

combinFormation: Exploring Multiple Searches Together through the Mixed-Initiative Composition Space

Eunyee Koh

Interface Ecology Lab
Computer Science Department
Texas A&M University, 77843-3112, USA
eunyee@cs.tamu.edu

Andruid Kerne

Interface Ecology Lab
Computer Science Department
Texas A&M University, 77843-3112, USA
andruid@cs.tamu.edu

ABSTRACT

combinFormation enables people to search, browse, collect, and compose information on the Web, supporting information exploration and discovery. Through mixed-initiatives, a visual representation of news feeds, searches, and document crawls is assembled. A generative agent collects information, forms visual surrogates and continuously develops a composition that represents them as a connected collection. The composition space enables direct manipulation of the results that the agent is generating.

We present a usage scenario, in which a participant engages in mixed-initiative composition in order to explore topics in the day's news that she is particularly interested in. A composition develops iteratively, over time, using information from news feeds, and popular web searches and sites. The participant manipulates the information sources, and the visual design of the composition. In addition to the scenario, the mixed-initiative composition methodology is validated with data from a field study.

Author Keywords

Exploratory search, search interface, information discovery, compositional hypermedia, surrogates, visual interfaces, semantic web

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

Sometimes, people lack the knowledge or contextual awareness to formulate queries that meet their information needs. As affordances, the text field and results list be

inadequate for successful task completion. Users may require browsing, exploring, and learning, rather than just looking up search results. In such exploratory search tasks, the human formulates a tentative query and moves forward, by browsing retrieved information, seeking, and obtaining cues about where the next steps lie, and iterating [9].

In order to support exploratory search, Marchionini has suggested more active user involvement in the search interface, by blending querying and browsing strategies [7]. Cutrell and Dumais developed the Phlat interface for personal search, which combines keyword search and metadata browsing in a seamless manner [2]. This allows people to quickly and flexibly find information based on whatever they may remember about the information that they are looking for.

COMBINFORMATION MIXED-INITIATIVES

Like the prior research, combinFormation integrates searching, browsing, and navigating. However, combinFormation is a mixed-initiative [3] exploratory search and creativity support tool that integrates processes of searching, browsing, collecting, mixing, organizing, and



Figure 1. re:mix interface: users can put together any number of searches (using multiple engines), sites, and feeds.



Figure 2. An early stage of the composition developed in the scenario. At this point, the seeds come from the *news re:collection*.

thinking about information [4, 5]. Digital information functions to support creative idea generation. By mixed-initiative, we mean that generative agents work in partnership with a human participant to collect relevant information, and represent the collection in the form a visual composition that changes over time. Images and text engage complementary cognitive subsystems. Each collection of information resources is represented as a connected whole. This promotes information discovery, the emergence of new ideas in the context of information.

combinFormation redesigns the searching and browsing interface to better support exploratory search. Use of combinFormation typically begins with launch, in which seeds are selected, to feed the generative agents. Each seed specifies an information source. The *re:mix* interface enables the participant to specify and assemble any number of seeds (see Figure 1). Each seed specifies either a document (HTML, PDF, or file system directory) location, an RSS feed, or a search query. For each search query, a particular search engine source may be selected, including Google, Yahoo, Yahoo Images, Yahoo News, del.icio.us, or Flickr. The *re:collections* interface enables the selection of

previously curated seed collections. One such collection, *news*, assembles material from online news sources, including The New York Times and The BBC. Another, *popular buzzelicious*, combines the searches currently most popular on Yahoo with the items most tagged in del.icio.us.

combinFormation enables people to explore, compare, collect, and compose multiple information resources in a composition space. In the composition space, multiple seeds are mixed and visualized with image and text surrogates. Metadata details on demand are provided in-context [5]. Intermediate results can be directly manipulated by the user at any stage in the process.

Image and Text Surrogates

combinFormation represents search results with image and text clippings from documents. These clippings function as surrogates that the participant can use to navigate back to source web pages. Dual coding of results with image and text surrogates promotes cognition [8]. Each image surrogate is formed from a significant image from the document. Instead of using a summary, we build texts surrogate by extracting significant textual phrases. When a

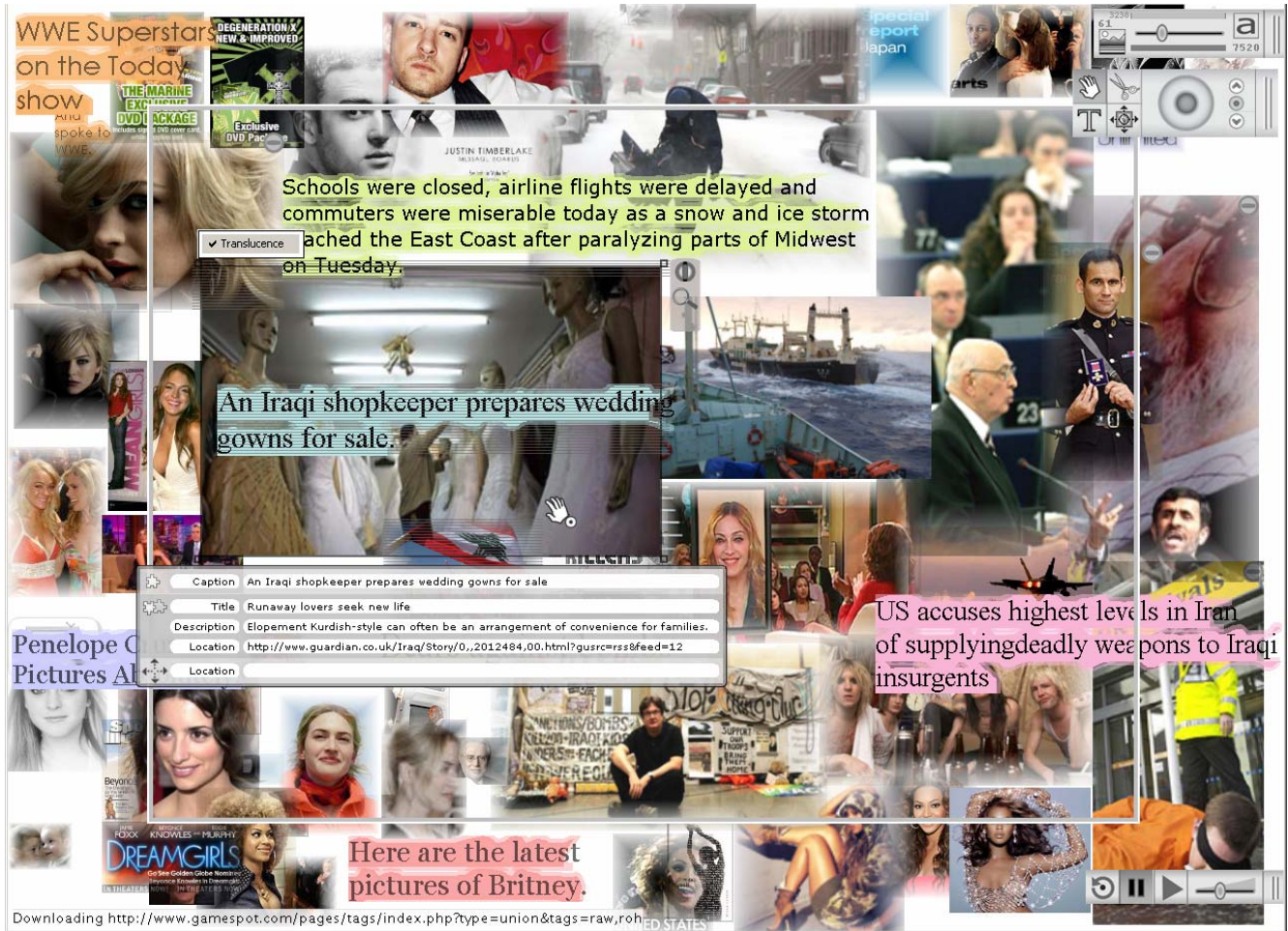


Figure 2. Later in the usage scenario, the participant has mixed the popular buzzelicious in with the news. In-context metadata details on demand are displayed below an element activated through brushing by simple mouse over.

surrogate is brushed with mouse over, the visual representations are augmented by metadata details on demand [1], such as the title of a document, and the caption of an image (see Figure 2). The details are provided in-context, to reduce cognitive load. The goal is to use surrogates to focus the representation of finer-grained ideas that lie within documents, as appropriate, in ways that reflect the intentions of document authors, and the needs of people collecting information. The surrogate serves to represent significant ideas from the document, in the context of the emerging collection and the associated task. The surrogate, in the composition space, also affords interest expression, which enables the participant to provide relevance feedback that affects the system's semantic model of the information and the participant's interests. This, in turn, affects the agent's on-going generation and exploration of the collection.

Composition Space

The composition space is a means for developing a representation of the collection as a connected whole. It is a mixed-initiative space, in which the agent and the human

participant can work concurrently. Within the composition space, the human can directly position, resize, color, edit, and remove surrogates. S/he can also navigate from each surrogate in the composition space back to its hyperlink and source document, in a traditional web browser. A tape recorder transport enables the participant to pause the agents, taking more control of the composition. When the human finds information that s/he wishes to collect in a source document, s/he can click and drag to select it in the web browser, then drag it over to combination, and drop it into the composition space. For the drag and drop operation to represent material that is selected in the source web page as surrogates in the composition, the program needs the context of the web address of the source container document. We developed a Firefox plug-in to pass this contextual metadata through drag and drop.

USAGE SCENARIO

A usage scenario was developed, with the simple exploratory search goal of exploring what is going on today in the world and on the internet, developing a focus on material of interest to a participant. The participant begins

by selecting the *news* collection using the *re:collections* launch interface. As you can see in Figure 2, materials flowed into the composition on diverse topics, including the war in Iraq, a winter storm, and Madonna's appearance on the Oprah Winfrey Show. The participant decides that she is interested in simultaneously exploring the current most popular topics on the internet, so she returned to *re:collections*, while the composition was running, and selected *popular buzzelicious*. *combinFormation* obtained the 15 most popular Yahoo searches via XML, then issued the searches, and downloaded 15 documents for each search. Likewise, a feed specifying the most popular documents from *del.icio.us* was obtained, and these documents were downloaded. All the documents were processed to form image and text surrogates, which then became surrogate candidates for the composition. While this was happening, the participant used *combinFormation*'s direct manipulation tools to reposition and resize surrogates of interest, and eliminate those that seemed uninteresting. A later stage of the exploration is presented in Figure 3.

DATA

Beyond this usage scenario, to validate the efficacy of the mixed-initiative composition methodology, we are conducting an on-going field study, in which *combinFormation* is being used on assignments in an undergraduate course on The Design Process [5]. The information discovery [6] assignments, which require the development of new inventions, require open-ended information exploration. Students used either *combinFormation* or Google and Word for information exploration. We investigated how students performed on their assignments, according to grades assigned by the teaching assistant. The ones using *combinFormation* performed better than those using Google and Word, and the results were statistically significant [5]. The TA also found that representations of collections assembled in the medium of composition of image and text surrogates are better than textual lists for understanding, developing ideas, and the communication of meaning.

Recently, student participants described the experience:

PI: When using this program, I learned that after I have all information gathered in one area, it is easier to play around and come up with something new. It takes in all the different ideas in our brains and just basically puts it on the table for us.

P199: The current and future application of ambient intelligence was very interesting especially as it applied to the medical field and to our everyday lives. It was something I did not know very much about, but there was a lot of info on. The really neat thing about the program was that instead of reading tons of long articles to find ideas or

get concepts, it pulls quick quotes that allow you to connect lots of ideas.

CONCLUSION

combinFormation supports people in exploring, understanding and learning about information on the Web, by providing a mixed-initiative composition space to search, browse, compare, and compose information. The composition space promotes exploration and information discovery, by juxtaposing relevant surrogates. This provokes the human mind to consider alternative explanations about the relationships between connected elements. This paper identifies significant interface components to support exploratory search, which are search query mixing, image and text surrogates, the composition space, and in-context details on demand. *combinFormation* development is conducted through an on-going iterative design process, based on feedback from field and laboratory studies. We will continue to advance the interface, visualization, and semantic model to better support exploratory search and information discovery.

ACKNOWLEDGMENTS

Support is provided by NSF grant IIS-0633906.

REFERENCES

1. Ahlberg, C., Shneiderman, B., Visual Information Seeking: Tight Coupling of Dynamic Query Filters With Starfield Displays, *Proc of CHI 1994*, pp. 313-317
2. Cutrell, E., Dumais, S.T., Exploring Personal Information, *Communications of the ACM*, 49(4), 2006, 50-51.
3. Horvitz, E., Principles of Mixed-Initiative User Interfaces, *Proc CHI 1999*, pp. 159-166.
4. Interface Ecology Lab (2006), *combinFormation*, <http://ecologylab.cs.tamu.edu/combinFormation/>, last accessed 01/08/06.
5. Kerne, A., Koh, E., Dworaczyk, B., Mistrot, M.J., Choi, H., Smith, S.M., Graeber, R., Caruso, D., Webb, A., Hill, R., Albea, J., *combinFormation: A Mixed-Initiative System for Representing Collections as Compositions of Image and Text Surrogates*, *Proc JCDL 2006*, 11-20.
6. Kerne, A., Smith, S., The Information Discovery Framework, *Proc DIS 2004*, 357-360.
7. Marchionini, G., Exploratory Search: From Finding to Understanding, *Communications of the ACM*, 49(4), 2006, 41-46.
8. Pavio, A., Osapo, K., Concrete Image and Verbal Memory Codes, *Journal of Experimental Psychology*, 80(2), 279-285.
9. White, R.W., Kules, B., Drucker, S.M., Schraefel, M.C., Supporting Exploratory Search, *Communications of the ACM*, 49(4), 2006, 37-39.