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

Andruid Kerne, Eunyee Koh, Blake Dworaczyk, J. Michael Mistrot, Hyun Choi[2],
Steven M. Smith[2], Ross Graeber, Daniel Caruso, Andrew Webb, Rodney Hill[3], Joel Albea[3]

Interface Ecology Lab

Computer Science Department[1] Psychology Department[2] Architecture Department[3]

Texas A&M University, College Station, TX 77843, USA

{andruid, eunyee, blake, ross, dcaruso}@csdl.tamu.edu, mistrot@blueskystudios.com

{stevesmith, rgraeber, hyun-choi, porosis, rhill, jalbea}@tamu.edu

ABSTRACT

People need to find, work with, and put together information. Diverse activities, such as scholarly research, comparison shopping, and entertainment involve collecting and connecting information resources. We need to represent collections in ways that promote understanding of individual information resources and also their relationships. Representing individual resources with images as well as text makes good use of human cognitive facilities. *Composition*, an alternative to lists, means putting representations of elements in a collection together using design principles to form a connected whole.

We develop combinFormation, a mixed-initiative system for representing collections as compositions of image and text surrogates. The system provides a set of direct manipulation facilities for forming, editing, organizing, and distributing collections as compositions. Additionally, to assist users in sifting through the vast expanse of potentially relevant information resources, the system also includes a generative agent that can proactively engage in processes of collecting information resources and forming image and text surrogates. A generative temporal visual composition agent develops the collection and its visual representation over time, enabling users to see more possibilities. To keep the user in control, we develop interactive techniques that enable the user to direct the agent.

For evaluation, we conducted a field study in an undergraduate general education course offered in the architecture department. Alternating groups of students used combinFormation as an aid in preparing one of two major assignments involving information discovery to support processes of invention. The students that used combinFormation were found to perform better.

Categories and Subject Descriptors

H.5.1 [**Multimedia Information Systems**]

General Terms: Human Factors, Design, Experimentation.

Keywords

mixed-initiative systems, collections, information discovery

JCDL'06, June 11–15, 2006, Chapel Hill, North Carolina, USA.

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 [23], 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 prior work investigation and comparison shopping are prevalent and significant. In *information discovery* tasks, finding relevant information resources is one part of the process [16]. Users also need to develop understanding of the connections among many diverse information resources. Thus, information discovery involves the emergence of new perspectives and new ideas in the context of the stimulus of found information. The present research develops new methods for supporting processes of information discovery, as well as new methods of validation.

We are developing a mixed-initiative system, combinFormation (cF) [13], 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 (See example, figure 1.). The clippings act as visual, semiotic, and navigational surrogates for the documents from which they are extracted; that is, they function as enhanced bookmarks. We need to represent surrogates in ways that promote users' quick understanding of the ideas inherent in information resources. A *mixed-initiative system* is one in which the actions of the user and an agent, working on a joint task, are interleaved [11].

In combinFormation, the user and agent collaborate to find relevant information resources, form image and text surrogates, and compose the information surrogates in a visual and navigational interactive space. 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 – clipping and collecting information elements to form surrogates, and composing them visually - are conducted iteratively over time, based on a semantic model of the surrogates, their relationships, and the user's interests. One of the user's initiatives is to directly manipulate the composition through interactive design operations, which enable surrogates to be displaced, layered, resized, annotated, and removed. The user can also take initiative to direct the agent by using interactive tools; this includes expressing positive or negative interest in any surrogate. Expressions of interest in surrogates affect

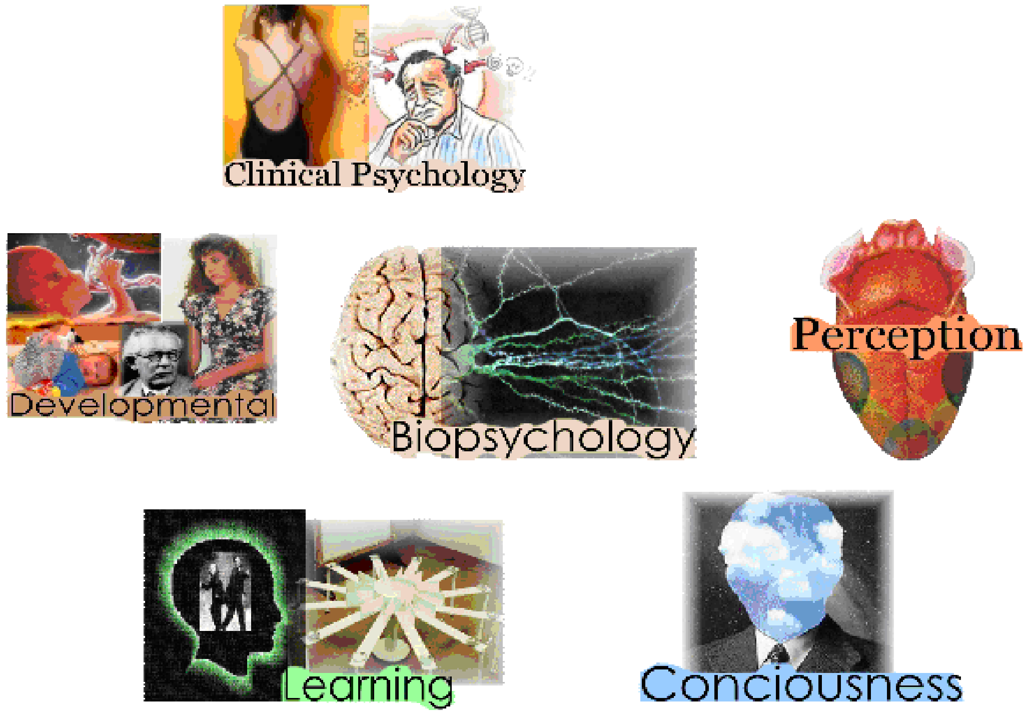

Figure 1. A composition of image and text surrogates representing sets of information resources for an overview of part of the undergraduate psychology curriculum. Each surrogate is formed by clipping information elements from source containers.

the semantic model, creating a relevance feedback loop through the composition (Figure 2).

This paper addresses new work completed in the course of a long term project. The focus is on how the mixed-initiative system and its agent initiatives work. A new field study is also presented. Prior research has addressed the design of the interest expression interface [17]; late-breaking results have addressed the capability for creating and publishing visual metadocuments [15], and the use of open-ended "divergent browsing" tasks for evaluation of pre-authored compositions of image and text navigational surrogates [18], without describing the mixed-initiative system as a whole, or investigating its use in practice.

We begin with a background section, which motivates our approach to representing collections, by drawing from the digital libraries, cognitive psychology, human computer interaction, and visual design literatures. Next, we develop combinFormation's mixed-initiative structure. Here, we begin by describing the user's direct manipulation initiatives. Then, we present the semantic model and the agent initiatives that collect information on behalf of users, form image and text surrogates, and generate navigable visual composition. This portion of the paper closes by describing interactive methods that enable the user to direct the agent. The next section presents our field study, which was conducted in an undergraduate course on The Design Process. Finally, we discuss the implications of our findings, and consider differences and connections between information discovery tasks and those that emphasize retrieval, foraging, and seeking.

2. BACKGROUND

2.1 Surrogates

Surrogates play a significant role in everyday information seeking experiences. A surrogate represents an information resource and enables access to that resource [4]. 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 [12].

In combinFormation, 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. Examples of containers includes information resources such as web pages, PDF documents, search query result sets, RSS feeds, and file system directories. The container is a composite, in the sense that it is made up of smaller atomic entities, such as images, sentences of text, hyperlinks, and metadata fields, which can be extracted.

During searches and recall of resources from personal collections, surrogates form the basis of our decisions about which documents to browse and which to pass by. Surrogates can also help us to think about the relationships among the significant ideas in information resources. Thus, they play a fundamental role in people's processes of comparing and choosing. *The significance of the representation of surrogate collections grows in importance over time, rivaling that of original documents.* In the popular forms mentioned above, the typical format of the individual surrogate is a textual element, and that of the collection is a linear list. Yet, better representations are available.

2.2 Image & Text Representations

In the working memory system, the visuospatial buffer, which stores mental images, and the rehearsal loop used for words have been shown to function as complementary subsystems [2]. They support each other in combined image and text representations. Likewise, research has shown image and text knowledge representations are more effective than text only. For example, cognitive research by Glenberg [10] establishes that the combination of an image and descriptive text promotes the formation of mental models [9].

Thus, combining images and text while forming surrogates makes excellent use of cognitive resources. Since text disambiguates images while engaging complementary cognitive subsystems, combined surrogates provide clearer navigation. Ding *et al* investigated the use of multimodal surrogates for video browsing [7]. 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.

Woodruff et al investigated the efficacy of "enhanced thumbnails" as navigational surrogates for documents [33]. 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 represent surrogates in ways that promote users' quick understanding of the ideas inherent in information resources. Our approach to forming image and 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.

2.3 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" [24]. Composition of image and text surrogates extends the organizing of information afforded by spatial hypertext [21]. 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. Figure 1 shows a composition that represents areas of the undergraduate psychology curriculum. This research addresses the processes through which collections are assembled, and how the resulting forms function as artifacts for communication and navigation, and stimuli for cognition. By *composition space*, we mean the interactive environment in which the process of putting the composition together occurs. The use of collected elements in compositional hypermedia enables the shift to more visual representations, based on images as well as text, without requiring these surrogates to be created anew. Through mixed-initiatives, the composition space serves as a basis both for the agent's generative representation of search query result sets and surrounding information, and for users' authoring of personal collections.

Composition uses visual design techniques, layering and blending elements, to form a coherent whole [32]. In combinFormation, these design techniques include relative size relationships, colors, type faces, text stroking, and image compositing. Stroking dark text with a thin light background guarantees contrast with overlapped elements, while maximizing how much of what is underneath can read through. 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 combinFormation, both the user and the generative agent can create compositing effects.

3. 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 [11]

In our mixed-initiative system, 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 relevant surrogates. The agent's actions are based on a semantic 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

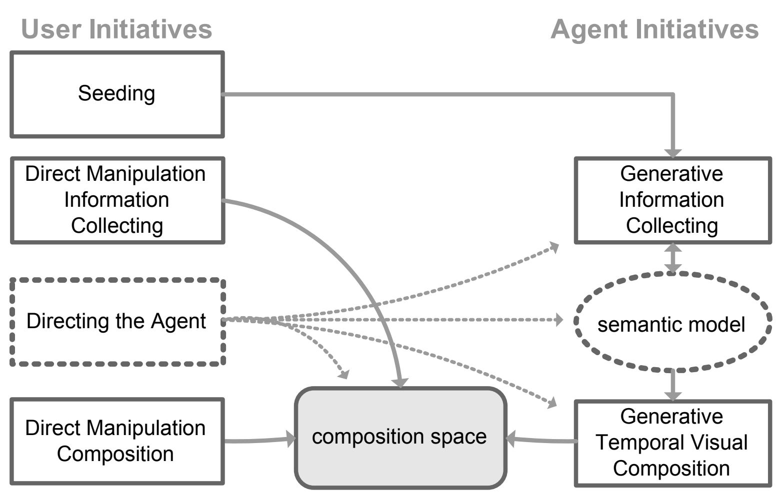

Figure 2: Mixed initiative processes in combinFormation.

collects and shares information resources. The semantic model of user interests and information structure serves a similar intermediary function, but in a purely computational form.

Figure 2 provides an architectural process overview of the mixed-initiatives that the user and agent can engage in, through the composition space. It shows their relationships, how they affect each other, the composition, and the semantic model. Through seeding, the user points the agent at particular information sources. Through direct manipulation information collecting, s/he brings surrogates, and their underlying semantics, directly into the composition space and the model. Through direct manipulation composition, the user changes how the composition looks, in order to facilitate her own understanding of the information resources and their connections, and perhaps to communicate such understanding to others. By directing the agent, the user can turn the agent initiatives off and on, and also provide relevance feedback through interest expression that affects the semantic model. The semantic model, in turn, serves as the basis for the generative agent's initiatives of collecting information and temporal visual composition.

3.1 User Initiatives

3.1.1 Launch / Seeding

The user can launch combinFormation in one of several ways [13]. In *re:open*, s/he begins with an empty composition space, which can be used to initiate direct manipulation information collecting or to reopen a saved composition for browsing and further mixed-initiative collecting and composing. Through another launch method, *re:mix*, the user can put together any number of *seeds* for the agent. These are specifications for searches and web sites that the agent uses as starting points for information collecting. In the *re:collection*, method, the user can select a set of seeds that has been curated by members of our team. An example is the "news collage," which puts together material from news feeds such as CNN, the New York Times, The Guardian, and Al Jazeera.

3.1.2 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 through this method. 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 standard web browser, then drag it over to combinFormation, and drop it into the composition 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.

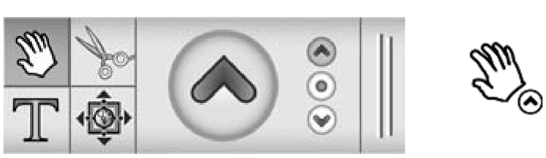

Figure 3. Design/Expression Toolbar + Cursor. On toolbar left, the design+ tools, clockwise, are grab, cut, navigate, and text edit. On toolbar right, the interest expression possibilities are positive, neutral, and negative. Cursor = positive grab.

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 to 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.

3.1.3 Direct Manipulation Visual Composition

The combinFormation interface provides the user with a set of visual composition capabilities. Some are activated like modal Photoshop tools (Figure 3, toolbar left). These include removing unwanted surrogates, spatial positioning, and text editing. Additionally an edit palette is activated on mouseover (Figure. 4). This enables adjusting the size of text and images, changing the stroke color and font of text surrogates, and compositing images.

3.2 Semantic Model

An information collection agent forms information elements from text chunks and images as documents are processed. Each element is associated with its source container documents and hyperlinks, to form the hypermedia graph. As it discovers such elements, the agent may add them to pools of candidate elements, which it then uses in its generative operations (figure 5). Thus, these information elements function as *surrogate candidates*. Those that are selected for visual inclusion in the composition space function explicitly as surrogates.

The semantic model consists of a structural representation of the relationships between documents, the components of documents that form surrogate candidates, associated metadata, and citation

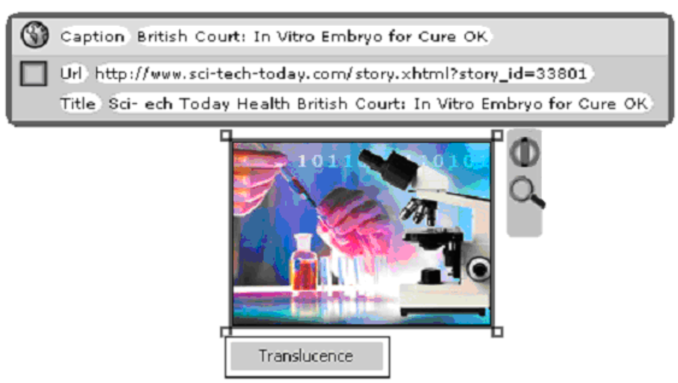

Figure 4: *In-context metadata details on demand* above an image surrogate. We see fields for the image (top row) and for its container (rows 2, 3). In this case, there is no hyperlink. In-context tools to the right afford latching and search. The edit palette, below, enables editing of visual attributes, such as translucence in the case of image surrogates.

links. Semantic models of relationships form the basis for spreading activation algorithms that utilize user expressions of interest (i.e. relevance feedback [28]). Metrics that utilize features and relationships of the semantic model compute weights that drive the agent initiatives that act on behalf of the user. Agent initiatives employ the metrics in algorithms that select new surrogates to present in the composition space from pools of surrogate candidates, select documents to crawl, and select layout positions of the new surrogates as they are added to the temporal visual composition.

3.2.1 Modeling Metadata

A set of metadata fields is associated with each surrogate candidate. 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. Metadata can be acquired from semantic web, HTML markup, and digital library sources.

When the user places her/his mouse over a surrogate within the composition, *in-context metadata details on demand* are displayed (Figure 4). The metadata details are presented directly above or below the surrogate they describe, rather than in a separate dialogue box, or in reserved screen real estate in the peripherae of the display, in order to make the best use of human cognitive attention and of screen real estate. The in-context metadata details can be edited in place by the user, through simple click and type. *In-context tools* are proximally presented at the same time. These include the latch (see Section 3.4.3), and a search tool. The user can invoke this search tool to generate a new query to Google, using the surrogate's composite metadata.

3.2.2 Information Retrieval Model

The primary model structure in combinFormation is the vector space model of information retrieval (IR) [29], which connects surrogate candidates (and visible surrogates) by common terms. A composite term vector is formed for each surrogate candidate. An inverted index, which associates a set of surrogate candidates and surrogates with each term object, is also formed, with entries for each element that refers to the term. For each surrogate candidate and surrogate, the composite term vector and inverted index entries are formed through the union of the associated metadata fields. Additionally, for text surrogate candidates and surrogates, the text itself becomes part of the composite term vector. The associated words are stemmed [27] 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 [29]. 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 interfere with the operation of the semantic model, only terms found in the dictionary are added to the *mined keywords* field.

3.2.3 Modeling User Interests

A participant object is associated with each surrogate candidate and surrogate, as well as with each term and each container. The set of these forms a profile of the user's interests. Interest level is modeled as an integer value on [-10, +10]. The interface for interest expression is described in section 3.4.1, below.

When the user expresses interest in a surrogate, this expression is propagated through the model into the appropriate participant objects by spreading activation [26] 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 the participant objects of 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 metrics the agent utilizes in its generative actions.

3.3 The Agent's Generative Initiatives

The semantic model is used to drive decision-making in several generative threads of execution. This set of software components comprises the agent initiatives. Two threads perform generative information collecting, while two others perform generative temporal visual composition of surrogates. When the user enables their operation, these threads run gradually over time, generating evolving state within the semantic model and the composition. They utilize the candidate pools and weighting metrics of the model. The metrics themselves utilize 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 [14] used weighted random select in order to create indeterminacy. However, it was found that the variability of network download times and the changing structure of the web interject sufficient variability. Thus, even using maximum select, the agent's operations are still stochastic, rather than purely deterministic. The resulting experience is consistent with Amar and Stasko's recent call for information visualization systems that incorporate uncertainty and respond to change [1].

3.3.1 Generative Information Collecting

The generative information collecting initiative identifies information resources that are relevant to the user, and forms surrogate candidates that represent these resources. One of this

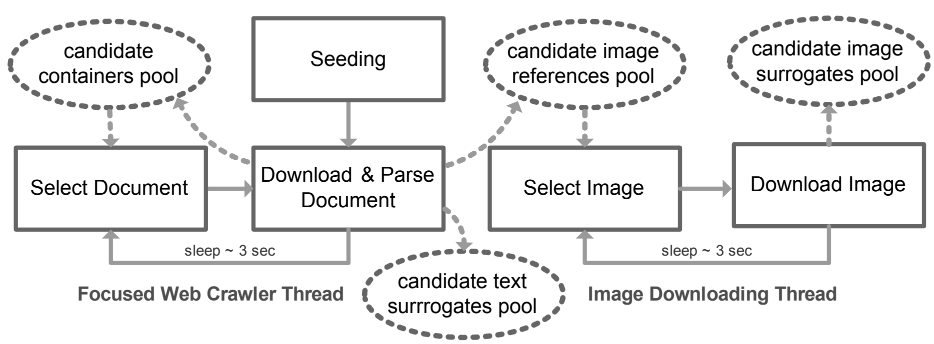

Figure 5. Generative Information Collecting agent initiative threads flowchart.

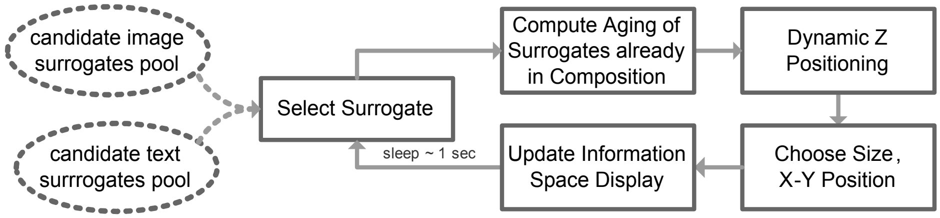

Figure 6. Generative Temporal Visual Composition: the composition building thread.

initiative's agent threads is a focused web crawler [6], which periodically chooses a candidate container, 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 uncertainties of network I/O. Thus, these agent initiative threads request the downloads, which are, in turn, performed by a pool of other threads dedicated to downloading operations. The software objects that manage these threads have been engineered to handle I//O errors gracefully. The generative information collecting initiative processes the pool of candidate containers to evolve the states of the 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, in response to factors such as network response time, and already achieved rate of success in retrieving relevant information. For example, when the agent seems to have already collected a large supply of relevant surrogate candidates, the rate of crawling will be reduced.

3.3.2 Generative Temporal Visual Composition

Like generative information collecting, the combinFormation agent's generative visual composition is not performed and presented all at once. Rather, it develops gradually over time. This mechanism for automatic layout uses time as a continuous dimension. In this way, *temporal visual composition* is a time-based visual medium, like video.

The primary thread of the visual composition agent initiative (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 3.4.1) 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 3.4.2). Through a series of such cycles, the layout emerges.

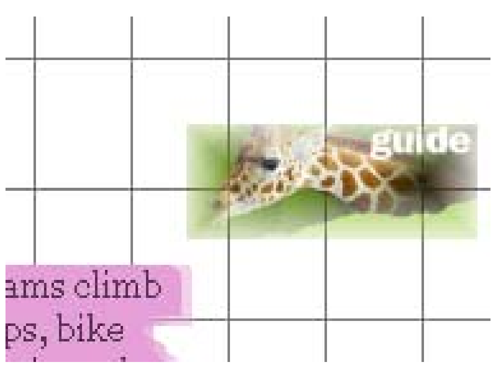

Figure 7. The composition space is divided into a matrix of cells, with each cell keeping track of the surrogates that are placed upon it.

To support the placement algorithm, the visual workspace in combinFormation is (internally) divided into a matrix of rectangular cells (Figure 7). Each cell is aware of the surrogates that substantially overlap it. A weight is assigned to the cell, which is simply the weight of the surrogate that is already on top within the cell. A size, in grid cells, is assigned to the new surrogate based on its importance relative to those already in the composition. Based on this size, the agent can establish a set of candidate locations in which to perhaps place the new surrogate. 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, the agent derives a set of candidate positions for placement, and associated grid region candidate weights. The new surrogate 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 surrogates in the space are sorted, based by weight. They are then restacked so the most important surrogates are on top. If the system considers the space to be full, based on a preset threshold of surrogate density, the least important surrogate will also be removed from the composition space as part of each cycle of generative temporal visual composition.

3.4 The User Directing the Agent

As the Shneiderman - Maes debates made clear, the promise of agents that assist the user is tenuous [30]. 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 [17]. Thus, we develop the role of surrogates and direct manipulation in the composition space as means to direct the agent.

3.4.1 Interest Expression

A primary means of directing the agent is through the interest expression interface [17], which affects a profile of interests. Through this interface, the user can activate up and down arrows (See Figure 3, toolbar right.), which signify the intent to increase or decease interest level values in the participant objects associated with a surrogate. After the user has selected a positive, neutral or negative interest expression setting, the cursor changes to make the now activated interest expression setting visible. This setting is applied in combination with subsequent design and navigation operations (Figure 3, left), until the setting is changed again in the toolbar (or via up and down arrow keyboard accelerators).

By expressing interest in a surrogate, the user provides relevance feedback, which effectively edits her interest profile of "rankings" in the semantic model. In order to facilitate this expression, no dialog box or other cognitive context switching is imposed on the user experience. Providing feedback is never required, and always possible. This is our solution to the problem of elicitation of user interests [22]. 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 and navigation operations. We interpret this result to demonstrate that users are able to express interest successfully, and motivated to do so.

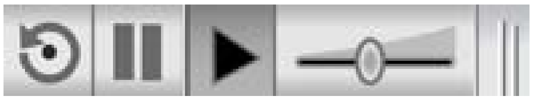

Figure 8. Tape recorder transport controls the agent's generation of the visual composition.

3.4.2 Affecting the Agent's Flows of Control

Since the temporal visual composition agent initiative, when it runs, is continuously changing the look and feel of the composition space, it is important to give the user direct control over this process. A tape recorder transport (Figure 8) enables pausing the agent's process of generative temporal visual composition. A slider in the same floating window enables adjusting the rate of this process.

There are controls for the other agent initiatives, as well. 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.

3.4.3 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 [17]. 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 composition space, typically located in the center, may be reserved for the user's composition actions (See example, Figure 9.). This area is off-limits for the agent to use during generative temporal visual composition. During the course of a 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* is an in-context tool (Figure 4, right) that enables the user to create the equivalent of a single-surrogate floating cool region. The generative temporal composition agent will not remove surrogates latched by the user from the composition, or cover them with new surrogates.

3.5 Visual Metadocuments

Compositions serve as a medium for exchanging collections. They may be saved and published on the web. Saving produces a file in XML format, which can be reopened with combinFormation, and also a file in 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 in-context 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. Elements of the semantic model are also saved.

4. VALIDATION

4.1 Information Discovery Tasks

The tasks used to evaluate interactive systems must match the tasks, processes, and needs that correspond to real world usage scenarios. Convergent thinking tasks involve questions that have a single correct answer. A problem is very explicitly specified, and the criteria for the sought-for solution are very clear [7] [16]. These tasks correspond to scenarios of information retrieval and foraging [25]. Divergent thinking tasks use open-ended questions that involve collecting and connecting multiple possible answers [7]. These are fundamentally different cognitive processes from convergent thinking, and so require different evaluation methods [30]. They correspond more closely to the cognitive experiences of the intellectual and creative tasks that comprise comparing, choosing, learning and research.

We need to discover how to build and evaluate tools, systems, and information environments that support creative processes, which are not addressed by convergent thinking tasks. Found information sometimes serves as a catalyst for the emergence of new ideas. *Information discovery tasks* involve divergent thinking in the context of browsing, searching for, and collecting digital

Figure 9. Example of a screen shot from a composition space state created by a student in The Design Process class. In the center, the cool space border is rendered explicitly. In this state, the cool space is empty.

information resources [16]. Information discovery tasks explore how people use information not simply to find, but to form and express ideas. How can we measure discovery? This is an important question for long-term research. Our current methods involve combining subjective user experience reports with evaluations by peers and experts.

4.2 Field Study: The Design Process in Undergraduate Education

We conducted a field study to validate the use of mixed-initiative composition of image and text surrogates for representing collections of information resources. In this study, alternating sets of members of an undergraduate course on The Design Process used combinFormation to create collections of prior work information resources relevant to their new inventions. The two mutually exclusive groups of students were both found to do better on the project, itself, when they used combinFormation to develop the prior work.

4.2.1 Experimental Method

In the field study, the mixed-initiative tool for composition of surrogates was utilized by students in two assignments in the interdisciplinary undergraduate course Environmental and Design Science (ENDS) 101/200, The Design Process. There were 182 students in the class, of which 47% were women and 53% were men. Academic majors were distributed, including 44% science and engineering, 33% architecture and liberal arts, and 17% business. The course engages these diverse students in an intensive process aimed at developing creative innovation in design.

The assignments in The Design Process course were already structured as information discovery tasks. In one assignment, the *Hybrid*, students are asked to

> Create the future by combining and connecting any services or objects that have never been linked before and illustrate your new service or idea. Search the Internet and the Patent and Trademark Library to see what the most relevant prior work is, as well as how your idea is original, and to collect the source materials for the existing services and objects that are being combined.

The description of a second assignment, the *Invention*, begins, "From your group's creative depths, journals or a posted Bug List, create at least three original inventions." The assignment continues with the same prior work collection requirement.

This was a real class, and not a controlled laboratory experiment environment, so we couldn't arbitrarily assign students to conditions in a way that was unfair. Therefore, we developed a method in which groups of students alternated using combinFormation for developing their prior work collections. For the first assignment, half of the students in the class were asked to volunteer to use combinFormation; in the second assignment, the other half of the class was asked to use the systems for representing their prior work collections as compositions of image and text surrogates.

In keeping with the ethical needs of the course, participation was not mandatory. Across the two assignments, two thirds of the students elected to participate by using combinFormation on one of the assignments. The groups of students that used combinFormation on the two assignments were mutually exclusive. Those students that did not use combinFormation used Google for searching, and Word for representing their results in the form of a traditional bibliography (Google+Word). For both assignments, we investigated both how the students performed on prior work collection development, and how they performed on the inventive deliverables.

4.2.2 Data and Results

We worked with the professor and teaching assistants for The Design Process course to develop evaluation criteria for evaluating both the collection deliverable, and the project itself. These criteria articulate the values of the course. and the evaluation process that was already in place. Additionally, a new 1-5 scale was instituted for the study. This scale corresponds directly to the letter grades that are assigned in the course. For the prior work, the criteria involve how informative, communicative, expressive, the collection is, as well as the variety of the collected resources. For the actual inventions, they involve originality, novelty, practicality, broad impact, and commercial transfer ability. While these measures are in some sense subjective, they are directly correlated and integrated with the evaluation process of the course. The evaluations were performed by the TAs as they were assigning grades based on the same criteria.

Approximately 81% of the students performed the Hybrid assignment (See Figure 10). 32.4% used combinFormation to develop the prior work collection, and 48.4% used Google+Word. Those who used combinFormation scored an average of 3.08 on the prior work, compared to 2.32 for those who used Google+Word, and the difference was significant [$t(118) = 3.528$, $p = 0.001$]. Likewise, those who used combinFormation also scored higher (3.32 vs. 2.85) on the actual Hybrid assignment, and again, the result was statistically significant [$t(145) = 2.227$, $p = 0.028$].

	cF	Google +Word	not involved	total / average
# of students	59	88	35	182
% of students	32.4%	48.4%	19.2%	100%
prior work	3.08	2.32		2.63
Hybrid	3.32	2.85		3.04

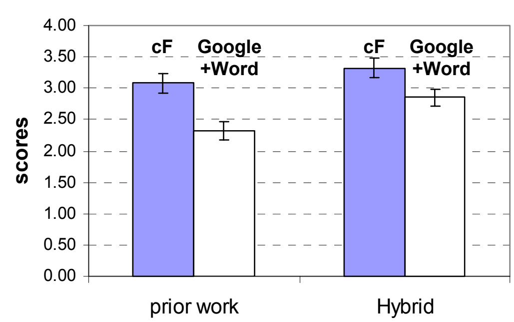

Figure 10. Student scores on the Hybrid assignment

The findings were similar for the Invention assignment. This time, 33.5% of the students used combinFormation, out of a total of 87% who did the assignment (See Figure 11). None of these were students who used combinFormation on the Hybrid. 53.9% used Google+Word for creating their prior work. The scores for the prior work collection were 3.13 for the combinFormation users vs. 2.38 for Google+Word [$t(141) = 3.843$, $p < 0.001$]. For the actual

Invention, the scores were 2.85 vs. 2.38 [$t(157) = 2.716$, $p = 0.007$]. The score differences of both for the prior work and for the actual invention assignment were statistically significant.

	cF	Google +Word	not involved	total / average
# of students	61	98	23	182
% of students	33.5%	53.9%	12.6%	100%
prior work	3.13	2.38		2.66
Invention	3.41	2.85		3.06

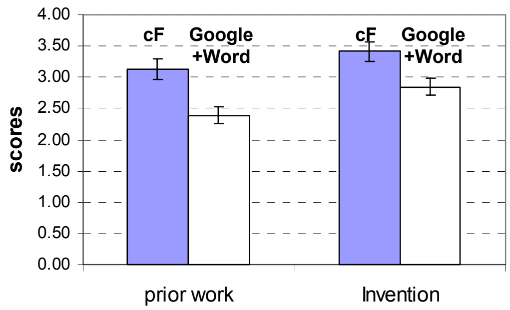

Figure 11. Student scores on the Invention assignment

From the field study, we found that combinFormation better supports students engaged in information discovery tasks in collecting and putting together prior works. According to the scores, the T.A. has 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. This is consistent with Chang et al's findings, that collage promotes collection understanding [5]. Further, subsequent to developing prior work collections with combinFormation, students performed better on the actual Hybrid and Invention assignments than those who used Google+Word.

5. DISCUSSION

We are just beginning to deploy combinFormation in The Design Process Course, and gather data from students about their practices and needs, and how the mixed-initiative system can serve them better. One finding is a need to improve the responsiveness of the agent, to enable more focused information retrieval, and more meaningful forms of generative composition. Another is the need for direct manipulation design features, such as grouping. Additionally, there have been some difficulties with field deployment, involving memory allocation by the Java Plug-in when invoked inside the Internet Explorer web browser.

In light of these shortcomings with the present deployment of combinFormation in The Design Process Course, we are particularly encouraged by the results of the present field study. Our explanation is that using the mixed-initiative system for collecting relevant information, and representing the collection as a composition of image and text surrogates stimulates the students to think about possibilities for their hybrids and inventions that are outside of the realm of what they would otherwise consider. Adding the temporal dimension to visual composition increases the set of information resources that the user is exposed to. The use of complementary image and text surrogate representations promotes cognition of this larger set of representations. Additionally, the accessibility of these surrogates in the composition space enables quick expressions of interest, which tune the semantic model, and thus the performance of the agent, to retrieve information that is more relevant to the user's emerging sense of the invention process. This traversal of a wider emergent space of relevant possibilities promotes information discovery; students create better hybrids and inventions.

Our investigation of how to support design processes by developing and representing collections of information resources through compositions of image and text surrogates is only beginning. We will continue to develop our relationship with The Design Process Course, investigating user needs, conducting formative and summative evaluations, and iterating mixed-initiative system designs. Through further dialogue with students in the class, and further development of the mixed-initiative system, we expect to articulate and develop further methods to promote information discovery. We can already see, for example, that support for digital libraries and semantic web repositories, such as patent collections, is directly relevant to the course's goals. Further, we intend to investigate support for information discovery in other research task contexts, such as graduate students developing papers and theses.

We are particularly interested in the relationships between support for information discovery, and for other related paradigms, such as information retrieval [3], information foraging [25], and information seeking [20]. This involves the differences between tasks in which the goal is to find relevant information, and those in which the goal is to put together, that is, to compose, relevant information. The user's needs in some tasks to understand connections, to compare and abstract, as well as to choose, are highlighted. Accounting for changes in the user's sense of what is relevant in the course of work on an information discovery task is part of this. We wonder what balances between finding exactly what the user has specified, and exploring a somewhat larger space of related but perhaps unexpected work will be most effective. We expect this will vary with task contexts, so it will be necessary to discover human centered means for supporting adjustment of the range of information exploration. Unlike the user's attention, compute power and network bandwidth continue to become less expensive. Thus, the incremental costs for locating a broader set of relevant information resources may be inconsequential. But the limits of cognitive attention dictate that we need to discover methods that balance these factors to enable users to see, understand, and connect some set of relevant information resources while performing their particular tasks.

6. ACKNOWLEDGMENTS

Support has been provided by NSF grant IIS-0429469. Thanks to the students of Texas A&M University ENDS 101/200 for working with us.

7. REFERENCES

[1] Amar, R., Stasko, J., A Knowledge Task-Based Framework for Design & Evaluation of Information Visualizations, *Proc InfoVis 2004*.

[2] Baddeley, A.D., Is working memory working?, *Quarterly Journal of Exp Psych*, 44A, 1-31, 1992.

[3] Belkin, N., ASK for Information Retrieval: Part I. Background and Theory, *Journal of Documentation*, Vol 38, Num 2, 1982.

[4] Burke, M., *Organization of Multimedia Resources*, Hampshire, UK: Gower, 1999.

[5] Chang, M., Leggett, J.J., Furuta, R., Kerne, A., Williams, J.P., Burns, S.A., Bias, R.G., Collection Understanding, *Proc JCDL 2004*.

[6] Diligenti, M., Coetzee, F., Lawrence, S., Giles, C.L., Gori, M., Focused Crawling using Context Graphs, *Proc VLDB 2000*.

[7] Ding, W., Marchionini, G., Soergel, D., Multimodal Surrogates for Video Browsing, *Proc DL 1999*, 85-93.

[8] Finke, R., Ward, T., Smith, S.M., *Creative Cognition*, Cambridge, MA, MIT Press, 1992.

[9] Gentner, D., Stevens, A.L., eds., *Mental Models*, Hillsdale, NJ: Erlbaum, 99-129, 1983.

[10] Glenberg, A.M., Langston, W.E., Comprehension of illustrated text: Pictures help to build mental models, *Journal of Memory & Language*, 31(2):129-151, April 1992.

[11] Horvitz, E., Principles of Mixed-Initiative User Interfaces, *Proc CHI 1999*, pp. 159-166.

[12] Jones, W., Dumais, S., Bruce, H., Once found, what then?: a study of 'keeping' behaviors in personal use of Web information. *Proc ASIST 2002*, November 18-21, 2002, 391-402.

[13] Interface Ecology Lab, combinFormation, http://ecologylab.cs.tamu.edu/combinFormation/

[14] Kerne, A., CollageMachine, A Model of Interface Ecology, NYU Ph.D. Dissertation, 2001.

[15] Kerne, A., Khandelwal, M, Publishing Evolving Metadocuments on the Web, *Proc Hypertext 2003*, 104-105.

[16] Kerne, A., Smith, S., The Information Discovery Framework, *Proc DIS 2004*, 357-360.

[17] Kerne, A., Smith, S.M., Mistrot, J.M., Sundaram, V., Khandelwal, M., Mapping Interest and Design to Facilitate Creative Process During Mixed-Initiative Information Composition, *Proc Interaction: Systems, Practice and Theory, Creativity & Cognition Symposium 2004*, 1-25.

[18] Kerne, A., Smith, S.M., Choi, H., Graeber, R., Caruso, D., Evaluating Navigational Surrogate Formats with Divergent Browsing Tasks, *Proc CHI 2005 Extended*, 1537-40.

[19] Kerne, A., Koh, E., Smith, S.M., Choi, H., Graeber, R., Caruso, D., Composition of Image and Text Surrogates Supports Expression During Information Discovery, submitted to DIS06

[20] Marchionini, G., *Information Seeking in Electronic Environments*, Cambridge Univ Press, 1997.

[21] Marshall, C.C., Shipman, F.M., VIKI: Spatial hypertext supporting emergent structure, *Proc ECHT 1994*, 13-23.

[22] McNee, S.M. Lam, S. Konstan, J.A., Riedl, J. Interfaces for Eliciting New User Preferences in Recommender Systems. *Proc User Modeling 2003*, 178-188.

[23] Morrison, J.B., Pirolli, P., Card, S.K., A Taxonomic Analysis of What World Wide Web Activities Significantly Impact People's Decisions and Actions, *Proc CHI 2001 Extended*.

[24] *Oxford English Dictionary on Compact Disk, 2nd Edition*. Oxford: Oxford University Press, 1992.

[25] Pirolli, P., Card, S.K., Information Foraging, *Psychological Review*, 106:4, Oct 1999, 643-675.

[26] Pirolli, P., Pitkow, J., Rao, R.. Silk from a Sow's Ear: Extracting Usable Structures from the Web, *Proc CHI 1996*.

[27] Porter, M. F., An algorithm for suffix stripping, *Program*, 14(3), 1980, 130-137.

[28] Rocchio, J., Relevance feedback in information retrieval, In The Smart Retrieval System-Experiments in Automatic Document Processing, Prentice-Hall, Englewood Cliffs, NJ, 313-323, 1971.

[29] Salton, G., McGill, M.J., *Introduction to Modern Information Retrieval*, New York, McGraw-Hill, 1983.

[30] Shneiderman, B., Maes, P., Direct Manipulation vs. Interface Agents, *ACM Interactions*, 4:6, Nov 1997.

[31] Shah, J. J., Smith, S. M., Vargas-Hernandez, N., Gerkens, R., & Wulan, M., Empirical studies of design ideation: Alignment of design experiments with laboratory experiments, *Proc Am Soc Mechanical Engineering*, 2003.

[32] Tufte, E., *Envisioning Information*, :Cheshire, Connecticut: Graphics Press, 1990.

[33] Woodruff, A. Rosenholtz R., Morrison, J., Faulring, A., Pirolli, P., A Comparison of the Use of Text Summaries, Plain Thumbnails, and Enhanced Thumbnails for Web Search Tasks, *JASIST* 53(2):172-185, 2002.