Supporting Ideation by Integrating Exploratory Search, Browsing, and Curation

Yin Qu
Interface Ecology Lab
Department of Computer Science and Engineering
Texas A&M University
yin@ecologylab.net

ABSTRACT

This research integrates exploratory search, browsing, and curation in information-based ideation (IBI) environments, to support users finding, discovering, and combining web content. Information-based ideation refers to open-ended activities in which humans work to imagine, generate, and develop new problems and solutions. We use web semantics as a basis for summarizing and representing heterogeneous content from diverge sources involved in ideation tasks. To enable working with web semantics, we develop a novel type system that brings together data models, dynamic extraction, and presentation of semantic information. Based on the web semantics type system, we build interfaces that preserve contexts during exploratory browsing, and plan to build integrated interactive environments to address exploratory search, curation, and ideation. We will investigate how integrated environments affect people's practices with finding and combining information, and derive new principles for supporting exploratory search, browsing, curation, and ideation. Methods, techniques, and findings developed through this research have the potential to transform infrastructures and human experiences of the web.

Categories and Subject Descriptors

H.5.2 [**Information interfaces and presentation: user interfaces**]: Graphical user interfaces

1. MOTIVATION

To generate and develop new problems and solutions, people perform open-ended ideation tasks, such as planning a wedding or designing a course. These tasks require exploratory search for and browsing of diverse content. The present research develops a novel web semantics type system as the basis for representing diverse, contextual searches, results, and linked documents, to support exploration.

Prominent researchers have emphasized the importance of developing new methods for synthesizing different types of information from different sources into unified representations, to help users focus on *meaning* in the context of tasks [23]. Kerne et al call this process of synthesis, *curation*, and show how curation is inherent in the performance of ideation tasks [10]. We apply the new semantic type system to generate consistent representations across exploratory search, browsing, and curation interfaces, integrated to form an ideation support environment.

Ideation tasks involve searching for new information. Found information can be used to remedy incomplete or inaccurate understandings [1], solve task-related problems, or stimulate the generation of new ideas [21, 9, 24, 10]. Search and browsing are important methods people use to find information on the web, both purposefully and serendipitously [13, 22]. Exploratory search extends search from a one-time query-document matching act to an iterative process, involving browsing, learning, and discovery [25, 12]. Prior work finds exploratory search useful for understanding collections and relationships [20, 27, 5, 28]. Exploratory browsing and search are intrinsic to the performance of sensemaking [19] and information-based ideation tasks [10, 18].

Interfaces supporting ideation need to address fundamental cognitive limitations. For instance, human working memory, which is essential to comprehension, learning, and ideation, only holds approximately 4 elements at a time [15, 4]. However, the number of elements that must be understood, reflected upon, and synthesized in the performance of contextualized ideation tasks is often larger. Typical browsers present information in separate tabs and windows, making it difficult to keep track of relationships among documents, or think about and combine information across documents. As the result, users may lose orientation during navigation, or digress from the task at hand [14, 6]. We need to discover interactive systems that help users maintain context during exploration, sensemaking, and ideation.

Curation is the creative process of gathering, assembling, annotating, and exhibiting found information elements [10]. It is valuable for curation environments to provide external representations to help users—who by nature possess limited working memory—think about and combine information encountered during search and browsing to form new ideas. Curation environments can promote synthesis by supporting iteratively gathering found objects, representing and interpreting objects and relationships, integrating thoughts and explanations as annotations, formulating solutions across objects, and presenting formulated solutions as holistic exhibits for reflection and communication. During synthesis, new information needs may emerge, leading

CHIIR '16 March 13-17, 2016, Carrboro, NC, USA

ACM ISBN 978-1-4503-3751-9/16/03.

DOI: http://dx.doi.org/10.1145/2854946.2854948

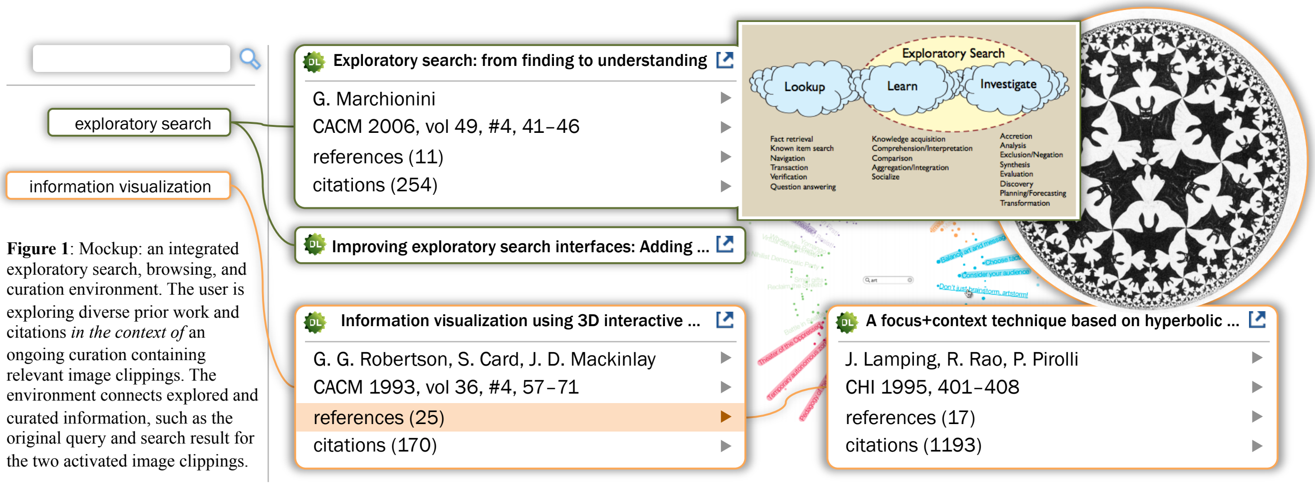

Figure 1: Mockup: an integrated exploratory search, browsing, and curation environment. The user is exploring diverse prior work and citations *in the context of* an ongoing curation containing relevant image clippings. The environment connects explored and curated information, such as the original query and search result for the two activated image clippings.

to further exploratory search and browsing [8, 16]. Thus, search, browsing, curation, and synthesis mutually motivate one another, and are inherent to information-based ideation.

However, despite this inherent relationship, prior tools separate exploratory search, browsing, and curation. This forces users to split attention, resulting in extraneous cognitive load [3, 26]. The motivation of this research is to investigate and develop methods for integrating and contextualizing exploratory search, browsing, and curation, to reduce cognitive load, and support synthesis for ideation.

2. RESEARCH QUESTIONS

This research investigates the following research questions:

1. How do people find, discover, and combine web information to form new ideas in practice?
2. How to integrate exploratory search, browsing, and curation in one environment? How to support a large and growing number of types, sources, and use cases?
3. How do an integrated environment and consistent representations affect users in exploratory search, browsing, curation, and ideation?
4. How to contextualize exploratory search and browsing with curation? How does seeing search results and linked documents in context of curated information affect cognition and experience?
5. What are implications for the design of future web interfaces?

3. METHODOLOGY

The present research investigates integration of exploratory search, browsing, and curation, using a novel web semantics type system as a basis. *Web semantics* involve significant attributes, descriptions, representations, and relationships of objects people find on the web. During exploratory search and browsing, web semantics can support derivation of summaries of search results and linked web pages, enabling users to see diverse significant information in a single context. Derived summaries can be brought into curation products, providing external representations for users to understand, combine, synthesize, and reflect on.

We use a web semantics type system as the basis for deriving document summaries [17]. With this method, we do not rely on websites to publish structured semantic information, such as RDF [11]. Instead, structured semantic information is extracted from regular, semi-structured web pages, using code snippets called *wrappers*. To derive summaries, wrappers integrate data models, extraction rules, and presentation semantics, and can be reused through inheritance and polymorphism [2]. This research will extend extraction rule capabilities to support diverse semantic formats and schemes used on the web, including microdata, Facebook Open Graph, and Twitter Cards. Presentation semantics regulate how semantic information is presented to end users, ensuring visual consistency and usability. Types enable dynamic resolution of a web page to a semantic type at runtime, containing a contextually specific data model, extraction rules, and presentation semantics. Polymorphism enables the development of rich interfaces that operate on a base type to dynamically present detailed semantic information from heterogeneous web pages encountered at runtime.

This research involves the following activities:

1. Develop a web semantics type system that integrates data models, extraction, and presentation. The working hypothesis is that the new web semantics type system can be used as the basis for extracting and presenting usable summaries of webpages for many information types, sources, and use cases.
2. Use extracted summaries to build integrated exploratory search and browsing interfaces. The working hypothesis is that the integrated exploratory search and browsing interface will mitigate issues with typical browsers, such as disorientation and digression.
3. Use the new web semantics type system as the basis for integrating dynamic exploratory search and browsing interfaces into a curation environment (Figure 1). We hypothesize that the integrated environment will reduce the cognitive load of splitting attention, and promote synthesis and ideation.
4. Conduct laboratory and field studies to evaluate how the integrated environment affects people's exploratory search, browsing, curation, and ideation experiences.

4. CURRENT PROGRESS

We have developed a functional prototype of the web semantics type system, and built an example dynamic exploratory browsing interface using it. The example interface presents summaries from linked web pages in one space, starting from a particular source document, to facilitate navigation without leaving the current context. A preliminary study shows that the example interface helps users maintain orientation and focus on the task at hand during exploratory browsing [17]. We also used the type system to provide contextual information in a curation environment, which helps users understand, explore, and reflect on their curations [24].

5. FUTURE PLAN

We are working on a dynamic interface that integrates exploratory search and browsing, using the new web semantics type system to support deriving summaries for search results and linked pages. The interface presents summaries in a single context, to help users maintain orientation and focus. Queries, search results, and browsing history will be preserved across sessions, to support iterative exploratory search and browsing spanning an extended period of time. We plan to further integrate the exploratory search and browsing interface with a web-based curation environment (Figure 1) to enable seeing new information in the context of the task at hand, and so support synthesis and ideation.

We will evaluate the integrated environment's impact on people's exploration and ideation experiences through laboratory and field studies. For controlled laboratory studies, we will invite researchers to explore prior work, curate, and ideate, to form project ideas related to pre-selected topics. We will build on Kerne et al [10] and Jain et al's [7] ideation metrics, and develop practical methods for measuring exploratory search, browsing, curation, and ideation. We will collect qualitative data from subjects and corroborate with quantitative metrics, to gain a deeper understanding of their practices and experiences. We will distill implications for designing future exploratory search, browsing, curation, and ideation environments from study results.

6. REFERENCES

[1] N. Belkin. Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information and Library Science*, 5:133–143, 1980.

[2] G. Booch et al. *Object-Oriented Analysis and Design with Applications*. Pearson Education, 2007.

[3] P. Chandler and J. Sweller. Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4):293–332, 1991.

[4] N. Cowan. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *The Behavioral and brain sciences*, 24(1):87–114, 2001.

[5] S. Dumais et al. Stuff i've seen: A system for personal information retrieval and re-use. In *Proc of ACM SIGIR*, pages 72–79, 2003.

[6] C. L. Foss. Detecting lost users: Empirical studies on browsing hypertext. Technical report, 1989.

[7] A. Jain et al. Evaluating TweetBubble with ideation metrics of exploratory browsing. In *Proc. of ACM Creativity and Cognition (C&C)*, pages 53–62, 2015.

[8] K. Jarvelin and P. Ingwersen. Information seeking research needs extension towards tasks and technology. *Information Research*, 10(1), 2004.

[9] A. Kerne et al. combinFormation: Mixed-initiative composition of image and text surrogates promotes information discovery. *ACM TOIS*, 27(1):1–45, 2008.

[10] A. Kerne et al. Using metrics of curation to evaluate information-based ideation. *ACM TOCHI*, 21(3):14:1–14:48, 2014.

[11] O. Lassila and R. R. Swick. Resource description framework (RDF) model and syntax specification. 1999.

[12] G. Marchionini. Exploratory search: From finding to understanding. *Commun. ACM*, 49(4):41, 2006.

[13] G. Marchionini and B. Shneiderman. Finding facts vs. browsing knowledge in hypertext systems. *Computer*, 21(1):70–80, 1988.

[14] R. McAleese. Navigation and browsing in hypertext. In *Hypertext: Theory into practice*, pages 6–44. Intellect Books, Oxford, 1989.

[15] G. A. Miller. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(1):81–97, 1956.

[16] P. Pirolli and S. Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proc. of International Conference on Intelligence Analysis*, volume 5, pages 2–4, 2005.

[17] Y. Qu et al. Metadata type system: integrate presentation, data models, and extraction to enable exploratory browsing interfaces. In *Proc. of ACM EICS*, pages 107–116, 2014.

[18] Y. Qu and G. W. Furnas. Sources of structure in sensemaking. In *Proc. of ACM CHI EA*, pages 1989–1992, 2005.

[19] D. M. Russell et al. The cost structure of sensemaking. In *Proc. of ACM CHI*, pages 269–276, 1993.

[20] m. c. scheraefel et al. mSpace: interaction design for user-determined, adaptable domain exploration in hypermedia. In *Proc. of AH*, 2003.

[21] J. J. Shah et al. Collaborative sketching (C-Sketch) – an idea generation technique for engineering design. *The Journal of Creative Behavior*, 35(3):168–198, 2001.

[22] J. Teevan et al. The perfect search engine is not enough: A study of orienteering behavior in directed search. In *Proc. of ACM CHI*, pages 415–422, 2004.

[23] J. J. Thomas and K. Cook. A visual analytics agenda. *Computer Graphics and Applications*, 26(1):10–13, 2006.

[24] A. M. Webb et al. Promoting reflection and interpretation in education: Curating rich bookmarks as information composition. In *Proc. of ACM Creativity & Cognition*, pages 53–62, 2013.

[25] R. W. White et al. Introduction to supporting exploratory search. *Commun. ACM*, 49(4):36–39, 2006.

[26] M. L. Wilson and m. schraefel. Improving exploratory search interfaces: Adding value or information overload? In *Proc. of 2nd Workshop on HCIIR*, pages 81–84, 2008.

[27] K.-P. Yee et al. Faceted metadata for image search and browsing. *Proc. of ACM CHI*, pages 401–408, 2003.

[28] J. Zhang and G. Marchionini. Evaluation and evolution of a browse and search interface: Relation Browser++. In *Proc. of National Conference on Digital Government Research*, pages 179–188. Digital Government Society of North America, 2005.