Standing on the Schemas of Giants: Socially Augmented Information Foraging

Aniket Kittur, Andrew M. Peters, Abdigani Diriye, and Michael R. Bove
Carnegie Mellon University
5000 Forbes Ave, Pittsburgh, PA 15213
{akittur, andrewpe, adiriye}@cs.cmu.edu, mbove@andrew.cmu.edu

ABSTRACT
People spend an enormous amount of time searching for complex information online; for example, consumers researching new purchases or patients learning about their conditions. As they search, people build up rich mental schemas about their target domains; which, if effectively shared, could accelerate learning for others with similar interests. In this paper we introduce a novel approach for integrating the schemas individuals develop as they gather information online and surfacing them for others with similar interests. Through a controlled experiment we show that having access to others' schemas while foraging for information helps new users to induce more useful, prototypical, and better-structured schemas than gathering information alone.

Author Keywords
Schema; social search; foraging; search; structure.

ACM Classification Keywords
H.5.2 [Information Interfaces and Presentation]

General Terms
Human Factors; Design; Measurement.

INTRODUCTION
People spend an enormous amount of time online searching for information about unfamiliar domains, whether they are patients trying to make sense of their symptoms; consumers deciding which digital camera to buy; scientists learning the literature of an unfamiliar field; or voters trying to understand which issues they should support. The amount of time spent online globally is estimated at 35 billion hours per month (comscoredatamine.com), of which approximately ⅓ is estimated to be spent on such complex information foraging activities [23]. Thus accelerating online knowledge acquisition could have significant and widespread benefits.

As people engage with an unfamiliar domain they learn not only about the *content* of a domain but also its *structure*. Specifically, users build up mental models and rich knowledge representations that capture the structure of a domain in ways that serve their goals [45, 21, 25]. Such schemas change over time as a user moves from being a novice to an expert, with greater expertise enabling schemas that focus on deeper and more meaningful features [6, 25]. For example, a novice photographer may begin their search for a new digital camera focusing on how many megapixels a camera has, but upon learning more about the domain might come to the conclusion that the size of the sensor and low-light sensitivity are more important dimensions.

How can schema acquisition for novices be augmented? One approach is to leverage the schemas that other novices have already induced as they learn about a new domain [12, 47]. Enabling people to build on each others' learning could increase the speed and depth of their sensemaking across a variety of fields [11, 17]. However, such an approach also raises potential challenges. Capturing structured information can be time consuming for users, who may be focused on their own interests rather than on helping others. Even if users were motivated to capture structured information, mismatches with the goals and expertise of those consuming their information could make it too costly to consume and with little benefit. In the limit, if each persons' schema were idiosyncratic, there would be no easy way to aggregate across schemas.

In this paper we begin to explore these challenges by augmenting an online sensemaking tool [25] with features that enable the integration and surfacing of schemas from multiple users during information foraging. Our key contributions include:

- A scalable approach for capturing and aggregating schemas across individuals
- Methods for surfacing the schemas of others during sensemaking
- Empirical evidence of the value of others' schemas during information foraging

APPROACH
How can we capture the schemas that people build up as they make sense of an unfamiliar domain, integrate the schemas of multiple people, and usefully surface those schemas to others? One promising line of work highlights the utility of capturing and iterating on schemas across us-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CSCW'14, February 15–19, 2014, Baltimore, Maryland, USA.
Copyright © 2014 ACM 978-1-4503-2540-0/14/02...$15.00.
http://dx.doi.org/10.1145/2531602.2531644

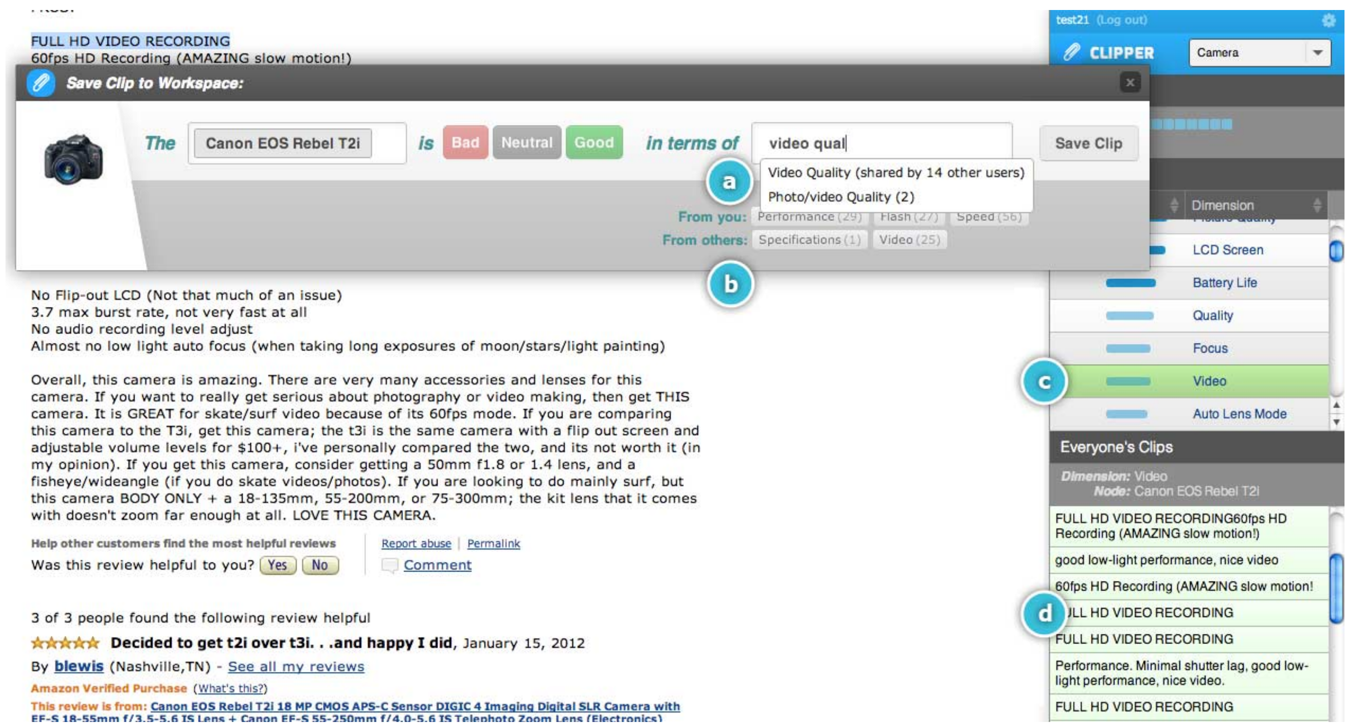

Figure 1. A novel interface to support social sensemaking. Social autocompletion (a) and dimension hints (b) assist users while they create clips. At right, the workspace panel shows information about the domain's dimensions (c) and their corresponding clips (d).

ers. Fisher, Counts & Kittur [12] engaged users in the open-ended creation of "knowledge maps", and found that users often created two-dimensional structured schemas involving options (e.g., *Canon T2i*, *Sony A55*) and dimensions (e.g., *lens, picture quality*) of the information space. Furthermore, they found that having users iteratively build on the maps of others led to improved maps that new users found useful. However, new users preferred to start from scratch rather than use maps created by only one individual. This raises the challenge of getting over the "hump" that raw, individual maps are not found useful, yet generating useful iterated maps requires someone to start from a raw, non-iterated map. Another challenge for this approach is how to aggregate maps across users, as each map in their study had an idiosyncratic spatial layout.

Instead of trying to combine entire maps across users, here we introduce the approach of breaking down schemas into their constituent elements and aggregating those elements. Given that users consistently structured their maps according to options (e.g., Canon Rebel T2i) and dimensions (e.g., picture quality) [12], here we use *dimension* as the element on which to aggregate across users' information spaces. To capture options and dimensions in a lightweight way we build on the sensemaking system described in [25], which elicited structure using an {*item*, *valence*, *dimension*} paradigm framed as a sentence (e.g., "The *Canon T2i* was *good* on *picture quality*"). This approach was shown to be useful and robust in eliciting people's structure of an information space. However, the system in [25] was designed for individual sensemaking, and users foraged for information completely independently from others. In this paper we add *asynchronous social aggregation*: aggregating the dimensions of previous users as potentially useful ways for a new user to structure their information workspace.

On the one hand, such an approach could expose users to more useful and expert dimensions that could augment their schema acquisition. Conversely, it is possible that the dimensions of others might not be valuable and just add to the burden of information to consume. Indeed, [25] found that eliciting dimensions from users during the information foraging process was likely to produce dimensions that are obsolete, irrelevant, novice, or idiosyncratic. For example, the previously mentioned novice photographer might start off creating dimensions such as *megapixels*, but upon learning more about the domain realize that *sensor size* or *lens selection* were more important factors. Aggregating and surfacing obsolete dimensions could be worse than showing none at all. However, if more useful or expert dimensions are more commonly input by users than obsolete or idiosyncratic dimensions, then aggregating dimensions across users could be a way to find valuable dimensions to surface.

We hypothesize that augmenting the approach introduced by [25] so that the captured dimensions could be integrated across individuals and then surfaced to new users might simultaneously address a number of fundamental challenges: 1) the foraging interface could capture and aggregate schemas across individuals without requiring them to interact with raw schemas, 2) aggregating by dimensions provides a scalable and segmentable method for combining multiple users' data; and 3) aggregating schemas across individuals could filter out noisy, obsolete, or idiosyncratic schemas. Below we introduce novel asynchronous social aggregation features to the Clipper interface [25] that enable capturing, aggregating, and surfacing schemas across individuals. We then describe a controlled experiment aimed at understanding the conditions under which socially augmented information foraging is effective.

RELATED WORK

Information Seeking

Users engage in a variety of information exploration tasks online ranging from finding the price of a camera they have in mind to deciding on the model of a carseat to purchase [7, 23, 33]; there is a large body of work that has looked at understanding and supporting how users navigate the web to find information (e.g., [5, 32, 42, 44, 50]). During online information seeking people engage in a process of information foraging and integration [45, 24]. To support these behaviors, a number of tools exist – for example, WebBook and WebForager by Card et al. [4], which use a book metaphor to find, collect, and manage web pages; Elastic Windows, which provided information overview and location context [22]; Webcutter, which collects and presents URL collections in tree, star, and fisheye views [31]; SenseMaker [2] for evolving collections of information; and Scatter/Gather [8] a text-clustering interface for iteratively navigating through document collections. Some of the more recent tools include Hunter Gatherer [46] which supports fine-grain clipping of information from web pages; and Dontcheva et al.'s suite of tools to assist users' browsing, collecting and sharing information found on web pages [10]. Our approach is similar in spirit to Dontcheva's in organizing collected information in a structured way, but instead of having predefined templates and extraction patterns, in our system structure continually emerges from and is refined by the aggregate behavior of many individuals. Furthermore our structure aims to reflect the schemas that underlie the structure of the conceptual information space rather than the elements of the web page.

Social Data

Data created by people using one of the above tools has started to be repurposed for other means. Social annotation and data is used by most systems to either augment the content surfaced to the user or to enrich the content presented. Social annotation and data is derived from user behavior (e.g. selecting a specific search result) or generated content (e.g. tags assigned to a given web page). Social data can benefit users to make sense of information and search and browse more effectively. Social data is becoming more visible; we see social footprints and annotations in the forms of "likes", "shares", +1's and reviews on sites like Facebook, Yelp, and Google+. Several different approaches have been proposed on how to use social data to personalize and re-rank search results, and enrich information sharing [17, 18, 34, 3, 54, 53].

Systems developed by Golder et al. looked at the structure of systems leveraging user annotations and tags [17]; Wu et al. leveraged social data and annotations to build statistical models in search and browsing [52]; Dmitriey et al. used social data to enrich anchor text and improve intranet search [9], and Muralidharan et al. studied the content and presentation that make social annotations useful [39]. Research into social data during information seeking can inform us of what type of information is useful during users' sensemaking.

Social and Collaborative Sensemaking

Other aspects of sensemaking such as the social and collaborative dimensions are seeing more attention. Collaborative sensemaking involves people interacting, sharing, searching and understanding information with each other while collaborating [42], and according to Morris and Horvitz [37], the support provided by existing tools for collaborative sensemaking falls under two categories: awareness features, that is features intended to support the group's status and work product, and division of labour, features intended to support how work is divided up. One notable piece of work is by Sharma et al. [47], who studied how artifacts handed-off and available online affect subsequent sensemaking. They found that resources that were accessed online resulted in more structured sensemaking whereas information that was handed-off to people resulted in less structure added.

We are also seeing work that looks at how people turn to others for recommendations [11, 36] or actively work together to search and filter information for each other [28, 36, 51]. Some work has looked at saving information from the search process so that others can re-use it in the future [1, 37, 41]. Research in personalized web search often uses the implicit traces of others' opinions of behaviors to train collaborative filtering systems to recommend items matching individuals' information seeking needs [15, 16, 19, 27, 49]. Similarly, social navigation uses the traces of others' behaviors to point people to useful places that others have found [38, 13]. An increasing number of social web sites are collecting and aggregating information from many users, including social bookmarking systems such as del.icio.us [29], DogEar [35], and Spartag.us [20]; social news aggregators such as Digg [30]; and many other commercial systems (e.g., diigo, CiteULike, Connotea, ConnectBeam). Our work builds on these approaches but instead of focusing on navigation and suggesting items to look at, we aim to make use of the deeper cognitive processing and myriad judgments that individuals engage in while sensemaking and using them to help others build up rich relational schemas of the information space.

PLATFORM – DESIGN RATIONALE

Overview

We started with the design of the existing Clipper web tool [25], which assists individuals in the gathering and structuring of information in a given domain. Clipper allows the user to save textual information ("clips") from a web page, along with a structured summary of the text. This summary is captured in a well-defined schema, where each contains an *item* (the topic of the clip), a *valence* (whether the sentiment is positive, negative, or neutral) and a *dimension* (the attribute of the item in question). This approach leads to summaries that can be surfaced in sentence form, such that "The *[item]* is *[valence]* in terms of *[dimension]*." We follow the approach of [25] in operationalizing schemas as the *dimensions* that structure the information space, due to the simplicity, tractability, and grounding in empirical evidence of the approach, though we recognize that real mental schemas can be richer and deeper. Thus, for a given domain, a collection of its clips' dimensions begins to define the important and differentiable criteria between items in the space.

We used two distinct versions of the interface that elicit this schema at different points of the information foraging process. The "full" clipper prompts the user to supply node, valence, and dimension when the Clipper is activated and text on the page is highlighted (see Figure 1); the "mini" clipper prompts only for node during clip creation and defers valence and dimension selection until a later review phase. Although we were interested in whether different versions of the interface would lead to differences in how social schema information was used, our results did not show any significant differences in our primary outcome measures; hence, going forward we collapse our data across the two interface conditions.

We then created a social version of the Clipper tool, which integrates and surfaces the schemas of multiple users. In order to provide users with assistance with creating dimensions, we leveraged schemas that had been collected using the Clipper interface from a previous study [25]. These data were comprised of 55 users, 2388 clips, and 560 distinct dimensions. We surface this information in a number of ways in the socially augmented interface:

Autocomplete

Dimensions are suggested to users as they begin to type in the dimension selection field while saving clips. In the social condition, users see autocomplete suggestions from other users' workspaces in addition to their own, along with a measure of dimension overlap (see Figure 2).

Dimension hints

We surface dimension hints to users whenever they are prompted to select dimensions while creating clips (see Figure 3). For each clip, we suggest up to five dimensions, including: the user's last used dimension; the user's most popular dimension; a random dimension from the user's workspace; a random dimension from all social workspaces; and the most overlapping dimension from other users, given the words comprising the clip. Specifically, the latter *social overlap* method involved summing the distribution of dimensions for each word in the clip across all social workspaces, and selecting the dimension with the highest total weight. Ties were broken by the heuristic of selecting the dimension with the higher average clip length. Dimension hints were generated independently, with one "slot" for each of the five methods. When the same dimension was generated for two or more slots it was randomly assigned to one of the slots so that no dimension was shown more than once. In the non-social condition, only the user's random, most popular, and last dimension were shown as hints.

Workspace Pane

The workspace pane is overlaid on the web pages where the user browses to gather information (see Figure 5). It contains a list of all dimensions the user has specified for the domain (marked with a head in the *You* column in Figure 5), and, in the social condition, the most popular dimensions specified by all users with their corresponding clips. The blue bar in the *Others* column (specific to the social condition) corresponds to the number of other users who have saved the dimension. It serves as a running summary of the most popular dimensions; users may refer to it while foraging for clips or hide it if desired.

Review Table

The review table, the final view of clips the user is shown after finishing their foraging (see Figure 4), contains a table of a user's nodes vs. dimensions, with corresponding clips in each cell organized by valence (e.g., good/neutral/bad). In the social condition, this table additionally contains nodes, dimensions, and clips from all users who have previously completed the task, which are shown in a less saturated hue than the users' own. Hovering over a node shows the actual text of the clip.

CORE HYPOTHESES

A core idea behind the interface features described above is that integrating and surfacing dimensions across individuals

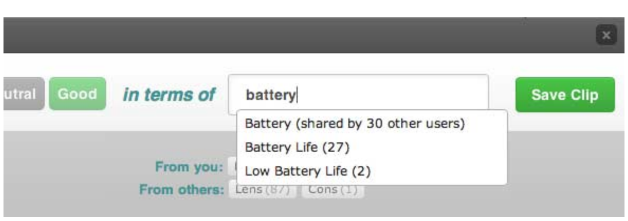

Figure 2. The social autocomplete feature suggests socially relevant dimensions to users as they type in the input field.

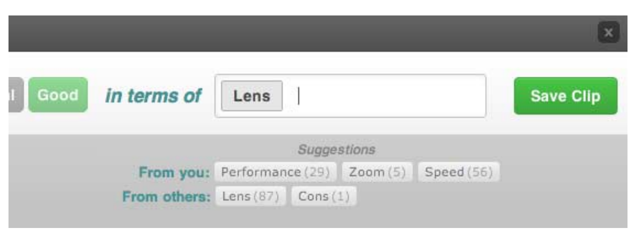

Figure 3. Dimension hinting provides users with suggestions derived from a variety of social and individual contexts.

will accelerate learning the important structure of a new knowledge domain. However, there are many reasons that this might not occur. Below we discuss three core challenges we examine through a controlled experiment.

Overlapping Dimensions

First, simply aggregating dimensions from multiple individuals might not be a useful approach for identifying useful dimensions. One challenge is that users may use dimension names that are too diverse to aggregate. For example, Furnas et al. [14] showed that the probability of two people generating the exact same name for an object was below 20% for a variety of tasks; even expanding to the three most popular names led to less than 50% overlap on most tasks. Furthermore, it is not clear whether dimensions that are more prototypical are more useful for learning about the structure of a domain. Thus the impact of social dimension aggregation depends on whether 1) there is sufficient overlap density of dimensions across individuals, and 2) more prototypical dimensions are found more useful.

Hyp 1.1: *Aggregating dimensions across users will result in overlapping dimensions.*

Hyp 1.2: *Dimensions that are shared across more people will be more useful.*

Noticing and Using Social Data

Second, users might not want to see or interact with the social data from other users. Information foraging is already a cognitively demanding task, and adding even more information from other users might prove overwhelming. We tried to address these issues by providing social data in an unobtrusive way, enabling (but not forcing) the user to drill down for further information (e.g., showing the clips associated with a dimension). However, this approach raises the risk that users would not even notice or look at social data. Conversely, if social data is found sufficiently useful then users might interact with the interface even more than they would if such data was not present.

Hyp 2.1: *People will notice and use social data in the interface.*

Hyp 2.2: *When social data is present people will interact with the interface more than when it is not present.*

Impact of Social Schemas

Finally, even if social schema information is present and noticed that does not mean that it will actually change users' mental representations of an information space. Users might not trust data from others, especially if those others are anonymous, with unknown expertise, motivations, or goals [26, 47]. Schemas from other users might be difficult to understand, especially if the schema provider is already familiar with the domain and the schema consumer is not -- exactly the situation we are trying to support. Relatedly, just seeing the abstracted dimension name may not be sufficient information to help users develop a better mental model of the information space.

However, if the social schema information conveys useful information it could help users to understand the structure of the domain, and do so faster than if such information was not available. We expect such understanding to manifest in more prototypical and useful dimensions generated, espe-

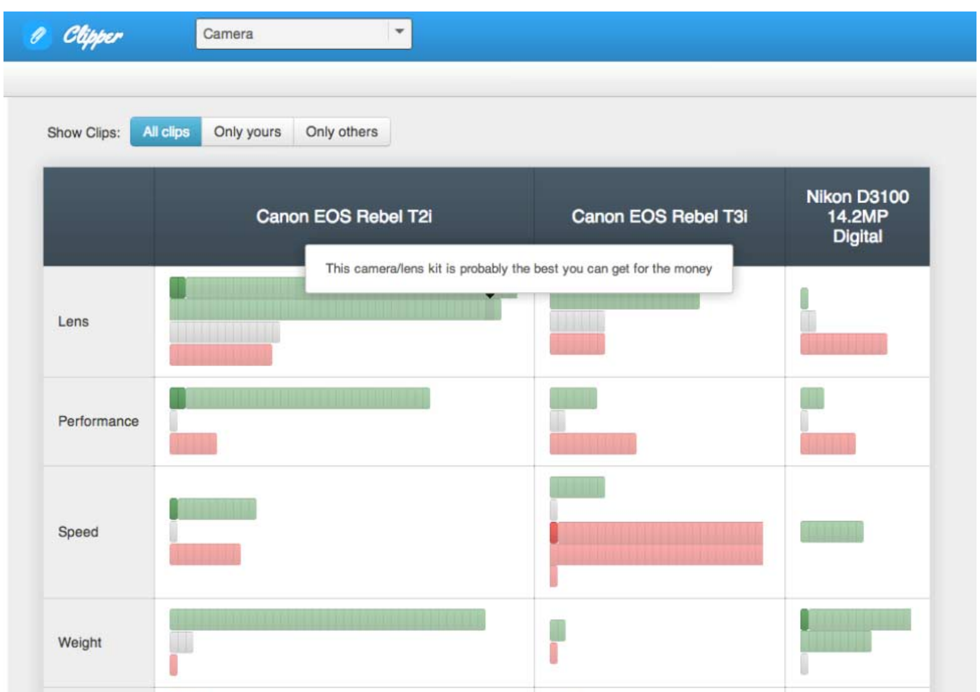

Figure 4. The review table shows a final, graphical summary of a user's workspace.

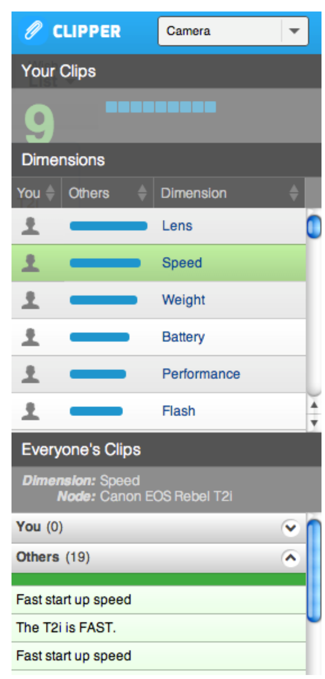

Figure 5. The workspace pane shows dimensions and clips.

cially early on in the process where people are still unfamiliar with a domain. It could also result in better structured workspaces involving fewer "throwaway" or singleton dimensions, which are often created when a person is unfamiliar with the structure of a domain and uses a dimension that later becomes obsolete or irrelevant [25].

Hyp 3.1: *The social condition will generate more prototypical and more useful dimensions than the non-social condition.*

Hyp 3.2: *The social condition will generate more prototypical and more useful dimensions* ***earlier*** *than the non-social condition.*

Hyp 3.3: *Workspaces in the social condition will be better structured than the non-social condition.*

In addition to improving the explicit structure of their workspace, we are also interested in how participants' mental models changed over the period of the task. If social schemas are useful and impactful we would expect participants' mental models to change more when in the social condition, for example in terms of the dimensions that they find important at the end of the task versus what they thought important going into the task. We might also expect them to converge faster on a common vocabulary, as they are exposed to an already-common vocabulary from others' schemas.

Hyp 3.4: *The social condition will lead to greater changes in participants' mental models from before the information foraging task to afterwards.*

Hyp 3.5: *The social condition will converge to a common vocabulary faster than the non-social condition.*

EXPERIMENTAL APPROACH

To test these hypotheses we conducted a within-subjects experiment comparing social vs. non-social versions of the interface. The social version of the interface included the various social features for capturing, integrating, and surfacing schemas, while the non-social version maintained the same basic functionality but was limited so that individuals only saw the results of their own work. Each participant engaged in two information foraging tasks, one with the social interface and one with the non-social interface (counterbalanced across participants).

Evaluating overall schema quality is challenging and unreliable: any single approach will have its weaknesses. Instead, we opted to operationalize multiple metrics in tandem (e.g. structure of the overall workspace, usefulness of dimensions, overlap of dimensions, etc.) to provide converging evidence for overall schema quality in the social and non-social conditions.

Participants

We recruited 64 participants from a university subject pool. Our participants were 28 females and 36 males, and their ages ranged from 18 – 63 years old with an average age of 24.8 years old (*SD* = 8.1 years). Our participants identified having a variety of backgrounds from undergraduate and graduate students to graphic designers and retired military staff. Almost all of our participants identified spending at least 5 hours online engaged in information foraging activities (61/64 participants). Our participants spent 90 minutes conducting our experiment and were paid $15 for their time.

Procedure

The experiment was a within-subjects design in which each participant interacted with both the social and non-social versions of the interface (counterbalanced for order). Participants were asked to perform information foraging tasks in two domains – cameras and carseats (fixed order, respectively). Half of the participants used the "full" interface (which captures node, valence and dimension during clipping) and the other half used a "mini" interface that captured only node during clipping, then valence and dimension in subsequent steps.

In each condition, participants were seated at computers and presented with a browser window with many open tabs comprising the experiment. The tabs contained a survey, two clipping tasks, and instructions and videos throughout to help assist participants during the experiment.

Participants were asked to proceed through the tabs in order. They began by entering demographic information in the survey, then proceeded to perform clipping tasks for both the camera and carseat domains. Each clipping task consisted of saving information from a group of three Amazon product pages using the tool, reviewing the resulting clips in spreadsheet form, and finally seeing a review table of the results. Participants could switch between the products pages as they wished. After completing each task, participants answered more survey questions about their search process and findings.

RESULTS

Overlapping Dimensions

We hypothesized that aggregating dimensions from multiple users would result in dimensions that had high overlap

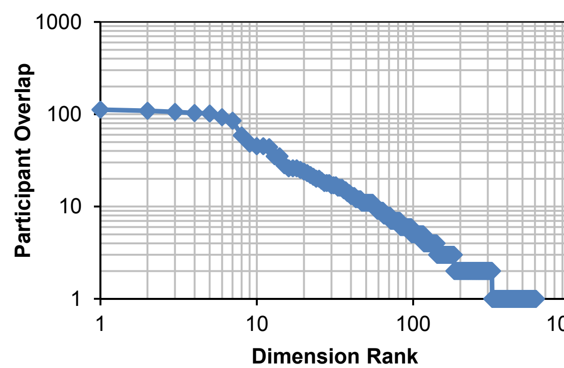

Figure 6. Distribution of dimensions by how many participants' workspaces they overlapped.

across users, and that those dimensions would be found more useful than low overlap dimensions.

Hyp 1.1: *Aggregating dimensions across users will result in overlapping dimensions.*

As shown in Figure 6, we found a highly skewed distribution in which a small number of dimensions were shared across many participants and many dimensions shared by only a few or unshared, similar to the long-tail distributions found in studies of unstructured tagging [17, 40].

Hyp 1.2: *Dimensions that are shared across more people will be more useful.*

Two domain experts independently labeled all 510 unique dimensions generated by our participants according to how useful they would be to new users engaging in the task on a scale of 1 – 5 (1 = low; 5 = high). We defined usefulness as how well the dimension described the domain to a novice. For example, dimensions such as "safety" or "affordability" would score highly on usefulness, while dimensions such as "operate" or "pro & cons" would score lower on usefulness. We compared our domain experts' ratings for the 510 dimensions, and found a weighted Cohen's kappa (κ) of 0.69 indicating a significant level of agreement between our raters on how useful the dimensions were.

Using this data, we performed a regression analysis which showed that increased dimension overlap was significantly positively associated with greater usefulness ($b = 0.21$, $p < .001$), indicating that dimensions that are shared across more people are also more useful. These results provide empirical evidence that surfacing high overlap dimensions may be indeed be a useful strategy for supporting social schema induction.

Noticing and Using Social Data

Hyp 2.1: *People will notice and use social data in the interface.*

We recorded the mouse hovering interactions of all users, and found that people did notice and interact with social data elements. This included both dimensions and clips in the workspace pane (M=10.3; M=2.8), and clips in the review table (M=10.4). These results suggest that many users did notice and interact with social elements of the interface.

We were particularly interested in how people would interact with the dimension hints, as we tested five different

	Hover counts	
	Social	**Non-social**
Dimension	545	116
Clip	169	11
Review	4757	1896

Table 1. Total interactions (as measured by hover counts) for interface elements in the social and non-social conditions.

methods for surfacing potentially useful dimensions, either from the participants' own workspace (their last used, most popular, or random) or from others' workspaces (highest social overlap or random). The social overlap method (i.e., selecting from others' dimensions given the text of the clip) was selected most frequently, although an ANOVA controlling for within-subject effects with post-hoc Tukey HSD comparisons showed that only the difference between the social overlap method and random was significant ($p < .05$), with the difference between social overlap and user-last of marginal significance ($p = .079$).

Hyp 2.2: *When social data is present people will interact with the interface more than when it is not present.*

Although ***Hyp 2.1*** indicates that people interacted with social data, many of the elements they interacted with exist for both the social and non-social conditions. To test whether participants interacted with the interface more in the social condition we aggregated counts of all of the instances when a participant hovered over an element in the dimension pane or in the review view, regardless of whether it was their own or another's clip or dimension (see Table 1). The social condition had significantly more hover counts than the non-social condition ($F_{1,63}=17.4$, $p < .001$). This effect was significant for hover counts for the review page, pane dimensions, and pane dimension clips – in other words, all of the elements that existed in both conditions and had social data embedded in them in the social condition. Interestingly, there was no difference between the social and non-social conditions as to how much they interacted with their own clips ($F_{1,63}=.20$, $p > .6$). This is a somewhat surprising and promising result: increasing their interaction with social data did not reduce participants' interaction with their own data. This suggests that social data was beneficial on top of participants' own data, rather than competing with it for attention.

Impact of Social Schemas

So far two of our prerequisite challenges have been addressed: aggregating dimensions across participants led to more useful dimensions; and participants noticed and interacted with social data. However, the most important question is whether interacting with useful social data actually changes participants' mental schemas. As it can be difficult to identify exactly when and how a mental schema has been changed and improved, below we address this question through converging evidence involving a number of variants operationalizing the question.

Hyp 3.1: *The social condition will generate more prototypical and more useful dimensions than the non-social condition.*

One question we looked was how the structure of a participant's workspace changed as a function of being in the social condition. In this hypothesis we look specifically at effects on prototypicality (i.e., participant overlap) and usefulness of the dimensions generated. We conducted nested

	Overlap		
	Coeff.	SE	P
Intercept	24.6	1.85	***
Social	5.55	2.31	*
Dimension order	-.849	.27	**
Social X Dimension order	-.278	.38	

Table 2. Regression model predicting dimension overlap.

linear regressions with participant as a random factor to determine whether being in the social condition would lead to more prototypical and more useful dimensions. Our results (see Table 2) show that participants' workspaces when in the social condition had dimensions with greater overlap with other participants' dimensions than when in the non-social condition ($b = 3.14$, $SE = .78$, $p < .001$). Dimensions generated in the social condition were also rated as more useful than in the non-social condition ($b = .12$, $SE = .035$, $p < .001$).

Hyp 3.2: *The social condition will generate more prototypical and more useful dimensions* ***earlier*** *than the non-social condition.*

We hypothesized that the social condition would lead to more prototypical and useful dimensions earlier in participants' information foraging process than the non-social condition. While participants did tend to produce higher overlapping dimensions earlier in the process, this effect did not differ between the social and non-social conditions Furthermore, there were no temporal effects associated with rated usefulness of dimensions. These results disconfirm our hypothesis, and suggest that the beneficial effects of surfacing social schemas continue even beyond the early period of information foraging.

Hyp 3.3: *Workspaces in the social condition will be better structured than the non-social condition.*

We operationalized workspaces as being better structured if they had fewer singleton dimensions, which often reflect dimensions that users created but later found irrelevant or obsolete [25]. To determine if workspaces in the social condition were better structured we conducted a nested logistic regression with participant as a random factor, predicting whether a dimension will be a singleton or not. We found that the social condition was a significant negative predictor of generating a singleton dimension ($b = -0.80$, $p < .001$). This effect was quite substantial: seeing others' schemas was associated with a .45 factor reduction in the likelihood of a dimension being a singleton.

Hyp 3.4: *The social condition will lead to greater changes in participants' mental models from before the information foraging task to afterwards.*

Previous research [25] has shown that people's mental models change depending on the support facilitated by a tool. We sought to extend this finding by studying the effects of social data. We hypothesized that the participants' mental models will change the most in the social condition. We used data pertaining to which dimensions our participants thought were important before and after the task to address this question. We find that on average participants in the social condition changed 0.98 dimensions compared to 0.77 for the non-social condition, and the participants in the social condition identified more dimensions (5.41 dimensions) than the non-social condition (5.14 dimensions). However, these findings were only marginally significant (change in dimensions: $t_{63} = 1.284$, $p = 0.103$; dimension sizes: $t_{63} = 1.619$, $p = 0.055$).

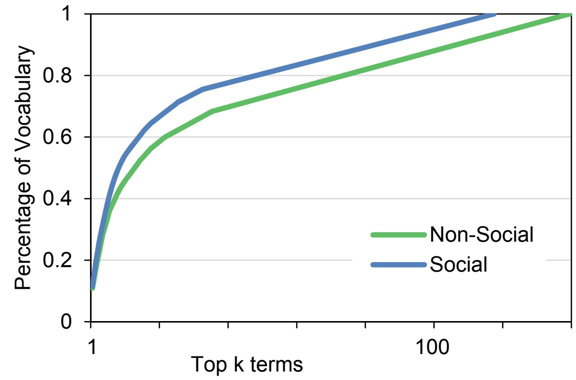

Figure 8. Vocabulary convergence over time in the social and non-social conditions. Users in the social condition achieved greater convergence in less time.

To shed some light on the rationale behind our participants change in mental models, we examined the qualitative data gathered from our surveys. Using our participants' responses to how their choice of dimensions changed over the task, we performed content analysis and categorized their responses into groups depending on the theme. We find that our participants' mental models changed during the task in four key ways: by certain dimensions becoming more important and visible in their mental models (27.3%) resulting in participants prioritizing them over others; by expanding through learning and encountering more information (28.9%); becoming more confident of their decision-making (2.3%); no change due to existing knowledge (28.9%); and other (12.5% participants). These findings support similar results published by [25].

Hyp 3.5: *The social condition will converge to a common vocabulary faster than the non-social condition.*

We hypothesize that the presence of social data will influence people to use a common vocabulary. Once again, using the dimensions identified by our participants as important, we compared the data from the social and non-social condition to understand the characteristics and differences of the vocabularies used in the two conditions. After stemming and cleaning the data (i.e. removing stopwords, e.g. 'the', 'a', 'were', etc.), we computed the frequency of each term within the social and non-social vocabularies.

	Social Condition	Non-Social Condition
Vocab. Size	346	329
# Unique Dimensions	118	140
# Non-Unique Dimensions	228	118

Table 3. Vocabulary size and dimension sharing, for social and non-social conditions. Users in the social condition built a larger vocabulary with more shared dimensions.

Table 3 shows that the characteristics of the social and non-social vocabularies differ: the social condition has a larger number of dimensions, but fewer unique ones used, and there is a large number of repeat dimensions found in the social condition vocabulary (228 vs. 118) compared to the non-social condition. What we can understand from this is that participants in the social condition are identifying more dimensions but are re-using ones used by other participants.

We tabulated this data, and in Figure 8 we see that fewer number of terms comprise the majority of the vocabulary in the social condition compared to the non-social condition. What is also of interest is that the social condition has a much smaller tail compared to the non-social conditions and fewer singleton dimensions suggesting that a smaller vocabulary with a high re-use of terms was found for the social condition.

Subjective Data

After performing both tasks, our participants provided feedback on the usability of the social Clipper tool via our survey. The feedback comprised the overall utility of the tool as well as the helpfulness and use of specific social features. Feedback took the form of evaluations on a 7-point Likert scale as well as open-ended responses. Overall results show that participants rated the system highly (5+) across all features and dimensions, including whether they felt effective and confident using it, whether it was helpful, enjoyable, and easy to use, and whether they would like to continue using it for their own use (Figure 9). Ratings for specific social features showed that participants found all features easy to use and useful, with ratings above 5 for the review table, dimension suggestions, autocomplete, and the dimension pane sidebar (Figure 10).

From our survey we also collected open-ended responses about what aspects of the tool our participants liked and what aspects they would improve. When asked what aspects of the tool our participants liked, a number of responses were given such as its ability to provide social awareness of others' information seeking and having this information at hand (15.6% participants):

"*The community aspect with others' clips was very helpful overall.*" – Participant 48

Most participants liked the features provided by Clipper and identified specific features like the review table and the dimensions (18.7% participants):

"*Review Table was very helpful. Interface for saving clips was intuitive.*" – Participant 8

"*Dimensions, seems to me like a time saver, very easy to use*" – Participant 29

More generally, a number of participants mentioned the positive aspects of how the tool organizes and structures information (17.2%) and the general functionality (18.7%), though this feedback was not specific to the social features we implemented. However, it is encouraging that despite the increased amount of information involved in using social data, participants still mentioned the tool's ease of use (15.6%):

"*It was so simple and intuitive – and I am thrilled to use it for my shopping purposes in the future*" – Participant 41

To assist in future iterations, we asked our participants what improvements they would make. Our participants identified a number of suggestions including additional features (23.4% participants) such as chat and Q&A features, previews of the product, snippets of content and summarizing the main content and dimension. A significant number of participants (23.4%) identified improvements being needed with the performance and responsiveness; these were primarily due to bugs that led to dimension suggestions taking

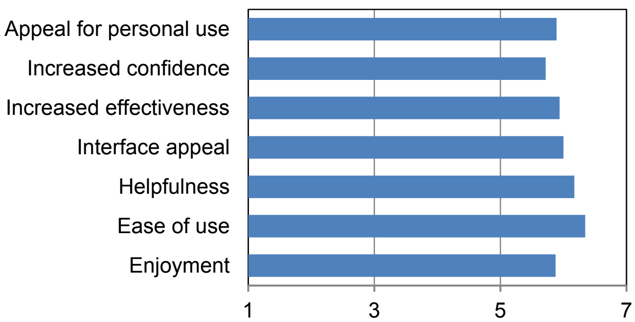

Figure 9. Participants gave the Clipper high marks for measures of usability.

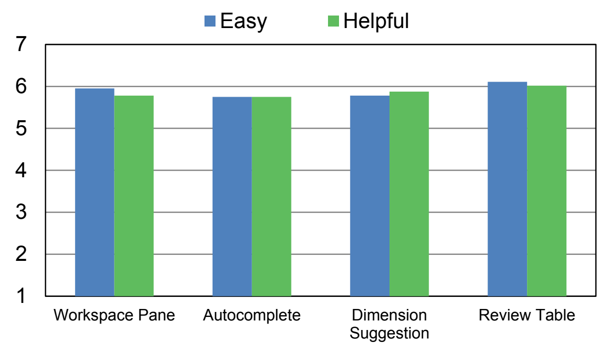

Figure 10. Participants found the interface's social features both easy to use and helpful.

2-3 seconds to generate. Other issues arose from constraints of the study in which we surfaced random suggestions as a baseline, which may have undermined perceived performance.

DISCUSSION

In this paper we examined how capturing, integrating, and surfacing the schemas of users as they engage in information foraging could help others to learn about the structure of a new domain. We designed and implemented novel social features into an information foraging interface which addresses the dual challenges of algorithmically integrating the schemas generated by multiple individuals while simultaneously filtering out less useful dimensions. Users found this interface and its various social elements useful and helpful for information foraging. We conducted an experiment to explore a number of potential challenges for socially augmented information foraging including: whether users would generate useful, overlapping dimensions; whether users would notice and use the dimensions of others; and whether others' dimensions would impact users' workspace structure, mental models, and vocabulary.

Our results indicated that the dimensions generated by users showed significant overlap, and that dimensions with more overlap across users were rated as more useful. This link between dimension overlap and usefulness is an interesting result that may have implications for interfaces aimed at supporting information foraging, and more generally for social tagging. We also showed that users noticed and used the social elements we introduced, with preliminary support for the usefulness of our social dimension suggestion method. These results are promising given our design rationale of making such elements unobtrusive and not forcing users to interact with them. Furthermore, users did not reduce their interaction with non-social elements of the interface, suggesting that the social elements were useful but not competing with the non-social elements.

Finally, we presented results indicating that being exposed to others' schemas changed users' behavior in a number of ways. In the social condition, users generated more prototypical and more useful dimensions than when they were in the non-social condition. They also had better structured workspaces with fewer singleton dimensions and converged on a common vocabulary faster than when in the non-social condition. Interestingly, these results were consistent whether participants added structure to their clips during (full version) or after (mini version) clip creation, suggesting that the workspace pane was effective in influencing their mental models even if they were not explicitly adding dimensions at the time of clipping.

Our approach demonstrates a scalable and effective way for users to learn from each others' schemas while learning about a new domain online. In many cases when users search they are not looking for a single piece of content but instead trying to understand the structure of a new information space and how multiple pieces of content fit into that structure. Framing search as a process of learning and schema induction [12, 25, 33] highlights this as a critical need that we begin to address here.

Our results provide an important step towards a future of distributed sensemaking [12], in which the effort that individuals put forth in learning about a new domain is not lost but instead accelerates the learning process for individuals coming after them. A key challenge for that future is determining how to capture, aggregate, and surface schemas so that they are useful to others. By algorithmically aggregating schemas and showing that more overlapping schemas are also rated as more useful, we identify a method of overcoming the "hump" paradox introduced by [12] -- that individual schemas are less valuable than starting from scratch, but iterated schemas are more valuable. Our approach also naturally deals with the challenge raised in [25], that early dimensions generated before a user is familiar with a domain become obsolete and irrelevant: by aggregating dimensions across users, these idiosyncratic and less relevant dimensions drop to the bottom of the pile.

At the same time, our study suggests a number of areas that would profit from further research. While we aim to minimize user effort through features such as dimension hinting, computational approaches to automatically infer important dimensions could further streamline the process. Aggregating and summarizing similar clips is another area that computational approaches could help. However, any such approaches would need to provide clear orientation, provenance, and the ability for the user to override the machine in order to support and privilege the process of learning and schema induction in information foraging.

There may also be drawbacks to using the social Clipper system for users with differing viewpoints than the common view. For example, people may use different dimensions for the same information if they have different goals, background, experience, or cultures, and valuable minority views could be lost through aggregation. However, there may be ways of addressing this, such showing dimensions from people like themselves, where similarity could be determined by browsing or clipping characteristics.

In summary, we have demonstrated an approach for capturing, aggregating, and surfacing schemas from others and its benefits for information foraging. Further reducing the costs and increasing the benefits for schema acquisition by novices will help us move to a future where new learners no longer have to start from scratch, and can stand on the schemas of their fellow giants.

ACKNOWLEDGMENTS

This work was supported by NSF grants IIS-1149797, IIS-1217559, OCI-0943148, IIS-0968484, IIS-1111124, Bosch, Google, and Microsoft.

REFERENCES

1. Amershi and Morris, M.R. CoSearch: A system for colocated collaborative web search. In *Proc. CHI 2008.*
2. Baldonado, M.W. and Winograd, F. SenseMaker: an information-exploration interface supporting the contextual evolution of a user's interests. In *Proc. CHI 1997.*
3. Bao, S., Wu, X. Fei, B., Xue, G., Su, Z. and Yu, Y. Optimizing web search using social annotations. In *Proc. WWW 2007.*
4. Card, S.K., Robertson, G.G. and York, W. The WebBook and the Web Forager: Video use scenarios for a World-Wide Web information workspace. In *Proc. CHI 1996.*
5. Catledge, L.D. and Pitkow, J.E. Characterizing browsing strategies in the World Wide Web. In *Proc. WWW 1995.*
6. 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.
7. Choo, C.W., Detlor, B. and Turnbull, D. Information seeking on the Web: An integrated model of browsing and searching. *First Monday* 5, 2, (2000).
8. Cutting, D.R., Karger, D.R. and Pedersen, J.O. Constant interaction-time scatter/gather browsing of very large document collections. In *Proc. SIGIR 1993.*
9. Dmitriev, P.A., Eiron, N., Fontoura, M. and Shekita, E. Using Annotations in Enterprise Search. In *Proc. WWW 2006,* 811-817.
10. Dontcheva, M., Drucker, S., Wade, G., Salesin, D. and Cohen, M. Collecting and Organizing Web Content. In *Proc. SIGIR 2006.*
11. Evans, B.M. and Chi, E.H. Towards a Model of Understanding Social Search. In *Proc. CSCW 2008.*
12. Fisher, K., Counts, S. and Kittur, A. Distributed sensemaking: improving sensemaking by leveraging the efforts of previous users. In *Proc. CHI 2012,* 247-256
13. Freyne, J., Farzan, R., Brusilovsky, P., Smyth, B. and Coyle, M. Collecting community wisdom: integrating social search & social navigation. In *Proc. UIST 2007,*
14. Furnas, G.W., Landauer, T.K., Gomez, L.M. and Dumais, S.T. The vocabulary problem in human-system communication. *CACM 30,* 11, ACM Press (1987), 964-971.
15. Glance, N. S. Community search assistant. In *Proc. IUI 2001.*
16. Goldberg, D., Nichols, D., Oki, B.M. and Terry, D. Using collaborative filtering to weave an information tapestry. *CACM 35*, 12, (1992), 61–70.
17. Golder, S.A. and Huberman, B.A. Usage Patterns of Collaborative Tagging Systems. *Journal of Information Science 32,* 2, (2006)*,* 198-208.
18. Hammond, T., Hannay, T., Lund, B. and Scott, J. Social book marking tools (i) - a general review. *D-Lib Magazine 11,* 4, (2005).
19. Hofgesang, P.I. Web personalisation through incremental individual profiling and support-based user segmentation. In *Proc. IEEE 2007.*
20. Hong, L, Chi, E.H., Budiu, R., Pirolli, P. and Nelson, L. SparTag.us: Low Cost Tagging System for Foraging of Web Content. In *Proc AVI 2008*.
21. Hummel, J.E. and Holyoak, K.J. Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review 104*, (1997), 427-466.
22. Kandogan, E. and Shneiderman, B. Elastic Windows: evaluation of multi-window operations. I*n Proc CHI 1997.*
23. Kellar, M., Watters, C. and Shepherd, M. A Field Study Characterizing Web-Based Information-Seeking Tasks. *Journal of the American Society for Information Science 58,* 7, (2007), 999-1018.
24. Klein, G., Moon, B. and Hoffman, R.R. Making Sense of Sensemaking 2: A Macrocognitive Model. In *Proc. IEEE 2006.*
25. Kittur, A., Peters, A.M., Diriye, A., Telang, T. and Bove, M.R. Costs and benefits of structured information foraging. In *Proc. CHI 2013*, 2989-2998.
26. Kittur, A., Suh, B. and Chi, E.H. Can you ever trust a wiki?: impacting perceived trustworthiness in Wikipedia. In *Proc. CSCW 2008*, 477-480.
27. Konstan, J.A., Miller, B.N., Maltz, D., Herlocker, J.L., Gordon, L.R. and Riedl. J. GroupLens: Applying collaborative filtering to Usenet news. *CACM 40,* 3, (2007), 77–87.
28. Large, A., Beheshti, J. and Rahman, T. Gender Differences in Collaborative Web Searching Behavior: An Elementary School Study. *Information Processing and Management 38*, (2002), 427-433.
29. Lee, K.J. What Goes Around Comes Around: An Analysis of del.icio.us as Social Space. In *Proc. CSCW 2006.*
30. Lerman, K. Social Information Processing in News Aggregation. In *Proc. IEEE 2007.*
31. Maarek, Y.S., Jacovi, M., Shtalhaim, M., Ur, S., Zernik, D. and Ben-Shaul, I.Z. WebCutter: a system for dynamic and tailorable site mapping. In *Proc. WWW 1997,* 1269-1279.
32. MacKay, B., Kellar, M. and Watters, C. An evaluation of landmarks for re-finding information on the Web. *Ext. Abstracts CHI 2005.*
33. Marchionini, G. Information seeking in electronic environments. Cambridge, UK: Cambridge University Press, (2000).
34. Mathes, A. Folksonomies – Cooperative Classification and Communication through Shared Metadata. http://www.adammathes.com/academic /computer-mediatedcommunication/folksonomies.html. (December 2004).
35. Millen, D.R., Feinberg, J. and Kerr, B. Dogear: Social bookmarking in the enterprise. In *Proc. CHI 2006.*
36. Morris, M.R. A Survey of Collaborative Web Search Practices. In *Proc. CHI 2008.*

37. Morris, M.R. and Horvitz, E. SearchTogether: An Interface for Collaborative Web Search. In *Proc. UIST 2007.*
38. Munro, A.J., Hook, K., and Benyon, D.R. Eds. *Personal and Social Navigation of Information Space.* Springer-Verlag (1999).
39. Muralidharan, A, Gyongyi, Z. and Chi, E. Social annotations in web search. In *Proc. CHI 2012*, ACM Press (2012), 1085-1094.
40. Marlow, C., Naaman, M., boyd, d. and Davis, M. HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, To Read. In *Proc. HT06,* 31-40.
41. Paul, S.A. and Morris, M.W. CoSense: Enhancing Sensemaking for Collaborative Web Search. In *Proc. CHI 2009.*
42. Paul, S.A., Reddy, M.C. Sensemaking in collaborative web search. Human Computer Interaction 26, 72–122 (2011).
43. Pirolli, P. and Card, S.K. Information foraging in information access environments. In *Proc. CHI 1995.*
44. Pirolli, P. and Card, S.K. Information foraging. *Psychological Review 106*, (1999), 643-675.
45. Russell, D.M., Stefik, M.J., Pirolli, P. and Card, S.K. The cost structure of sensemaking. In *Proc. CHI 1993*, 269-276.
46. Schraefel, M., Zhu, Y., Modjeska, D., Wigdor, D. and Zhao, S. Hunter gatherer: interaction support for the creation and management of within-web-page collections. In *Proc. WWW 2002.*
47. Sharma, N. Role of available and provided resources in sensemaking. In *Proc. CHI 2011*.
48. Shneiderman, B. Designing trust into online experiences. *CACM 43*, 12, (2000), 57-59.
49. Smyth, B., Balfe, E., Boydell, O., Bradley, K., Briggs, P., Coyle, M. and Freyne, J. A liveuser evaluation of collaborative web search. In *Proc. International Joint Conference on Artificial Intelligence 2005*.
50. Tauscher, L. and Greenberg, S. How people revisit Web pages: Empirical findings and implications for the design of history systems. *International Journal of Human–Computer Studies 47*, (1997), 97–137.
51. Twidale, M., Nichols, D. and Paice, C. Browsing is a Collaborative Process. *Information Processing and Management 33*, 6, (1997), 761-783.
52. Wu, X., Zhang, L. and Yu, Y. Exploring Social Annotations for the Semantic Web. In *Proc. WWW 2006*, 417-426.
53. Yanbe, Y., Jatowt, A., Nakamura, S. and Tanaka, K. Can social bookmarking enhance search in the web? In *Proc. ACM/IEEE-CS joint conf. on Digital libraries 2007,* 107.
54. Zanardi, V. and Capra, L. Social ranking: uncovering relevant content using tag-based recommender systems. In *Proc. ACM RecSys 2008*, 51.