A Mixed-Initiative System for Representing Collections as Compositions of Image and Text Surrogates

Andruid Kerne, Eunyee Koh, Blake Dworaczyk, J. Michael Mistrot, Steven M. Smith2, Ross Graeber, Daniel Caruso, Hyun Choi2, Andrew Webb, Pranesh Joshi
Interface Ecology Lab
Computer Science Department1 Psychology Department2
Texas A&M University, College Station, TX 77843, USA
{andruid, eunyee, blake, ross, dcaruso}@csdl.tamu.edu, mistrot@blueskystudios.com
{stevesmith, rgraebber, hyun-choi, porosis}@tamu.edu

ABSTRACT
People need to find, work with, and put together information. A wide range of activities, such as comparison shopping, entertainment, and scholarly research involve collecting information resources. A surrogate represents an information resource and enables the user to obtain it. Typical systems, such as bookmarks, iTunes, and shopping carts, represent collections as lists of textual surrogates. However, it has been shown that representations that combine text with images make better use of human cognitive facilities. Visual design principles, such as spatial organization, compositing and fading, can represent relationships among surrogates. By composition, we mean putting the elements in a collection together to form a connected whole. The set of surrogates and their mutual relationships contribute to communication and understanding. We present studies that show users experience image and text surrogate compositions as more useful and better for expressing ideas than text-only formats.

We develop combinFormation (cF), a mixed-initiative system for representing compositions as compositions of image and text surrogates. The system provides a set of direct manipulation facilities for forming, editing, and distributing compositions. Additionally, to assist users in sifting through the vast expanse of information resources that are available, the system also includes an agent which can proactively engage in the processes of collecting information resources, forming image and text surrogates, and composing them visually. We develop interactive techniques that enable the user to direct the agent from within the context of the composition space.

Categories and Subject Descriptors
H.5.1 [Multimedia Information Systems]


Keywords
mixed-initiative interfaces, collections, divergent thinking tasks

1. INTRODUCTION
Due to the popularity of digital media devices and the abundance of information on the web, a broad cross-section of society grows more and more exposed to large numbers of digital documents and media elements. People are confronted with the problem of how to keep track of significant elements within the stream of this experience. According to [16], the reason why people use the web in 69% of cases is to understand or compare/choose. The method of users in 71% of cases, is to collect, that is, to assemble information from multiple sources. Scenarios such as comparison shopping and citation collecting are prevalent and significant. In these scenarios, users need to consider the connections among diverse information resources.

combinFormation [9] is a publicly accessible mixed-initiative system that uses composition to represent information while a person is searching for, browsing, collecting, and arranging image and text clippings from web pages and other documents. The clippings act as visual, semiotic, and navigational surrogates for the documents from which they are extracted; that is, they function as enhanced bookmarks. The initiatives are the user’s direct manipulation collecting and composition, the agent’s generative collecting and composition, and the user’s direction of the agent. The agent’s generative actions – collecting clipped information elements, and composing them visually - are conducted iteratively over time, based on a model of user interests and information relationships. The agent presents the ongoing temporal generation of the composition to the user in an interactive information space. In this space, one of the user’s initiatives is to directly manipulate the composition through interactive design operations, which enable elements to be displaced, layered, resized, annotated, and removed. The user can also take initiative to direct the agent. One part of this is to express positive or negative interest in any element. Expressions of interest in the composition affect the model, creating a feedback loop (Figure 1).

This paper addresses new work completed in the course of a long term project. Prior research has addressed the design of the interest expression interface [12]; late-breaking results have addressed the capability for creating and publishing visual metadocuments [11], and the use of open-ended “divergent browsing” tasks for evaluation [13], without describing the system as a whole. This paper addresses structure of the model and the agent that collects information on behalf of users, forms image and text surrogates, and generates navigable visual composition. It also presents new user study data that demonstrates the need for representing collections as compositions. Participants were found to prefer working with collections as compositions of image+text surrogates, both when using the collections to find relevant information, and when answering open-ended scholarly questions, such as, “What kinds of things can cause behavioral problems for children in school? Put together ideas to answer this question. Develop as many varied and unusual ideas as you can.”

2. SURROGATES
Surrogates play a significant role in everyday information seeking experiences. A surrogate represents an information resource and
enables access to that resource [3]. Hypermedia surrogates, which enable navigation, are formed systematically from metadata. One typical surrogate is the Google gist, an element of the result set returned by a search query. People make critical decisions based on these surrogates, such as choosing which documents to browse, and which to ignore. Other typical surrogates include the bookmark, the ITunes playlist entry, and the TV guide entry. Surrogates play a major role in keeping found things found [8].

In combination, the basis of the surrogate structure is the inherent relationship between a clipping and its source container, and also to optional hyperlinked documents. The container is a generalization of the source document, which also includes objects such as the search query and the file system directory.

Surrogates are the medium from which result set and personal collections are constructed; they play a fundamental role in people’s processes of comparing and choosing. The significance of the representation of surrogate collections grows in importance, rivaling that of original documents. In the popular forms mentioned above, the format of the individual surrogate is a textual element, and that of the collection is a linear list. Yet, better representations are available.

### 3. IMAGE & TEXT REPRESENTATIONS

In the working memory system, the visuospatial buffer, which stores mental images, and the rehearsal loop used for text are complementary subsystems [2]. They support each other in combined Image+Text representations. Thus, it makes sense that research has shown Image+Text knowledge representations are more effective than text only. For example, cognitive research by Glenberg [6] establishes that the combination of an image and descriptive text promotes the formation of mental models [5].

Thus, combining images and text while forming surrogates makes excellent use of cognitive resources. Text disambiguates images while engaging complementary cognitive subsystems, and thus combined surrogates provide clearer navigation. Marchionini et al investigated the use of multimodal surrogates for video browsing [4]. They compared users’ performance and experience using different surrogate formats for digital videos. Combined surrogates, in which images and text reinforce each other, lead to better comprehension and reduced human processing time.

[25] investigated the efficacy of “enhanced thumbnails” as navigational surrogates for documents. They formed these thumbnails starting with a reduced screen shot of an entire web page. Each thumbnail is annotated with a larger textual “call out,” which indicates the presence of a key phrase from a search task in the result set document. Users performed significantly better on convergent thinking search tasks with enhanced thumbnails, than they did with text summaries or plain thumbnails.

We need to discover how to represent surrogates in ways that promote users’ quick understanding of the ideas inherent in information resources. Our approach to forming Image+Text surrogates is similar to enhanced thumbnails. We utilize important textual phrases that come from documents. The difference begins with our extraction of a significant image from the document for each surrogate, instead of using a thumbnail overview. The goal is to use surrogates to focus the representation of finer-grained ideas in a way that reflects the intentions of document authors, and the needs of people collecting information.

### 4. COLLECTIONS AND COMPOSITION

As evidenced by formats returned by search engines, and those utilized by web browsers for bookmarks and digital audio players such as the iPod for playlists, the list of textual surrogates is currently the format typically used to represent collections. Composition is an alternative to lists; literally, it means, “the act of putting together or combining … as parts or elements of a whole” [17]. Composition of image and text surrogates extends the organizing of information afforded by spatial hypertext [14]. Like spatial hypertext, composition includes arranging and annotating elements in an information space. Our approach differs in its emphasis on visual design and communication, as well as its attention to finding and collecting elements that function explicitly as surrogates (See example, Figure 2). We focus on the processes through which collections are assembled, and how the resulting forms function as artifacts for communication and stimuli for cognition. By composition space, we mean the composition as situated in the interactive environment in which the process of putting the collection together occurs. The use of found elements in compositional hypermedia enables the shift in emphasis to more visual representations, which are based on images as well as text, without requiring such elements to be created anew. Through mixed-initiatives, the composition serves as a basis both for an agents’ generative representation of search query result sets, and for users’ authoring of personal collections.

Composition uses visual design techniques to achieve layering effects [24]. These include relative size relationships, colors, type faces, fading, and compositing. Compositing is a means for making visible strong connections among elements. It is accomplished through the image processing technique of alpha gradients. This technique renders the border area of an image as progressively translucent. The result is a visual crossfade. Compositing contrasts with the hard edged juxtaposition of placement without blending. In combination, both the user and the generative agent can create compositing effects. Stroked text is another visual technique used to blend elements together.

### 5. MIXED-INITIATIVES

“Mixed-initiative … refers broadly to methods that explicitly support an efficient, natural interleaving of contributions by users and automated services … allowing computers to behave like associates… Achieving … fluid collaboration between users and computers requires solving difficult challenges.” – Eric Horvitz [7]

We are developing a mixed-initiative system for finding...
information resources, forming image and text surrogates, composing the information surrogates in a visual and navigational information space, and publishing resulting compositions. The part of the computer that behaves like an associate is a software agent, that is, a subsystem that engages proactively in processes of finding, forming, collecting, and composing surrogates. The agent’s actions are based on a model of the information and of the user’s interests. The goals of the mixed-initiative approach are to enable the user to act independently, and to build the agent and its interface so that the agent acts effectively on behalf of the user. Overall, we want the user to be able to successfully interleave his/her work with that of the agent, as well as with processes of directing the agent. We are developing a visual language specific to the goals of the system and the users’ tasks, in order to represent state and the possibilities for interaction.

The composition serves as a visible medium for communication between the user and the agent, as well as one in which the user collects and shares information resources. The model of user interests and information structure serves a similar intermediary function, though its form is purely computational. Figure 1 provides an architectural overview of the mixed-initiatives that the user and agent can engage in, through the composition space. In this section, we describe these initiatives and structures. In the spirit of user-centered design, we begin with the user’s direct manipulations of the underlying information collection and the visual composition. Next, we address the structure of the model, which connects user and agent, serving as the basis for agent’s actions. Subsequently, we use this specification of the model while explaining the operation of the agent’s generative information collecting and composing initiatives. Given this description of how the agent works, we complete this section by describing mechanisms we have built in order to enable the user to direct the agent.

5.1 The User’s Direct Manipulations

5.1.1 Direct Manipulation Information Collecting
Typically, when users browse the web, they utilize a bookmarks mechanism to save references to important information resources. The problem is that it becomes difficult to see meaning in the bloated lists of textual surrogates that are easily produced in this process. The composition space can be used as an alternative vehicle for collecting. In direct manipulation information collecting with combinFormation, when the user finds information that s/he wishes to collect, s/he can click and drag to select it in the source web page, then drag it over to combinFormation, and drop it into an information space. The drag and drop operation represents the material that is selected in the source web page as one or more surrogates in the composition.

Our first step in supporting this behavior was to implement support for interapplication drag and drop in combinFormation. However, this proved insufficient. The goal is not just to collect the selected information, but further, to use it construct a surrogate. In order to conduct this operation, the program needs the context of the web address of the source container document. Unfortunately, in current implementations, Java does not receive the web address of the source document in regular drag and drop operations. We developed a Firefox plug-in to pass contextual metadata through drag and drop. The plug-in annotates the HTML elements passed through drag and drop with a container attribute.

The user also needs to be able to move from the composition back to the standard web browser. The Navigate Tool (figure 3) enables the user to move from any surrogate in the composition to browse the source container from which it was extracted. If there is a hyperlink, navigation to that container is also enabled by this tool. The two destinations can be toggled with the shift key. Changes in the cursor make the state change visible.

5.1.2 Direct Manipulation Visual Composition
The combinFormation interface provides the user with a set of visual composition capabilities. Some are activated like Photoshop tools (Figure 3, toolbar left). These include removing unwanted surrogates, spatial positioning, and text editing. Other capabilities are activated on mouseover (Figure 5). These include adjusting the size of text and images, changing the stroke color and font of text elements, and compositing images by creating border regions with translucent gradients.

Figure 3. Design/Expression Toolbar + Cursor. On toolbar left, the design+ tools, clockwise, are grab, cut, navigate, and text edit. On toolbar right, the interest expression possibilities are positive, neutral, and negative. Cursor = positive grab.

5.2 Model
The system processes the containers specified as seeds, and those discovered by the agent. The currently supported container formats are HTML, PDF, RSS, and file system directory. Each container is translated into a set of text chunk elements (Text chunks are delimited by markup such as table cells, blank lines, and the period character that ends sentences.), image element references, and hyperlink references (In the case of a file system directory, there are only “hyperlink” references to other containers, but no text chunk or image reference elements.). These are the information elements of the model. The association of these elements with their source container document is maintained. This is the basis of the hypermedia graph model, which is one of two fundamental model structures the system maintains in order to assign floating point significance weights to the elements. Processed elements that are unique are added to pools of candidate elements, which are then used by the agent in its generative operations (see section 5.3). In such cases, we consider these text chunk and image element reference elements as surrogate candidates. The surrogate candidates are associated with properties of their containers and hyperlinks, as well as maintaining their more specific property sets. Weights are computed such that the more hyperlinks have been traversed, the lower the resulting weight, which favors breadth-first traversal.

5.2.1 Modeling Metadata
A set of metadata fields is associated with each information element. The underlying metadata system is extensible. It associates types with field labels. Currently supported types include colors, numbers, and URLs, in addition to text strings.

There are several possible sources for metadata fields. The RSS container type associates metadata fields directly with hyperlinked documents. Likewise, HTML documents themselves may contain metadata at two levels. One is the head element, which may include meta elements, such as description, as well as a title element; these become associated with the container. Also, the alt attribute of the img element, when present, is translated into a “caption” metadata field, which becomes associated with the image element reference. Valuable contributions to an ‘anchor’ metadata field for a container are derived when the anchors of hyperlinks to that container are encountered, as per [18].

5.2.2 Information Retrieval Model
Along with the hypermedia graph, the other fundamental model structure in combinFormation is the vector space model of information retrieval (IR) [21], which connects information elements by common terms. A composite term vector is formed for each information element. An inverted index, which associates a set of information elements with each term object, is also formed, with entries for each information element that refers to the term. For each information element, the composite term vector and inverted index entries are formed through the union of the associated metadata fields. Additionally, for text chunk elements, the text itself becomes part of the composite term vector. The associated words are stemmed [20] and added into the composite term vector, except for stop words. Our stop word list includes usual terms, such as ‘a’ and ‘the’, and special web stop words, such as ‘adv’, ‘click’, and ‘e-mail’.

Dynamically constructed term vectors are supplemented by a pre-built term dictionary, which contains frequency counts for the set of terms discovered in 6000 random web pages. This enables the computation of significance weights using inverse document frequency (IDF) statistics [21]. The dictionary is enlarged as the program operates and discovers new terms. However, the discovery of terms in the course of a session does not contribute to IDF. The reason for this is that it would penalize the agent’s success in dynamically discovering relevant documents. Also, in case there is no explicit metadata for a container or image element reference, the system will attempt to mine terms from the URL. To reduce the occurrence of junk term associations, which can interfere with the operation of the model, only terms found in the dictionary will be added to the mined keywords field.

5.2.3 Modeling User Interests
A participant object is associated with each information element
and each term. The set of these forms a profile of the user’s interests. The interface for interest expression is defined in section 5.4.1, below. Interest level is modeled as an integer on [-10, +10].

When the user expresses interest in a surrogate, this expression is propagated by spreading activation [19] to semantically related nodes. Hyperlinked documents represent one form of semantic relationship through which activation is spread. Likewise, the term objects of the IR model refer to related surrogates and hyperlinked containers (through anchor texts) that contain the same terms, and so receive spreading activation. These expressions of interest contribute to the weighting calculations the agent utilizes in its generative actions.

5.3 The Agent’s Generative Threads
The model is used to drive decision-making in several generative threads of execution, which together can be said to comprise the agent that executes the system’s initiatives. Two threads direct generative information collecting, while two others effect generative visual composition of surrogates. When the user enables their operation, these threads will run gradually over time, generating evolving state within the model and the composition. They utilize the candidate pools and weighting measures of the model. The measures themselves depend on statistics such as IDF. In the current implementation, all selection operations choose the maximum, given the weights of elements in a candidate pool. In case of a tie, an element is chosen randomly. A prior system [10] used weighted random select in order to create indeterminacy. However, we have found that the variability of network download times and the changing structure of the web interject sufficient variability. Thus, the agent’s operations are stochastic. The resulting experience is consistent with Amar and Stasko’s recent call for information visualization systems that incorporate uncertainty and respond to change [1].

5.3.1 Generative Information Collecting
The agent engages in generative information collecting to collect information resources that are relevant to the user, and form surrogates that represent these resources. One of the agent threads is a web crawler, which periodically chooses a candidate hyperlink, and initiates the download of the associated document, and processes it as above (Figure 5, left). Another thread directs the downloading of images (Figure 5, right). Neither of these threads actually performs the downloads. For purposes of robust software engineering, it was necessary to insulate the agent threads from the vagaries of network I/O. Instead, these threads request the downloads, which are, in turn, performed by a pool of threads dedicated to downloading operations. The software objects that manage these threads have been engineered to handle I/O errors gracefully. The result of generative information collecting is to transform a set of seeds into evolving pools of candidate text surrogates, candidate image surrogates, and candidate containers. Note that the sleep times between iterations of these threads are not actually fixed. They change adaptively, based on factors such as how large the candidate pools are, and how many elements are already queued awaiting download.

5.3.2 Generative Visual Composition (Temporal)
Like generative information collecting, the combinFormation agent’s generative visual composition of the information space is not performed and presented all at once. Rather, it develops gradually over time. This mechanism for automatic temporal layout uses time as a continuous dimension.

The primary thread of the visual composition agent (Figure 6) iteratively selects surrogate candidates for placement in the composition. Next, the state of each element already in the composition, which the user has not already expressed interest in (section 5.4.1) is aged, to gradually reduce its weight. This thread conducts a cycle of further steps to generate the layout. Iterations through the cycle are typically 1 second apart, though the user can change the rate or pause the process (section 5.4.2). Through a series of such cycles, the layout emerges.

To support the placement algorithm, the visual workspace in combinFormation is divided into a matrix of rectangular cells (Figure 7). Each cell is aware of the information elements that substantially overlap it. A weight is assigned to the cell, which is simply the weight of the element that is currently on top within the cell. A size, in grid cells, is assigned to the new element based on its importance relative to those already in the composition. Based on this size, we can establish a set of candidate locations in which to perhaps place the new element. With each such location is associated a region of grid cells, known as the candidate region. A weight is assigned to each candidate region by simply integrating, or adding, the weights of its constituent grid cells. From these calculations, we derive a set of candidate positions for placement, and associated grid region candidate weights. The new element is placed at the location of the minimum weight, so that it covers the region of least importance.

At the end of this process, the elements in the space are sorted, based by weight. They are then restacked so the most important elements are on top. If the system considers the space to be full, based on a preset threshold of element density, the least important element will also be removed from the information space as part of each cycle of generative visual information composition.

![Figure 5. Generative Information Collecting threads flowchart.](image)

![Figure 6. Generative Visual Composition building thread.](image)
5.4 The User Directing the Agent

As the Shneiderman - Maes debates made clear, the promise of agents that assist the user is tenuous [22]. It depends on interactive mechanisms that enable the user to effectively affect the agent’s actions. Further, studies have shown that depending on the state of the task at hand, the user may need to turn the agent off, and engage solely in self-directed composition of surrogates [12]. Thus, we develop the role of surrogates and direct manipulation in the composition space as means to direct the agent.

5.4.1 Interest Expression

A primary means of directing the agent is through the interest expression interface (Figure 3, toolbar right) [20]. The user can adjust her profile of interests through an interactive interface. The user selects a positive, neutral or negative interest expression setting. This setting is applied during subsequent design+ operations (Figure 3, left). By expressing interest in a surrogate, the user provides relevance feedback, which effectively edits her interest profile of “rankings.” In order to facilitate this expression, no dialog box or other cognitive context switching is utilized. Providing feedback is never required, and always possible. This is our solution to the problem of elicitation of user interests [15]. Thus, in interest expression interactions, the surrogate clipping serves as an affordance for relevance feedback. In the course of a 21 minute authoring session, combinFormation users were found to conduct 92 interest level operations, in addition to 202 design+ operations. We interpret this result to demonstrate that users are able to express interest successfully, and motivated to do so.

5.4.2 Affecting the Agent’s Flows of Control

A tape recorder transport (Figure 8) enables pausing the agent’s process of generating the visual composition, and altering the rate of temporal layout generation. Menu entries enable the user to pause the web crawler that follows hyperlinks to download documents, and also the thread that utilizes references to image locations to download them and form surrogate candidates for possible inclusion by the agent in the composition.

5.4.3 User-controlled Subspaces

User feedback made it clear that in addition to the weighting system, users want more direct control of parts of the composition space [12]. They don’t want to share all of it with the agent. In response to this, we created two structural mechanisms to give them complete control of subspaces of the composition: the cool space and the latch.

A resizable cool region of the information space, typically located in the center, may be reserved for the user’s composition actions. During the course of the session, as s/he finds and composes relevant surrogates, the user may choose to enlarge the cool space. In this way, the agent’s composition actions may function as peripheral suggestions to the initiatives of the user’s central direct manipulation processes of collecting and composing.

The latch tool (Figure 4, right) enables the user to create a single-element floating cool region. The agent will not remove surrogates latched by the user from the composition.

5.5 Visual Metadocuments

Compositions serve as a medium for exchanging collections. They may be saved and published on the web. Saving produces an XML format, which can be re-opened with combinFormation, and a dynamic HTML format, which can be opened in a regular web browser. Both formats can be published on the web and exchanged with colleagues and students. The DHTML version is visually identical to the full combinFormation composition, and provides similar metadata details on demand. It includes a link at the bottom, which opens the XML in combinFormation. In the XML, text and metadata are saved by value. Images are saved by reference. Specifications of the model are also saved.

6. USER STUDIES

We report on two user studies, designed to determine the effectiveness of the composition of image and text surrogates as a format for navigational collections. For both studies, a set of collections of information resources representing the undergraduate psychology curriculum was authored. Equivalent surrogate collections were authored in three formats.

We developed a controlled study by defining tasks to match the cognitive processes found in real world scenarios. Prior studies of navigational surrogates cases [25] have utilized convergent thinking tasks. Convergent thinking tasks use well-defined problems with unique correct answers. Divergent thinking tasks are less well-defined, and demand a variety of creative responses [23]. By divergent browsing tasks, we mean divergent thinking tasks which require browsing to assemble information resources from multiple diverse sources in order to form a set of answers. Although convergent tasks are easier to construct for experiments, and their solutions are easier to score, divergent browsing tasks better exemplify the cognitive processes of common web usage scenarios that involve comparison of resources, and collecting.

6.1 Using the Composition for Navigation

6.1.1 Experimental Approach

In the first study, rather than making omnibus comparisons of systems for collection and navigation, we used an experimental approach to ask questions about navigational surrogate representations, examining the efficacy of specific dimensions of the surrogates [13]. We constructed the user tasks for these experiments to use fixed collections of information resources in
order to reduce the variability of the factors in the experiment. Two dimensions, the representational format of each surrogate (text vs. image with text), and the spatial layout of the presented navigational collection (serial lists vs. spatially arranged clusters), were singled out and examined. Both dimensions were hypothesized to be relevant to the user's cognitive abilities, such that spatial clusters would augment the formation of "chunks" in working memory [2], and images with text would augment rapid formation of mental models of the information resources linked to the surrogates [1][6]. In the second study, we added an additional dimension: the format of the answers users’ developed. Here we compared the usual textual list of answers with the composition of image and text surrogates, authored in combinFormation.

Prior studies have used convergent search tasks over wide, unclassified sets of web pages [25]. To isolate navigational surrogate format from the effects of dynamic components, such as search engines, the present study focuses on browsing, rather than searching. To integrate these factors, we authored a specific collection for this study, and developed equivalent versions in each of the 3 navigational surrogate formats. We designed divergent browsing tasks specifically to work with this collection.

6.1.2 A Psychology Resources Collection
We used our university’s psychology pool as a source of subjects for the experiment. These subjects are undergraduates in the introductory psychology class, who fulfill a course requirement for learning about what a psychology study is by participating in one as a subject. To maximize the educational value of our collection-authoring efforts, we assembled information resources based on the psychology curriculum. We intend to make this collection available to the community, as a pedagogical resource.

We divided the psychology curriculum into 12 topics. We collected information resources for 6 of these topics: biopsychology, learning, developmental, consciousness, clinical psychology, and perception. Resources were gathered from diverse websites, such as governmental agencies, medical centers, research institutes, internet journals, and universities.

We created equivalent surrogates and authored equivalent collections of psychology information resource surrogates, using each of three formats: linear text, spatial text, and composition of images and text. We used combinFormation to author the collections. We started with the composition format. Surrogates were positioned within the information space based on mutual interconnections. Sizes of elements and colors for text stroking were chosen to maximize layering and legibility [24]. Some images were processed with translucent border areas in order to visually represent strong interconnections. We utilized combinFormation’s DHTML output format to derive a simplest product to use in the study. We added to combinFormation a mechanism for automatically deriving spatialized text and linear text output from an image+text composition.

Each collection begins with a Psychology Overview, consisting of surrogates for each of the 6 topics (Figure 2). Navigation from here leads to the topic-specific collections, each of which contains an additional 6-10 surrogates. Clicking within each topic collection leads to the information resources, themselves.

6.1.3 Experimental Method
There were three divergent browsing questions (e.g., “What kinds of things can cause behavioral problems for children in school?”), which were manipulated within-subjects. Each question was designed to be answered by navigating the collection of psychology resources. The instructions for each question stated, “Gather all of the possible facts relevant to this topic from the websites provided, and list the facts. For each fact write a brief description (e.g., one sentence) of its relevance to the topic. List as many varied and unusual facts as you can.”

Each subject received one divergent question in each format. The order of the 3 questions and the 3 formats of navigational surrogates was counterbalanced between subjects, thereby producing 9 different combinations of question order and surrogate formats. Each participant was assigned to one of these counterbalancing conditions.

52 introductory psychology students were recruited. Each session was held in a group of 5-15 participants at a time. Participants were asked to find answers for questions chosen from various areas of psychology, by browsing websites that include possible answers. They were told to give as many answers as they could find using the time provided. The experiment consisted of three parts: pre-experimental questions, the three divergent questions, and post-experimental questions. Pressing a start button led to a framed webpage in which each question was presented in the top frame. There was also an “Overview” button in the upper right hand corner, for browsing the Psychology Overview in the bottom frame. Participants were given 8 minutes for each divergent question. They could see the time remaining counting down on a digital clock. At the end of 8 minutes, if a participant had not clicked “submit,” the program automatically submitted what they had written, and the next question was displayed. After all 3 divergent questions were answered, participants answered post-experimental questions about their preference among the 3 navigational formats, and the reasons for their choices. The whole procedure lasted 40 minutes. To avoid the variability of Internet download times, resource websites were mirrored locally.

6.1.4 Results
We collected subjective measures of user experience in relation to the 3 surrogate collection formats, liking, ease of use, and utility (Figure 9). Participants clearly liked the composition of image and text surrogates best [$\chi^2(2) = 34.88, p < .001$]; over 70% preferred that format. Linear texts were preferred second most, and spatial text surrogates were least preferred. Further, participants found the composition format easiest to use of the three [$\chi^2 (2) = 16.12, p < .001$], with linear texts second and
6.2 Creating Composition

We realized that for the divergent thinking questions, the answers that users create also form a collection, and that when one thinks of citations as part of the product of the collecting process, that the composition of surrogates is a sensible format for answering divergent thinking questions, as well as for providing pedagogical materials. In a follow-up study, we added the direct manipulation composition format for answers. We reduced the input format choices to linear text and composition of images and text, in order to reduce the number of subjects needed.

We increased the running time of the experiment to 80 minutes. We included a brief combinFormation training and a warm-up exercise, in which each participant created a composition using direct manipulation to collect information from popular media web sites, such as People Magazine, Sports Illustrated, and the Metropolitan Museum of Art. Each participant was then exposed to 4 experimental conditions, varying the 2 input vs. the 2 output formats, using questions similar to those in experiment 1.

The study has produced statistically significant results on user experience measures for output format with 11 participants. 91% of participants like the direct manipulation composition output format better than linear text [paired t-test: F(1,10)=4.5, p < 0.001]. 91% also think this format is best for expressing their ideas. 72% find the composition of image and text surrogates format to be most useful [F(1,10)=1.614, p < 0.138].

7. CONCLUSION

Users experience composition of image and text surrogates as a more effective means for representing collections of surrogates. Users were also found to like to manipulate the visual form of their collections. Compositions of image and text surrogates could prove to be particularly beneficial in information discovery tasks, in which the user’s goal is not just to find relevant information, but to synthesize found information, and develop new insights about relationships and applications. Further work is needed on systems to support this, and processes and methods for evaluation.

8. ACKNOWLEDGMENTS

Support has been provided by NSF grant IIS-0429469.

9. REFERENCES