items on average ($M = 5.88$, $SD = 5.62$), with a range of zero to 17. Participants also tended to add and delete organizational groups of items with an equivalent rate (add: $M = 2.17$ groups, $SD = 2.22$; delete: $M = 1.88$ groups, $SD = 2.33$).

Potential Advantages & Disadvantages

Descriptive statistics for all of the analyses described below are in Table 1 (survey data) and Table 2 (eye tracking data). Unless stated otherwise, for all comparisons described below, ANOVAs with planned comparisons (with Bonferroni correction) were used to compare means when
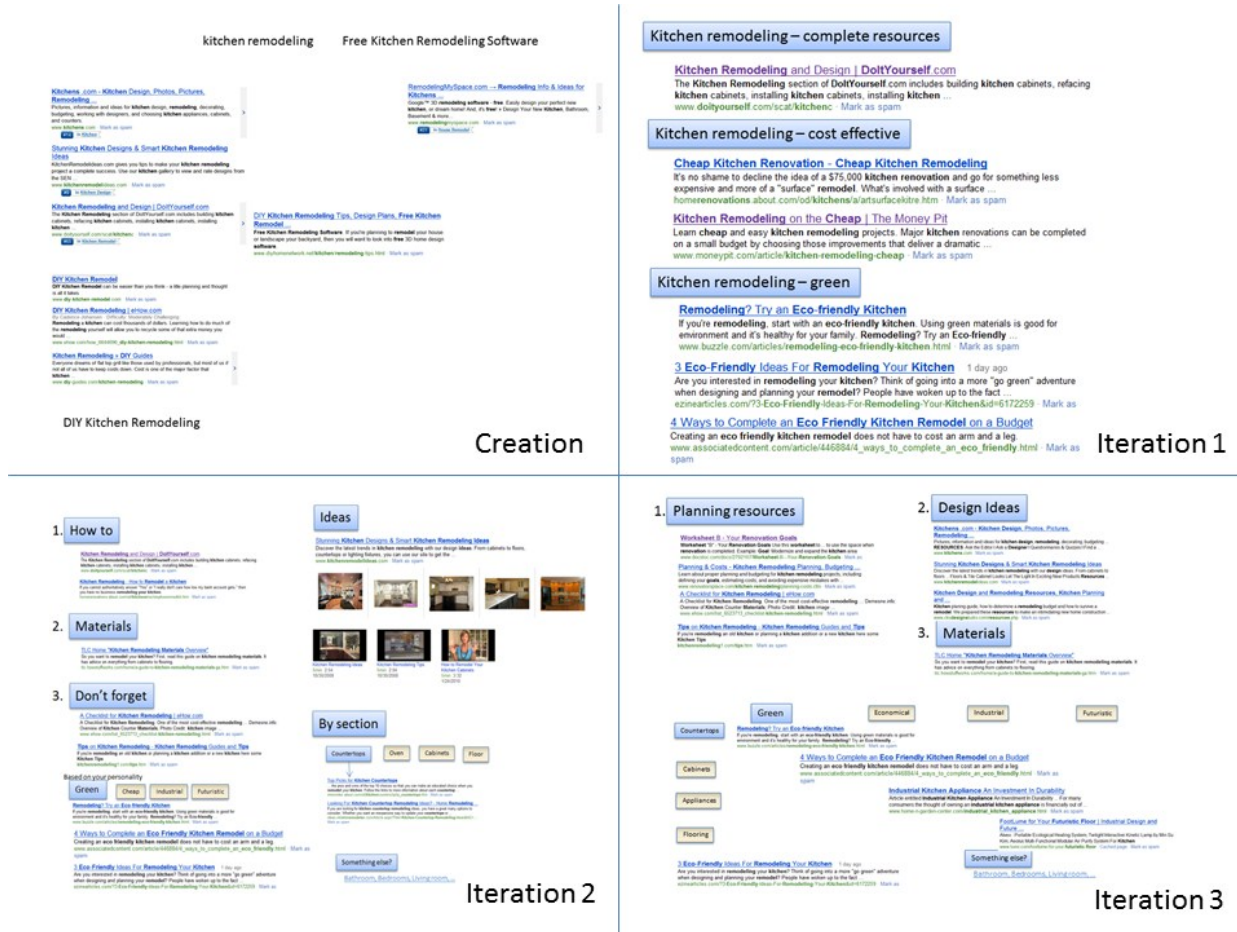

Figure 2. Example of a "knowledge map" that was created and then iterated on over time by three subsequent users. The topic of the map was *Exploring options for "do it yourself" kitchen remodeling*. The first user found some relevant resources and made an attempt to start organizing them into groups, but the second user was able to greatly improve on the organization. The third user expanded on the organizational groupings, adding a numbered step-by-step process for how to sort through the resources. The final user changed little of the organizational structure, and tried to synthesize it. The fourth person also started building an arrangement of rows and columns based on labels from the third user in order to provide a resource for many different combinations of options reflecting different needs (e.g., "cost effective," "industrial"). Note: for space reasons, we use this example to highlight schema emergence through iteration, with smaller text not meant to be readable.

Map Creation Task (Distributed Sensemaking)						
	Solo		Other		Iterated	
	M	*SD*	*M*	*SD*	*M*	*SD*
Time to Create Map (minutes)	15.52	3.49	14.33	3.95	13.67	4.52
Rating of Cognitive Effort for Creation	4.14	1.42	3.71	1.45	3.24	1.37
Self-rated Map Quality	4.90	1.26	4.29	1.19	5.48	1.12
Helpfulness of Other Users' Maps	-	-	3.57	1.86	4.86	1.85
Self-rated Helpfulness to Others	4.90	1.26	4.74	0.86	5.62	0.97

Map Examination (Eye-tracking)				
	Other		Iterated	
	M	*SD*	*M*	*SD*
Map Clarity	4.19	2.09	6.24	0.89
Understanding Author Intent	4.67	1.46	5.95	0.67
Likelihood of Creating a Similar Map Organization	2.95	1.75	5.19	1.17
Likelihood of Creating a Map with Same Content	4.76	1.00	5.67	1.02

Table 1. Descriptive statistics for survey and creation time data for the two tasks in Phase 2 of the study. For all survey items, participants used 1-7 rating scales, with 1 representing the negative valence (e.g., "High effort," "Very unclear") and 7 representing the positive valence (e.g., "Low Effort," "Very clear").

there were three conditions of interest for a measure, and dependent t-tests were used when there were two conditions. Note that the results do not change when data are analyzed with non-parametric tests.

Time and Effort
Turning to results of the second phase of the study, people spent the most time creating a knowledge map in the Solo Condition and the least time creating a knowledge map in the Iterated Condition. However, the difference among these three conditions was not statistically significant, $F(2, 40) = 2.10$, $p = .14$. A similar pattern was observed for users' ratings of the "cognitive effort" that creating a knowledge map required in each condition. On a scale of 1 (very low effort) to 7 (lots of effort), the average rating for the Solo Condition suggests it was neither very easy nor very difficult. Effort in the Other and Iterated conditions were rated as being somewhat less effortful, but the ratings in these three conditions were not statistically different from each other, $F(2, 40) = 2.20$, $p = .12$. A direct comparison between the Solo and Iterated conditions was borderline significant, $t(20) = 2.38$, $p = .054$. Thus, while the effects are in the direction of distributed sensemaking being beneficial, we cannot draw conclusions about time and effort savings, with the exception of a trend-level reduction in effort when compared to starting from scratch. We note that particularly for time spent, the variance was high and therefore these effects may only be teased out with considerably greater numbers of subjects.

Quality and Helpfulness
Though time spent was similar across conditions, what users were able to accomplish within that period of time appears to be of higher quality and greater use when working with the iterated knowledge maps. Using a scale of 1 (low quality) to 7 (high quality), participants rated the quality of the maps they created the highest in the Iterated

Condition, followed by the Solo Condition, and lastly the Other Condition. These quality ratings were significantly different from one another, $F(2, 40) = 5.68$, $p = .007$, and paired contrasts with a Bonferroni correction revealed a significant difference between Other and Iterated, $p = .005$. Participants felt *their own* sensemaking superior when starting with content organized and iterated on by multiple previous participants.

Participants also rated the knowledge maps of others in the Iterated Condition as being significantly more helpful than those in the Other Condition, $t(20) = 2.38$, $p = .027$. Furthermore, participants were asked to rate how helpful they thought the knowledge map would be to another user if that user was asked to work from the knowledge map in a sensemaking task for the same topic. Again, participants thought their map would be most helpful to others when they had constructed it in the Iterated Condition, followed by the Solo Condition, and lastly the Other Condition.

Value of Organization
We asked participants to describe *why* they rated the knowledge maps given to them in the Other and Iterated conditions as helpful or not helpful. In the Other condition, four of the participants who rated the map helpful (i.e., greater than 4 on the rating scale) mentioned the quality of the organization as the primary reason for their rating, while only one person mentioned the content. For the Iterated Condition, nine people who rated the map as helpful cited organization of the map versus three due to content. Most of the reasons cited when the map was rated un-helpful were due to the content the map contained (5 people) rather than the organization of the map (1 person). Some quotes from this free-form section of the survey that highlight the importance of well-organized content over the content itself are listed below:

- *Although I deleted 50% of what was on the other knowledge map, it gave me some ideas right away on where to look and which factors or questions I should consider during my search.*

- *Helpful only because there were groupings of ideas that I used to form my ideas for what would be useful.*

- *I liked the structure - it was helpful things to see the thought process laid out.*

Understanding the Sensemaking of Others

To further probe the idea that structure was most helpful to subsequent users, we turn to results of the eye-tracking portion of the study. Recall that we tracked eye movements of participants as they read knowledge maps created by previous participants (see Figure 3). The purpose of this task was to understand how, and how well, participants could make sense of other people's sensemaking activities. We focus on the comparison between working with knowledge maps generated by a single other person versus those iterated by four people.

Participants looked for a roughly equal amount of time at both the Other knowledge map and the Iterated knowledge map; the difference was not significantly different, $t(20) = 1.03$, $p = .32$. They were also highly accurate in both conditions at discerning the topic of the knowledge map, though slightly more so in the Iterated condition (Other: 81% correct, 5% close, 14% incorrect; Iterated: 86% correct, 14% close); the difference in performance proportions was in the expected direction, but not significant, as revealed by a chi-square homogeneity of proportions test, $X^2(2, N = 21) = 0.82$, $p = .66$.

Diving in deeper, we examined looking time across different aspects of the knowledge maps. Specifically we compared time spent looking at content (the *links* in the search results, and content *boxes* containing clusters of results, such as groups of images) versus organizational elements like *labels*. Looking times for links, labels ($M = 6.52$, $SD = 6.33$), and boxes ($M = 6.80$, $SD = 5.57$) are not significantly different from one another for Iterated maps, $F(2, 36) = .44$, $p = .65$. However, there is a significant difference among looking times for links, labels, and boxes for the knowledge maps in the Other condition, with links looked at considerably longer than organizational elements, $F(2, 26) = 6.38$, $p = .006$. From this we infer that iteration allows a schema to emerge, reducing visual focus on content elements (links).

After viewing each knowledge map, participants completed a survey. Participants were asked whether it was clear how the items in the map were related to one another (rated from 1 – not clear, to 7 – very clear), and they rated the Iterated maps as being significantly more clear than the Other maps, $t(20) = 4.22$, $p < .001$. Participants also said that they could more easily understand what the author was thinking in the

Iterated maps compared to the Other maps, $t(20) = 3.29$, $p = .004$.

Participants were also asked to rate how likely they would be to create a similar map if they were given the same topic (1 – not at all likely to 7 – very likely). Participants said that they would be much more likely to *organize* content in the same way that the Iterated maps were organized relative to the Other maps, $t(20) = 4.03$, $p = .001$. They also said that they would be more likely to include the same content as the Iterated map compared to the Other map, $t(20) = 3.10$, $p = .006$, though this difference was a bit smaller than that for organization.

DISCUSSION

We examined the process of distributed sensemaking in which participants iteratively engaged in a sensemaking task using the results of others' work. Overall we found statistically significant increases in self-rated quality and helpfulness, and a modest reduction in effort, when using iterated maps with no corresponding increases in time spent. Our eye tracking data, rating data, and other qualitative feedback suggest that these gains were due to the accumulation of an organizational framework, or schema, across users and over the rounds of iteration. While the usefulness of the content often changed from person to person, the structure of that content remained consistent and useful. For example, two people looking to start a garden may live in different cities, or one in the country and the other in the city, so specific content such as which seeds to plant will change, but they both benefit from the content dimensions of "design ideas", "how to", and so on. That this structure did not start to emerge until a map had been iterated once (see Figure 2) speaks to the value of building on others' work: this organizational structure is not necessarily something a person is going to think of in an initial 15 minutes of sensemaking.

Our survey results also indicated that many users would simply prefer to start from scratch rather than with the initial content collection of a single other user. A fruitful area for future research may be exploring how systems for distributed sensemaking can encourage users to get over the hump of the initial round of iteration. However, despite this "first iteration hump", it did not appear difficult for users to parse the meaning of another person's knowledge map, even when the content and/or organization of this knowledge map were less than ideal. Specifically, we found in the case of knowledge maps created by just a single user,

	Other		Iterated	
	M	*SD*	*M*	*SD*
Total	30.26	17.10	27.02	12.56
Links	15.07	12.27	8.36	6.55
Labels	5.95	3.18	6.52	6.33
Boxes	5.95	5.91	6.80	5.57

Table 2. Descriptive statistics for looking time data (in secs) from the Map Examination (eye-tracking) task.

participants could make an accurate assessment of that user's information goal in a short amount of time.

Importantly, these findings provide evidence that people do not have to collaborate directly in order to help one another make sense of information, nor even know each other. If one user can pass along the "structure" or "schema" of their thought process with respect to a topic—even by simply providing some meaningful sub-topics helpful in organizing relevant information—it can facilitate the sensemaking process of a subsequent user. Our eye tracking results also showed that with further iteration on knowledge maps, users shift to using a greater balance of structural dimensions versus content dimensions. As such they provide an interesting relation to the Chi et al. [4] findings discussed earlier in which the transition from novice to expert was accompanied by a shift in focus from surface to structural features. In the case of distributed sensemaking, the expertise can be thought of as embodied in the artifact itself, which develops deeper and more useful structural organization with further iteration.

The impact and generalizability of our results should be considered in light of the limitations inherent in our study design. First, the knowledge map creation procedure constrained behavior to some extent (e.g., that people had to use a single power point slide). Second, our participant sample could arguably be considered "experts" in sensemaking with digital information, as they were employees at a technology company (though again we point to the diversity in job descriptions). Finally, and perhaps most importantly, we were only able to put knowledge maps through a few stages of iteration. The effects of distributed sensemaking over tens to hundreds of iterations are left to future research.

In summary, we find both promise and challenge in the distributed sensemaking approach. Iteration on distributed knowledge maps can lead to greater improvements in quality and usefulness in a given amount of time than when starting from scratch, even when individuals are working for themselves and not for others. These benefits occurred with only four rounds of iteration and a simple interface not optimized for knowledge maps; exploring new interfaces for constructing knowledge maps and greater iterations are important future work. Furthermore, benefits due to greater structure and organization of the maps transferred across individuals even when the content of the maps did not, demonstrating the value of the approach across differences in contexts, goals, and expertise.

However, we also found challenges in first iteration adoption of knowledge maps, with many users preferring to start from scratch than to start with a map that has not been iterated on. Importantly, the system must encourage users to iterate beyond the initial creation, as significant improvements in quality and helpfulness were seen in the fully iterated knowledge maps. Approaches to doing so include both machine methods, such as automatic alignment of first-round knowledge maps to produce maps that look more like iterated maps; as well as human methods, such as leveraging paid crowds to bootstrap the system or enforcing constraints such as requiring users to integrate some first-round maps in order to benefit from iterated ones.

CONCLUSION

In this paper we examined the idea of distributed sensemaking, in which users who are unknown to one another iteratively engage in a web search-based sensemaking task. This form of sensemaking could harness effort that today is lost when users work in isolation. We
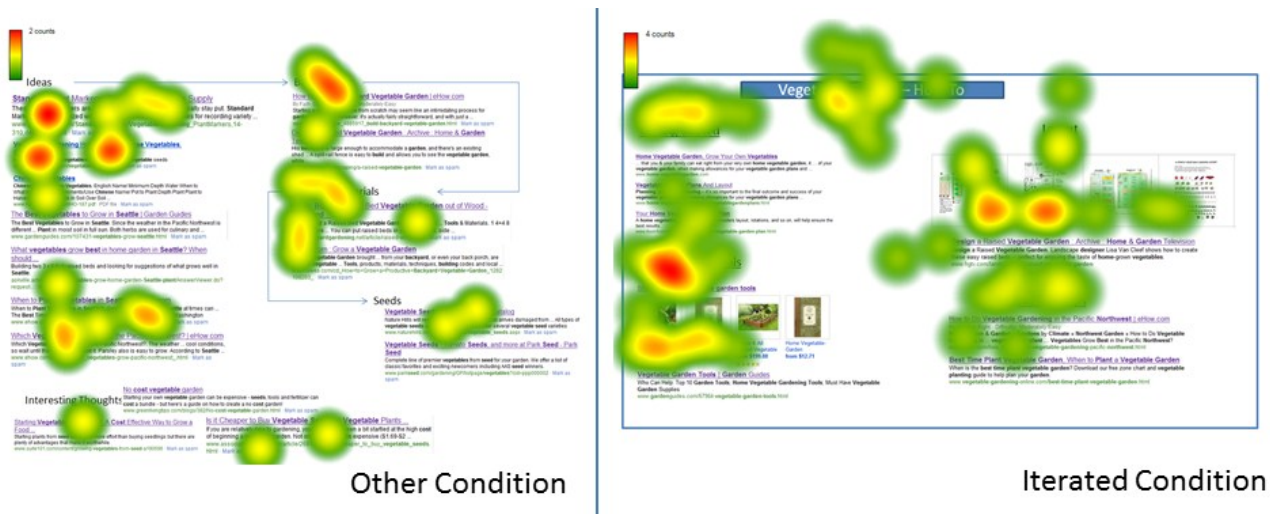

Figure 3. Example eye-tracking results for the topic, *Exploring options for starting a vegetable garden at your home*. Across all topics, in the Other condition (map created by one other user), participants spent more time looking at the search result content relative to the organizational labels. Here the labels "Ideas" and "Interesting Thoughts" receive little to no visual attention. In the Iterated condition (map created by one user and iterated on by three subsequent users), participants looked for equal amounts of time at the different types of content in the knowledge map. Here the labels given to groups of content items receive heavy visual attention.

showed that iterating knowledge maps over only four users allowed a collectively-generated schema to emerge, leading to significant improvements in sensemaking quality and helpfulness. We also identify a key startup obstacle in the distributed sensemaking process; users preferring to use maps that have been iterated on multiple times versus starting over, but prefer to start from scratch when given a map iterated on only once. Overcoming this startup cost will be a key factor in realizing the potential of distributed sensemaking.

REFERENCES
1. Amershi, S., & Morris, M.R. CoSearch: A system for co-located collaborative Web search. In *Proc. CHI, 2008*, 1647-1656.
2. Baldonado, M.Q.W., & Winograd, T. SenseMaker: An information-exploration interface supporting the contextual evolution of a user's interests. *Proc. CHI, 1997*, 11-18.
3. Billman, D., & Bier, E. Medical sensemaking with entity workspace. In *Proc. CHI, 2007*, 229-232.
4. Chi, M.T.H., Feltovich, P.J., and Glaser, R. Categorization and representation of physics problems by experts and novices. *Cognitive science 5*, 2 (1981), 121–152.
5. Davey, B. (1983). Think aloud: Modeling the cognitive processes of reading comprehension. *J. of Reading, 27*(1), 44-47.
6. Evans, B.M., & Chi, E.H. Towards a model of understanding social search. *Proc. CSCW, 2008*, 485-494.
7. Erickson, T., Smith, D., Kellogg, W., Laff, M., Richards, J., & Bradner, E. (1999). Social translucent systems: social proxies, persistent conversation, and the design of "babble". In *Proc. CHI 1999*.
8. Gentner, D. Structure-mapping: A theoretical framework for analogy. *Cognitive science 7*, 2 (1983), 155–170.
9. Giles, G. Internet encyclopaedias go head to head. *Nature, 438* (2005), 900-901.
10. Golder, S. & Huberman, B. (2006). Usage Patterns of Collaborative Tagging Systems. *Journal of Information Science*. 32(2).
11. Gotz, D. The ScratchPad: Sensemaking support for the web. In *Proc. WWW, 2007*, 1329-1330.
12. Hu, N., Pavlou, P., & Zhang, J. (2006). Can Online Reviews Reveal a Product's True Quality? Empirical Findings and Analytical Modeling of Online Word-of-Mouth Communication. In *Proc. EC, 2006*.
13. Hummel, J.E. and Holyoak, K.J. A symbolic-connectionist theory of relational inference and generalization. *Psychological Review 110*, 2 (2003), 220.
14. Keel, P.E. EWall: A visual analytics environment for collaborative sense-making. *Information Visualization, 6* (2007), 48-63.
15. Kellar, M., Watters, C., & Shepherd, M. (2007). A Field Study Characterizing Web-Based Information-Seeking Tasks. *J. of American Society for Information Science, 58*(7), 999-1018.
16. Kittur, A., Chi, E., & Suh, B. (2008). Can You Even Trust a Wiki? Impacting Perceived Trustworthiness in Wikipedia. In *Proc. CSCW 2008*.
17. Little, G., Lydia, B., Chilton, M., & Miller, R. (2010). Exploring Iterative and Parallel Human Computation Processes. In *Proc. HCOMP, 2010*.
18. Morris, M.R. A survey of collaborative Web search practices, In *Proc. CHI 2008*, 1657-1660.
19. Morris, M.R., & Horvitz, E. SearchTogether: An interface for collaborative Web search. In *Proc. UIST 2008*, 3-12.
20. Nielsen, J., Clemmensen, T., & Yssing, C. (2002). Getting access to what goes on in people's head: reflections on the think-aloud technique. In *Proc. CHI 2002*.
21. Paul, S.A., & Morris, M.R. CoSense: Enhancing sensemaking for collaborative web search. In *Proc. CHI, 2009*, 1771-1780.
22. Paul, S.A., & Reddy, M. (2010). Understanding Together: Sensemking in Collaborative Information Seeking. In *Proc. CSCW 2010*.
23. Pirolli, P. Rational analyses of information foraging on the web. *Cognitive Science, 29* (2005), 343-373.
24. Pirolli, P., Wollny, E., & Suh, B. (2009). So you know you're getting the best possible information: A tool that increases Wikipedia credibility. In *Proc. CHI 2009*.
25. Qu, Y. (2003). A sensemaking-supporting information gathering system. In *CHI EA 2003*.
26. Russell, D.M., Stefik, M.J., Pirolli, P., Card, S.K. The cost structure of sensemaking. In *Proc. CHI, 1993*, 269-276.
27. Shneiderman, B. (2000). Designing trust into online experiences. *Communications of the ACM, 43*(12), 57-59.
28. Suthers, D.D. Collaborative knowledge construction through shared representations. In *Proc. HICSS-38, 2005*.
29. Umapathy, K. Requirements to support collaborative sensemaking. In *CSCW CIS Workshop, 2010*.
30. Whitaker, S. Making sense of sensemaking, in *HCI remixed: Reflections on the works that have influenced the HCI community,* MIT Press, Boston, MA, USA, 2008.