combinFormation: Mixed-Initiative Composition of Image and Text Surrogates Promotes Information Discovery

ANDRUID KERNE
Texas A&M University
EUNYEE KOH
Texas A&M University and Adobe Systems Inc.
and
STEVEN M. SMITH, ANDREW WEBB, and BLAKE DWORACZYK
Texas A&M University

5

combinFormation is a mixed-initiative creativity support tool for searching, browsing, organizing, and integrating information. Images and text are connected to represent surrogates (enhanced bookmarks), optimizing the use of human cognitive facilities. Composition, an alternative to lists and spatial hypertext, is used to represent a collection of surrogates as a connected whole, using principles from art and design. This facilitates the creative process of *information discovery*, in which humans develop new ideas while finding and collecting information. To provoke the user to think about the large space of potentially relevant information resources, a generative agent proactively engages in collecting information resources, forming image and text surrogates, and composing them visually. The agent develops the collection and its visual representation over time, enabling the user to see ideas and relationships. To keep the human in control, we develop interactive mechanisms for authoring the composition and directing the agent. In a field study in an interdisciplinary course on The Design Process, over a hundred students alternated using combinFormation and Google+Word to collect prior work on information discovery invention assignments. The students that used combinFormation's mixed-initiative composition of image and text surrogates performed better.

Categories and Subject Descriptors: H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval—*Search process*; H.3.7 [**Information Storage and Retrieval**]: Digital Libraries—*Collection, user issues*; H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems

Support is provided by NSF grants IIS-0429469 and IIS-0633906.
Authors' addresses: A. Kerne, A. Webb, and B. Dworaczyk, Interface Ecology Lab, Department of Computer Science and Engineering, Texas A&M University, College Station, TX 77843-3112; email: {andruid, awebb, blake}@ecologylab.net; E. Koh, Adobe Advanced Technology Labs, 321 Park Avenue, San Jose, CA 95110-2704; email: eunyee@adobe.com; S. M. Smith, Department of Psychology, Texas A&M University, College Station, TX 77843-4235; email: stevesmith@tamu.edu.

General Terms: Human Factors, Design, Experimentation

Additional Key Words and Phrases: Creativity support tools, mixed-initiative systems, collections, information discovery, creative cognition, exploratory search, software agents, focused crawler, semantics, relevance feedback, clustering, field study

ACM Reference Format:

Kerne, A., Koh, E., Smith, S. M., Webb, A., and Dworaczyk, B. 2008. combinFormation: Mixed-initiative composition of image and text surrogates promotes information discovery. ACM Trans. Inform. Syst. 27, 1, Article 5 (December 2008), 45 pages. DOI = 10.1145/1416950.1416955 http://doi.acm.org/10.1145/1416950.1416955

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. People are confronted with the problem of how to keep track of significant ideas within the stream of this experience. According to Morrison et al. [2001], the reason why people use the Web 69% of the time is to understand or compare/choose. In 71% of Web use scenarios, an important method people use to meet their information needs is to collect, that is, to assemble, information from multiple sources. Information collecting is part of a broad cross-section of human activities, including education, research, work, and entertainment [Koh and Kerne 2006]. We call the creative process of finding information and forming new ideas *information discovery*. Finding relevant information resources is only one part of human engagement in information discovery tasks [Kerne and Smith 2004]. Humans also need to develop an understanding of the connections among many diverse information resources. Information discovery is based on creative thinking; it involves the emergence of new perspectives and new ideas in contexts of stimuli of found information. The present research develops new methods for supporting processes of information discovery and new methods of validation.

We are developing and making freely available [Interface Ecology Lab 2007] a mixed-initiative creativity support system, combinFormation (cF), which uses composition to represent information while a person is searching for, browsing, collecting, arranging, and thinking about information [Kerne et al. 2006]. Image and text clippings are extracted 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. We need to represent surrogates in ways that promote quick human understanding of the ideas inherent in information resources.

A mixed-initiative system is one in which the actions of a human and an agent, working on a joint task, are interleaved [Horvitz 1999]. The basic premise of combinFormation is to collect and compose information for and with the human participant. The long-term research goal is similar in its complexity to that of Breazeal's [2002] social robots, even though the combinFormation agents are not anthropomorphic. The goal is to build a searching and collecting companion that, when asked, can model what the human participant is thinking about and interested in sufficiently well that it can anticipate her information needs in information discovery tasks, optimally stimulating her creative process. The objective of this article is to report on the approach and methods of the present system, and how it has been used by undergraduate students engaged in information discovery tasks involving creating inventions. The current version has been shown to work well enough to promote information discovery on this real-life coursework.

In combinFormation, human 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 agent's generative collecting and composition, the human's direct manipulation collecting and composition, and the human'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 model comprised by the collected semantics of the documents, the surrogates, relationships, expressions of the human interests, and functions that operate on this data. One of the human initiatives is to directly manipulate the composition through interactive authoring operations, which enable surrogates to be displaced, layered, resized, annotated, and removed. Human beings can also take the 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 the model through relevance feedback, interjecting the human into the agents' generative loops through the visual interface of the composition.

1.1 Targeting Users and Their Tasks Means Supporting the Open Web

We are developing combinFormation as a creativity support tool for a wide range of users who need to find and understand information while developing ideas. Among our target users are the undergraduate students in The Design Process Course offered at Texas A&M University. They are diverse in their academic majors, representing the Colleges of Engineering, Architecture, and Liberal Arts, and the Business School. Gender distribution is approximately equal. These students do not have any particular background or expertise in using digital systems.

In The Design Process, the assignments are information discovery tasks that involve creating new inventions. The course requires students to develop new knowledge, rather than to learn a particular set of facts. The methodologies involve creative processes, rather than a particular domain. As part of creating new inventions, a subtask the students work on is to create prior work collections. Like the background section of this article, a prior work collection assembles needs that motivate an invention, ingredients that serve as building blocks in making something new, and precedents, the prior research that is most similar. On alternating assignments, a different half of the students use combinFormation to develop a collection of prior work. Their information needs and the associated collection emerge through processes of browsing, searching, collecting, and composing, using combinFormation or the usual Google + Word. The prior work collections, and the processes through which they emerge, are examples of the many practices of collecting that humans engage in to develop meaning in their work and in their lives [Koh and Kerne 2006].

In creating their prior work collections, to do their assignments, the students need to search the open Web. Thus, unlike in some IR research, we are not working with a fixed corpus, but rather one that is constantly changing in new and noisy ways. Like metasearch providers, combinFormation forwards search queries to typical search engines, including del.icio.us and Flickr, as well as Google and Yahoo. However, instead of creating a unified ordered list of results, we represent results using composition, in order to promote integrated readings. We also use time as a dimension of presentation, in order to be able to present more results. The resulting ongoing set of composition space states are intended to serve as provocative stimuli [Shah et al. 2003] for creative ideation. Search result documents are downloaded, and processed. Their constituents are incorporated into combinFormation's Information Collection of documents, surrogates, semantics, and user interests.

1.2 Structure of this Article

This research develops the interaction of human and agent as they engage in searching for, browsing, and collecting information through a visual interface. We begin with a scenario, to introduce the reader to the experience of mixed-initiative information composition with combinFormation. This is followed by a section that considers the theory and practice of human experiences of information by examining prior work and synthesizing the creativity-oriented information discovery framework. Next, we develop our approach to representing collections in an integrated visual form that the user can manipulate. After these introductory sections, combinFormation's mixed-initiative system structure is developed. Since the program takes an unusual approach and achieves an unusual result, we thoroughly describe how it works. This begins with an explanation of how to install and launch cF, and then a description of the interactive methods that enable the human to compose and collect information, and direct the agent. Next, we present the structures of the Information Collection, which integrates metadata semantics, hypermedia and term models, and the participant interest components of the user profile. These collected model components drive the generative agent initiatives, which collect information on behalf of users, form image and text surrogates, and generate a navigable visual information composition. We follow the development of the system by presenting a field study, in which undergraduate students engaging in information discovery invention tasks were found to perform better when they used combinFormation to collect prior work. Finally, we discuss implications of our findings and draw conclusions, while considering how information discovery relates to information retrieval, foraging, seeking, and exploratory search.

2. SCENARIO

Audrey is a student in The Design Process class. Her assignment is to develop an original invention that is a hybrid of different existing technologies. Audrey's big idea is to develop a wearable location-aware information appliance that

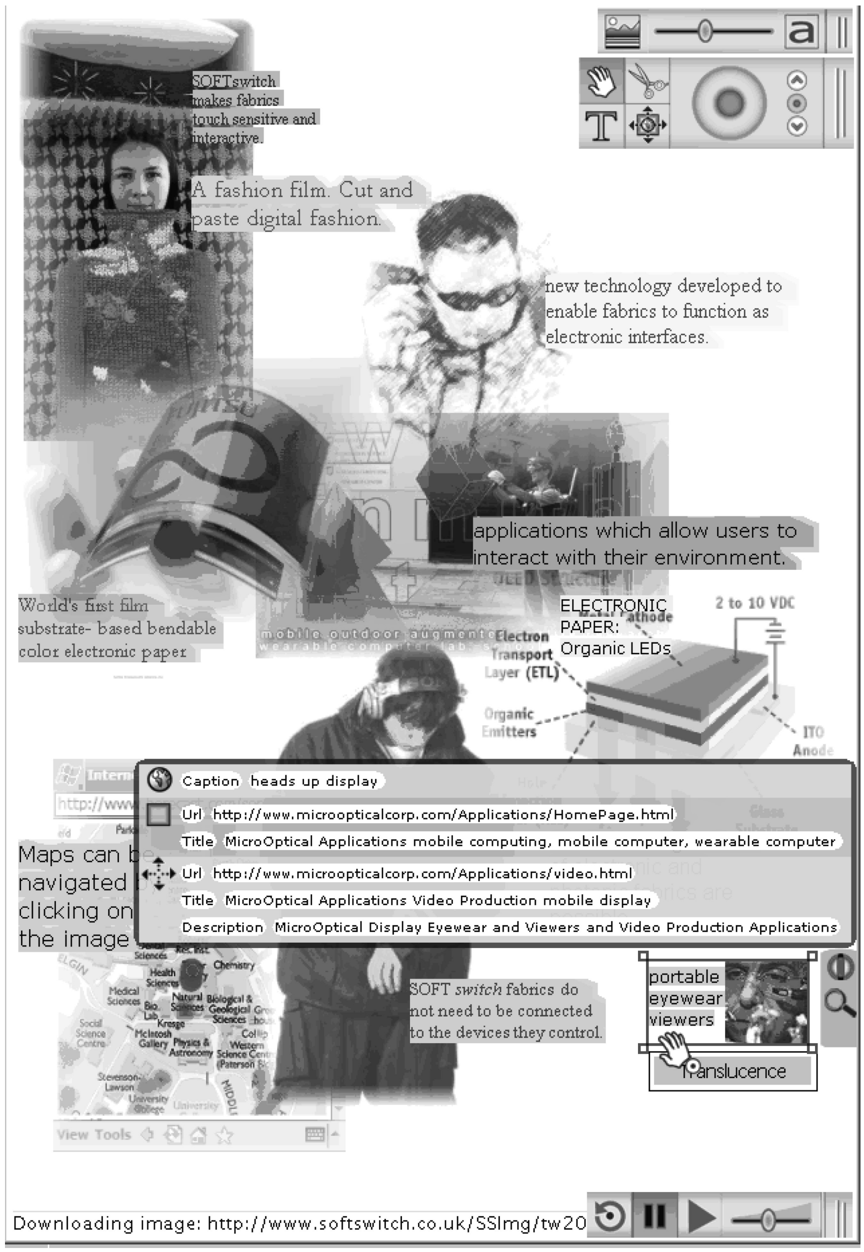

Fig. 1. Composition space for Audrey's research on a wearable location-aware information appliance. Each surrogate is a visual element that can be manipulated, and also enables navigation to source documents. In-context details on demand are shown for a surrogate about heads-up displays.

uses a high-resolution color display built into the fabric of a shirtsleeve. This device senses the user's current location, and provides contextual location-based information to enhance tourist activities.

To determine whether her invention can be developed in the next 3–5 years, Audrey needs to find current technologies that can be adapted and incorporated. She previously read an article about a new type of display that is paper-thin. Such a display, although not fabric, could provide a building block for the invention. She also remembers a friend telling her about new forms of fabric which are wired to carry an electrical charge; such fabrics can have designated areas function as control mechanisms like buttons.

To research the practicality and originality of her idea, Audrey uses combinFormation as a research tool. Audrey seeds combinFormation with searches based on the possible technologies she determined: paper thin display, electronic fabrics, and location-aware devices. combinFormation begins downloading and displaying images and text from the search results (Figure 1). As images and text appear, Audrey uses the navigate tool to open referenced pages and examine these pages for relevant information. One such is an image of a flexible display being bent by a hand. Upon examining this referenced page, she finds the article she had previously read about a paper-thin display. She rereads the article, noting important information.

Returning to the composition, Audrey uses the navigate tool on a text element: "applications which allow users to interact with their environment." This takes her to the Web site of the Tinmunth augmented reality project. Images show people wearing goggles that render augmented reality based on a person's GPS location and viewing direction. Another site shows portable eyewear displays that look like eyeglasses. Audrey notes that her invention can use the glasses (since goggles draw undesired attention to a tourist). Looking through the glasses could highlight tourist attractions in view (e.g., historic buildings, statues, art museums) and provide directions to those out-of-sight.

Audrey notices elements appearing about touch-sensitive cloth. Another inventive spark occurs: her fabric display can also be touch-sensitive. She uses the navigate tool, and explores the Web site for SOFTswitch, a manufacturer of electronic fabrics, noting possible devices that can be used for her hybrid invention.

Audrey begins moving interesting elements around, creating relationships. She moves elements into the central *cool space*, to protect them from combinFormation's agents. She juxtaposes the paper-thin display with location-aware devices and electronic fabrics. Audrey arranges elements for organization and for aesthetics. She is creating this composition not just for herself, but also as a deliverable for her assignment so that her professor and classmates can better understand her invention.

Audrey completes the composition and saves it. She brings it to class. Her composition functions as a visual aid in explaining her invention, its practicality, and its originality. Later, her classmates gather in-depth information by using the composition to navigate to her information resources, and using combinFormation's agents to find other related sources. In creating her final deliverable, Audrey refers back to her composition to help her gather information, organize her thoughts, and create a bibliography.

3. THEORY AND PRACTICE OF INFORMATION EXPERIENCES

We briefly survey prior work regarding people's experiences with digital information, relating processes of search to collecting and conceptualization. We develop information discovery as a framework of analysis, and as an approach to system-building and evaluation.

3.1 Prior Work

Many related research systems support collecting and visualizing information elements. Hunter Gatherer [Schraefel et al. 2002] is a tool and architecture to support Web-based, within-page component collections. The tool supports browsing, sorting, addition, editing, and deletion of components. Each component also has a link back to its source document for reference. VKB [Shipman et al. 2004] helps users access, collect, annotate, and combine materials from digital libraries and other sources into a personal information workspace. Users create a visualization of search results while selecting and organizing materials for their current activity and rapid addition of metadata to documents. Icon Abacus organizes citations into a compressed grid [Bier and Perer 2005]. PiggyBank [Huynh et al. 2005] is a tool integrated into the Web browser that lets users extract individual information items from within Web pages and save them in Semantic Web Format (RDF). These items, collected from different sites, can now be browsed, searched, sorted, and organized together, regardless of their origins and types. combinFormation builds on these approaches by connecting direct manipulation collecting with a generative agent that helps users find and assemble information. The present research differs in its emphasis on visual representations, which is integrated with giving the user the ability to manipulate the representation of the collection in the composition space. Additionally, combinFormation enables the user to provide relevance feedback and manipulate her interest profile through interaction in-context.

Cutting et al. [1992] and Hearst and Pedersen [1996] developed scatter/gather, which automatically clusters search results. The user can interact to focus on some clusters and eliminate others, generating relevance feedback. Baldonado and Winograd [1997] developed a system that enables users to collect intermediate search results in persistent folders, while iterating query definition, and related this to *sensemaking*. Marshall and Bly [2005] investigated the use of clippings in the physical world.

Stuff I've Seen [Dumais et al. 2003] investigates problems in personal information management, and develops solutions by using a timeline as the basis for aggregation of subsets of search results. Some researchers have suggested that rich search capabilities make explicit filing and organizing less important for retrieving personal information [Cutrell et al. 2006]. However, other studies have established that collecting and organizing remain important to users, no matter how good search tools become in personal information management [Nardi and Barreau 1997; Koh and Kerne 2006; Jones et al. 2005]. Reaccess to personal information is not necessarily the sole or the primary purpose of collecting. According to Marshall and Jones [2006], encountered information may be kept for pleasure. Users keep collecting to better see and understand their information. When people collect, they give attention as part of a practice of embodying significance, for example to remind themselves of things they experience as important, and as part of education [Koh and Kerne 2006]. Therefore, how the collection represents ideas can be valuable for meaning creation and enjoyment, as well as for using the material effectively.

Marchionini [1995] considered how humans iteratively define problems, form queries, and engage in searching and browsing to find information of value during *information seeking*. Pirolli and Card [1999] characterized humans as *inform-avores*, considering cognition, ecology, and statistics to develop computational models of *information foraging* based on optimizing costs and benefits. *Exploratory search* addresses situations in which users lack the knowledge or contextual awareness to formulate queries that meet their information needs, so browsing, exploring, and learning are required [White et al. 2006].

3.2 The Information Discovery Framework

Information discovery shifts the emphasis from finding information to having ideas, from search to human-centered creative cognition [Kerne and Smith 2004]. Cognitive psychologists since Maier [1931] have identified a shift in representations, that is, *cognitive restructuring*, as an essential step in insight and *creative ideation* (idea generation; or the emergence of new ideas) [Finke et al. 1992]. *Information discovery* means ideation in conjunction with information finding. The representation shifts associated with insight and ideation, such as changes in conceptual framing and information needs, are the crux of information discovery knowledge creation tasks, such as invention and the formulation of a thesis topic. Found information stimulates seeing new perspectives and formulating new models. Information finding and manipulation can help humans overcome fixation. Fixation means getting stuck in an unproductive mental set [Smith 1994].

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 [Finke et al 1992; Kerne and Smith 2004]. Among convergent thinking tasks are those used to evaluate information foraging [Pirolli and Card 1999], as well as those used by Woodruff et al. [2002] and Marchionini's group [Ding et al. 1999; Wildemuth et al. 2003] to evaluate search result representations.

Divergent thinking tasks are based on open-ended questions that involve collecting and connecting multiple possible answers [Shah et al. 2003]. These are fundamentally different cognitive processes from convergent thinking, and so require different evaluation methods. They correspond to the cognitive experiences of the intellectual and creative tasks that comprise comparing, choosing, learning, and research. In information discovery tasks, such as the invention tasks of The Design Process Course, the participant must also connect the set of answers, in a coherent form. To evaluate these answers, in the course, grounded evaluations are conducted by experts, the teaching assistants (TAs) who grade assignments. The present research augments these evaluations with user experience reports.

4. INFORMATION REPRESENTATIONS: COMPOSITION OF IMAGE AND TEXT SURROGATES

We develop the composition of image and text surrogates as an optimal representation for information discovery of search results and personal collections. Findings in cognitive psychology and human-computer interaction demonstrate the complementary cognitive roles played by text and image representations of information. The surrogate serves as a mainstay of collections by representing significant information from a document, and enabling access. The arts offer the integrative form of *composition*. We apply composition to the representation of collections to facilitate information discovery.

4.1 Prior Work

4.1.1 *Integrated Image-Text Representations Promote Cognition.* In the working memory system, the visuospatial buffer (which stores mental images) and the rehearsal loop used for words are complementary parallel processing subsystems [Baddeley 1992]. They support each other in combined image-text knowledge representations. Glenberg and Langston [1992] and Glenberg [2002] established that the combination of an image and descriptive text promotes the formation of mental models, extending working memory capacity. Mayer and Moreno [2002] found that dual coding strategies enhance cognition during educational experiences of digital media. Text disambiguates images while engaging complementary cognitive subsystems. For representing search results, Woodruff et al. [2001] used "enhanced thumbnails," annotating a thumbnail of a document with visual text callouts of search queries.

4.1.2 *Surrogates.* A *surrogate* represents an information resource and enables access to that resource [Burke 1999]. Hypermedia surrogates, which function as navigation, are formed systematically from metadata. One typical surrogate is the Google snippet, an element of the result set returned by a search query. 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 [Jones et al. 2002], that is, in remembering and relocating what has been collected during searching and browsing.

Marchionini's group investigated the use of multimodal surrogates for video browsing [Ding et al. 1999; Wildemuth et al. 2003] by comparing users' performances and experiences using different surrogate formats for digital videos. Combined surrogates lead to better comprehension and reduced human processing time. Our approach to forming and evaluating surrogates is oriented to supporting information discovery.

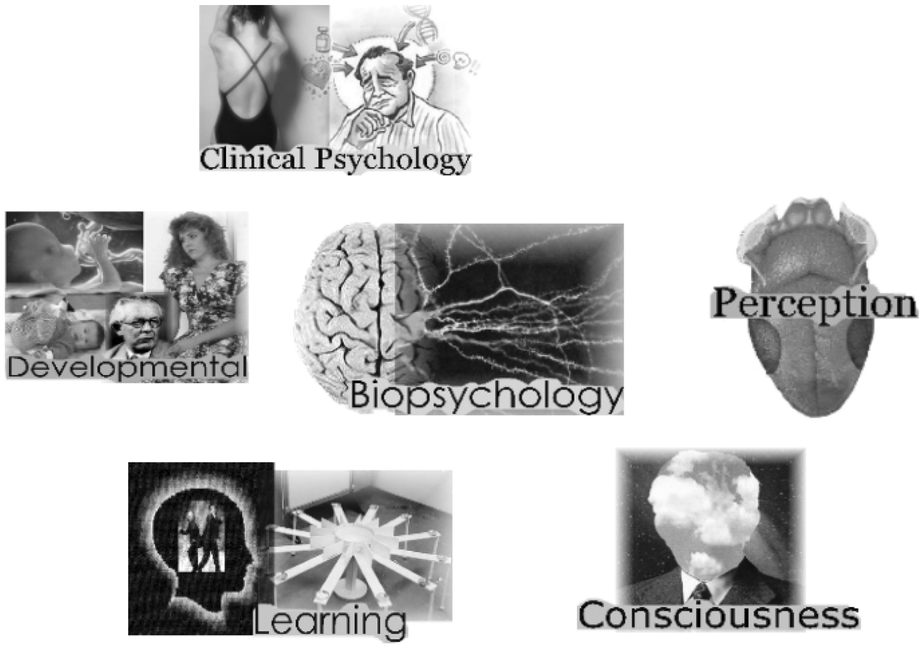

Fig. 2. A composition of image and text surrogates represents a collection of information resources, forming a partial overview of the undergraduate psychology curriculum. Each surrogate is formed by clipping information elements from source documents.

4.2 Image and Text Surrogates

During search and recall of resources from personal collections, surrogates form the basis of human decisions about which documents to browse and which to pass by. Surrogates can also help people to think about the relationships among the significant ideas in information resources. Thus, they can play a fundamental role in cognitive processes of comparing and choosing. As the significance of the representation of surrogates grow in importance, the surrogate comes to eclipse the original document by standing between the user and the document itself. For information discovery, we need to discover better representations for individual surrogates than text alone, and better representations for collections than the list.

We need to represent surrogates in ways that promote users' quick understanding of the ideas inherent in information resources, and connections among them. Using images and text to represent documents makes optimal use of parallel cognitive processing units. Our approach to forming image and text surrogates is more focused than enhanced thumbnails. Instead of using a thumbnail overview, which represents document layout more than constituent ideas, we use a *sample* [Kerne and Sundaram 2003], also known as a *clipping* [Marshall and Bly 2005] to represent an idea from a document (see Figure 2). To represent parts of documents, instead of using a summary, we build text surrogates by extracting significant textual phrases. These visual representations of the surrogate are augmented by metadata, such as the title of a document and the caption of an image. The goal is to use surrogates to focus the representation of finer-grained ideas that lie within documents, as appropriate, in ways that reflect the intentions of document authors, and the needs of people collecting information. The new surrogate will represent significant ideas from the document, in the context of the emerging collection and the associated information discovery task. In combinFormation, underlying the surrogate structure is the inherent relationship between a clipping and its source document, and (where present) a hyperlink reference. These relationships ensure that the clipping can function as a surrogate, affording navigation back to its source and out to its hyperlink.

4.3 Representing Collections as Compositions

The list of textual surrogates is typically used to represent collections, such as search result sets and bookmarks. Composition is an alternative to lists; literally, it means, "the act of putting together or combining ... as parts or elements of a whole" [Oxford English Dictionary 1992]. Composition of image and text surrogates extends the organizing of information afforded by spatial hypertext [Marshall and Shipman 1994] by emphasizing visual design, conceptual connection, and communication (for examples, see Figures 1 and 2). Spatial hypertext allows a participant to incrementally instantiate, organize, and structure elements and their relationships. The representation "talks back" to the user, supporting reflection in design [Yamamoto and Nakakoji 2005].

Composition is a fundamental artistic process. The arts, historically, are the intuition-driven domain in which creativity is primary. Thus, it makes sense, intuitively, to draw on artistic methods for supporting creativity. The artist Marcel Duchamp emphasized the creativity inherent in finding and choosing objects [Lippard 1971]. Collage [Spies 2006], montage [Eisenstein 1975], and hip-hop remix [Spooky DJ, that Subliminal Kid 2004] are artistic forms of composition based on juxtaposing collected elements to form a whole, and so they are particularly relevant to the representation of collections of surrogates. Generalizing across media, the creative forms of composition through juxtaposition of found and collected elements are known alternatively as *recombinant media* or *recombinant information* [Kerne and Sundaram 2003; Kerne et al. 2004a].

Composition uses visual design techniques that connect and layer elements [Tufte 1990] to form a coherent whole, including images, text stroking, and compositing, as well as relative size relationships, colors, and typefaces. Connecting elements visually is important because it promotes chunking, a means for overcoming the 7 ± 2 entities' capacity limitation of working memory [Miller 1956; Simon 1971]. Connecting elements is accomplished through layering and through compositing.

Compositing, a means for *making visible* strong connections among elements, is accomplished through the image processing technique of alpha gradients. This technique decreases the alpha level radially across the border area of an image surrogate, from opaque in the center to more translucent in the perimeter. When the translucent portions of such a surrogate overlaps others, the underneath layers become visible. Thus, compositing results in visual cross-fades. Compositing contrasts with the hard-edged contrast of juxtaposition without blending. With combinFormation's mixed initiatives, both the user and the generative agent can create compositing effects by adding translucence to an image surrogate or maintaining opacity.

4.4 Manipulable Composition of Image and Text Surrogates

The present research applies the composition form to represent collections of image and text surrogates, while bringing attention to the processes through which collections are assembled, and how the resulting forms function as artifacts for communication and navigation, and stimuli for cognition. Figure 2 shows a composition that represents areas of the undergraduate psychology curriculum. By *composition space*, we mean the mixed initiative environment in which the user and agents engage in the process of putting the composition together. The use of collected elements in the hypermedia composition space 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. The form of information received in response to search queries is immediately manipulable, instead of fixed. In response to the needs of users [Morrison et al. 2001], collecting is integrated with search. The mixed-initiative composition space is being designed to make it easy for nonexpert users to assemble collections of information resources as compositions of image and text surrogates.

5. MIXED-INITIATIVE INFORMATION COMPOSITION SYSTEM

> *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 [1999], page 17

In the mixed-initiative information composition system, combinFormation, 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 mixed-initiative composition space is shared by human and agent, integrating functions of search, presentation, organization, feedback, and discovery. The Shneiderman-Maes direct manipulation versus agents debates made clear that, if software agents are to be usable, we must develop interfaces that enable humans to easily control them [Shneiderman and Maes 1997]. Thus, we develop interactive methods to enable the user to manipulate the space, and to direct the agent. The agent initiatives are *generative*: they work proactively, over time, to procedurally create an evolving collection and visual composition representation. To do so, the

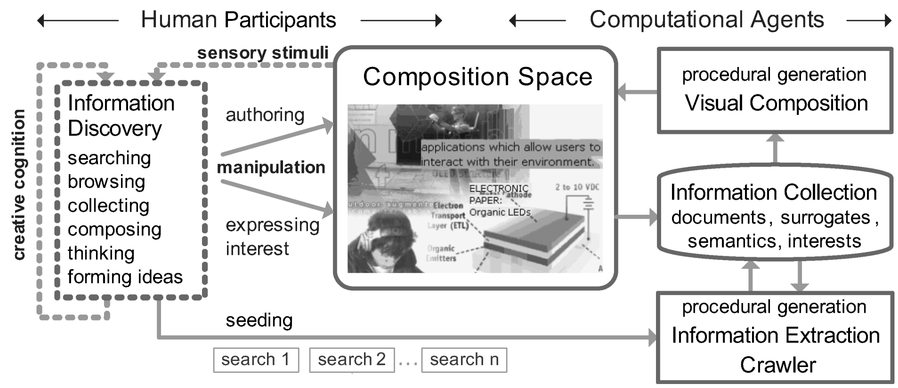

Fig. 3. Mixed initiatives in combinFormation, an experiential and functional overview.

software uses hypermedia and information retrieval techniques, developing the information collection, which models the information and the participant's interests. The interface enables the user to express interest, and thus provide feedback to the model. Generative information collecting initiatives use the model to run a surrogate extractor and focused crawler. Generative temporal visual composition initiatives are procedural algorithms that iteratively synthesize a visual representation of the evolving surrogate collection in the composition form.

The composition and its surrogate components serve as a visible medium for communication between human and agent, as well as one for collecting and sharing information resources. The semantic and user interest components of the information collection also serve as a medium between human and agent, but in a behind-the-scenes computational form. This section addresses the mixed initiatives in combinFormation. Figure 3 provides an experiential and functional overview of the mixed initiatives that the human participant and computational agents engage in through the composition space, and their impact on information discovery and creative cognition. It shows their relationships, and how they affect each other, the composition, and the semantic model. Seeding is the primary means through which the user first specifies information to be collected. Documents, clippings as surrogates, and associated semantics are extracted from the seeds, initiating the information collection. The agents use the collection as the basis for generative information extraction and temporal visual composition. A focused Web crawler continues to generate the collection by recursively processing selected hyperlinks. Through direct manipulation information collecting, the user brings surrogates and their underlying semantics directly into the composition space and the collection. Through direct manipulation composition authoring, the user can change how the composition looks and what it says, in order to facilitate her or his own understanding of the information resources and their connections, and to communicate such understanding to others. By directing the agent, the user-in-the-loop can turn the agent initiatives off and on, and also provide manual relevance feedback through interest expression that changes the model of the user's interests. The efficacy of the mixed-initiative composition system as a whole comes from the integrated functionalities of the collection and its semantics, the procedural information extraction and crawler, the procedural composition visualization, and the interface of the composition space.

In the information collection, each clipping functions as a surrogate by maintaining association with its source document, or *container*, and hyperlinks, to form the hypermedia graph. The container is a generalization of the source document. Examples of containers include 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 clippings, that form surrogates, metadata field semantics, and hyperlinks to other containers, which can be processed recursively. As it acquires clippings and container hyperlinks, the agent may add them to pools of candidate elements. The agent initiatives then use the pools in their generative operations.

5.1 Installation

combinFormation (cF) is distributed as a DHMTL seeding layer that connects to a signed Java Web Start (JWS) application. It requires Java 5 or 6 to be installed [Sun Microsystems 2006]. With Java installed, the user simply accesses the combinFormation Web site [Interface Ecology Lab 2007], chooses a seeding mechanism, and launches cF. The first time that the user activates launch from within the site, the latest versions of the appropriate Java Archive (JAR) files, and other configuration files will be installed automatically. Thereafter, launch will automatically download only components that have been modified and are out of date in the user's installation. The launch page passes parameters regarding seeds and configuration to the Web start application. The "fat client" architecture takes advantage of the growing power of contemporary PCs.

5.2 Seeding/Launch

The user launches combinFormation through one of several DHTML mechanisms that assemble seeds for the agent to use as starting points for information collecting [Interface Ecology Lab 2007]. 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 the *re:collection*, mechanism, the user can select a set of seeds that has been curated by members of our team. An example is the "news collection," which puts together material from news feeds such as CNN, *The New York Times, The Guardian*, and the BBC. The most general and powerful launch method, *re:mix*, lets the user specify and assemble any number of seeds for the agent. Each seed specifies either a document location (HTML, PDF, or file system directory), an RSS feed, or a search query (Google, Yahoo, Yahoo Images, Yahoo News, del.icio.us, or Flickr).

5.3 Composition Space: Participant Initiatives to Author and Direct the Agent

When the user presses Launch, the combinFormation Web Start application is executed, creating the composition space. The seeds are passed to the information extraction module. Surrogates are formed from image and text clippings with associated metadata semantics. These feed the visual composition agent, which selects the most relevant surrogates and flows them into the composition space. Extracted hyperlinks are passed to the crawler. The human participant interacts with the composition space through direct manipulation.

As Shneiderman and Maes [1997] have articulated, the promise of agents that assist the user is tenuous. It depends on interactive mechanisms that enable the user to effectively direct 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 [Kerne et al. 2004b]. Thus, we develop the role of surrogates and direct manipulation in the composition space as means to enable the user to author the composition and direct the agents.

5.3.1 *Composition Authoring, Navigation, and Collecting.* Within the composition space, the human can directly position, resize, color, edit, and remove surrogates. S/he can also navigate from each surrogate in the composition space back to its hyperlink and source document, in a traditional web browser. When the human finds information that s/he wishes to collect in a source document, s/he can mouse over and click to select it in the Web browser, then drag it over to combinFormation, and drop it into the composition space. For the drag and drop operation to represent material that is selected in the source Web page as one or more surrogates in the composition, the program needs the context of the Web address of the source container document. We developed a Firefox plug-in to pass this contextual metadata from the Web browser to cF during drag and drop.

5.3.2 *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 [Kerne et al. 2004b]. 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 parts of the composition: the cool space and the latch.

The cool space is a resizable rectangle in the center of the composition space (Figure 4). This space is for the user, only, to develop the collection as composition. The cool space is in the center, framing the user's direct work as the focus, while, in the outer area, the agent's temporal visual composition initiative in the mixed-initiative peripheral hot space provides context. The user can place surrogates into the cool space via drag and drop either from the hot space or directly from documents. The cool space is resized by simple drag and drop of its border. The cool space size is constrained to fall on grid cell boundaries (see Section 5.5.2). The generative visual composition initiatives will not remove or age any elements that the user places in the cool space.

The latch is an in-context tool, which enables the user to establish a single surrogate as a floating cool space. The agent will not remove, cover, or age any

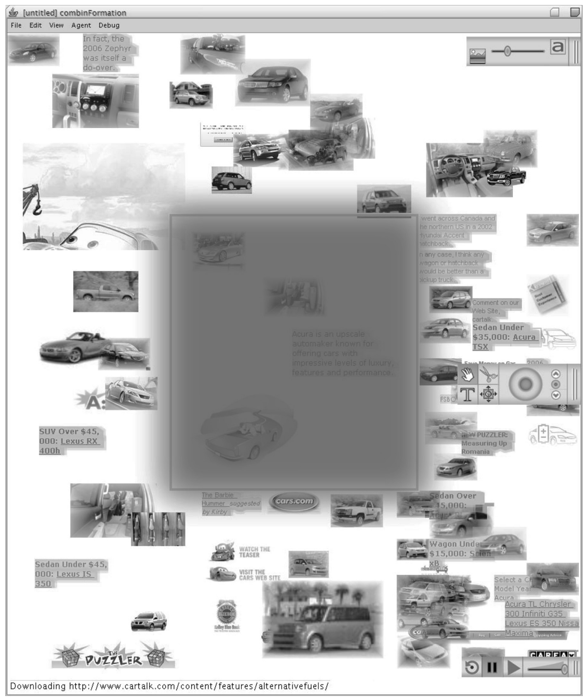

Fig. 4. The cool space is highlighted in the center. The hot space is at the peripherae.

Fig. 5. A latched element.

latched element. The latch tool turns from vertical to horizontal when activated. To make state visible, latched surrogates remain marked with the latch in their upper right hand corner (Figure 5).

5.3.3 *Interest Expression.* A primary means of directing the agent is through the interest expression interface [Kerne et al. 2004b], which affects a profile of interests. Through this interface, the user can activate up and down arrows, which signify the intent to increase or decrease interest level values in the participant objects associated with a surrogate. After the user has selected

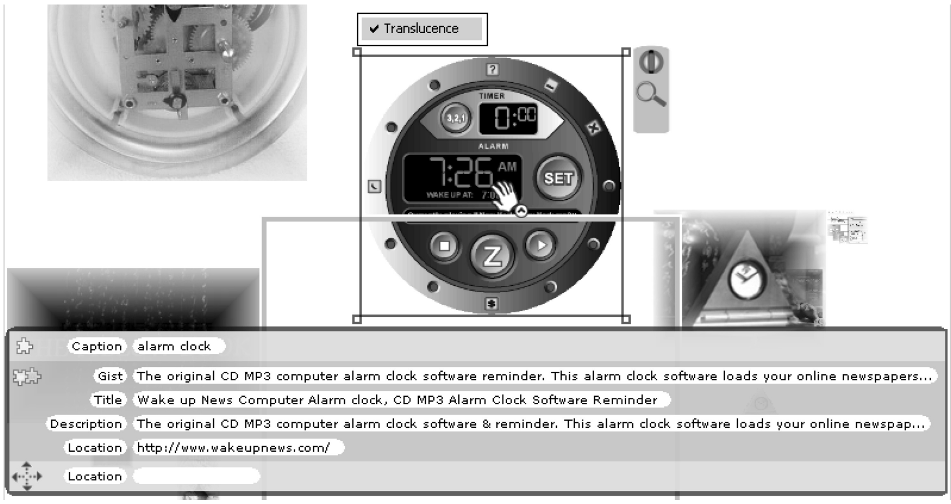

Fig. 6. In-context details on demand and in-context tools are displayed next to the element when activated through brushing by simple mouse over. Excerpt from a student composition in The Design Process.

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, 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 or his 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 [McNee et al. 2003]. Thus, in interest expression interactions, the surrogate clipping serves as an affordance for relevance feedback. In the course of a 21-min. authoring session, combinFormation users were found to conduct 92 interest-level operations, and 202 authoring and navigation operations. The number of times that the two kinds of operations were conducted was on the same order of magnitude. We interpret this result to demonstrate that users are able to express interest successfully, and are motivated to do so.

5.3.4 *In-Context Details on Demand and Tools.* When a user brushes a surrogate by mouse over, in addition to in-context metadata details, in-context tools are displayed. In Figure 6, on the right part of the image surrogate, the latch tool and the synthesize search tool are displayed. The synthesize search tool enables users to create a new search query in the midst of the session. When the user clicks the synthesize search tool, the agent will form a search query from the significant terms of the surrogate, and generate a new search for the user. The edit palette above the surrogate in Figure 6 enables the user to add or remove a translucent border gradient on the image surrogate (see

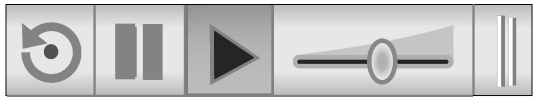

Fig. 7. Tape recorder transport controls the agent's generation of the visual composition.

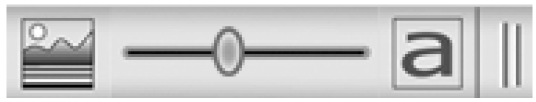

Fig. 8. Image-text surrogates mix slider.

Section 4.3). This creates compositing with underlying neighbors, for a sense of visual connection. For text surrogates, the edit palette enables changing stroke colors, font sizes, and font styles.

5.3.5 *Affecting the Agent's Flows of Control: Tape Recorder Metaphor.* Since the temporal visual composition agent initiative can continuously change the composition space visualization, it is important to give the user direct control over this process. A tape recorder transport enables pausing the agent's process of generative temporal visual composition. Play means the composition agent is activated and running. Pause means to halt its execution. A slider in the same floating window enables adjusting the rate of this process. The circular icon to left is Reset, which deletes all surrogates from the hot space. 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 (see Figure 7). An experimental version of the tape recorder transport enables the user to go backward in time, with a jog-shuttle control, as well as forward [Khandelwal et al. 2003]. However, this capability was not in the version used in the field study.

5.3.6 *Affecting the Agent's Flows of Control: Image-Text Surrogates Mix Slider.* The generative information collecting agent initiatives maintain separate pools of image and text surrogate candidates. The temporal visual composition initiative periodically chooses a new surrogate candidate to add to the composition. Text surrogates are much easier to collect, because documents contain more text than images. Further, image surrogates must be downloaded separately. Yet users often want to see more images. The mix slider allows the user to specify a desired ratio of image and text surrogates in the composition (see Figure 8). If the image candidate pool becomes empty, and an image surrogate should be added to the mix, according to this slider's ratio, the temporal visual composition agent may skip updating the composition. The skip can last up to 10 composition cycles. This method was arrived at as a compromise, through an iterative design process. When a prototype immediately added text even though an image was called for, users thought the slider was broken. Waiting indefinitely without updating the composition even though play is activated leaves users thinking the agent is entirely broken.

5.3.7 *Visual Metadocuments.* Compositions serve as a medium for exchanging personal and institutional collections. These collections of surrogates connect the process of searching with that of thinking about, and otherwise using the information. They may be saved and published on the Web as visual metadocuments [Kerne and Khandelwal 2003]. Visual and semantic components of surrogates are saved 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 maximizes accessibility to a broad range of users, as a Web browser is all that is needed to use it. This representation 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 full XML in combinFormation. The DHTML is a read-only snapshot, which only lacks the capabilities to author and generate.

5.4 Information Collection

The information collection and its model underlie operation of the generative agents. It contains all content, and represents the structural relationships between containers, the information elements clipped from containers to form surrogates, associated metadata, hypermedia links and the human participant's interests (see Figures 3 and 9). One model component is the hypermedia graph, which represents the referentiality of the authored and dynamically generated hypertext of the World Wide Web, and of the user's file system. Another primary component is based on the vector space model of information retrieval (IR) [Salton and McGill 1983]. It uses terms to connect surrogate candidates (and visible surrogates). The participant interest components of surrogates, documents, and terms utilize human expressions of interest to model the user and manifest relevance feedback [Rocchio 1971]. 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 containers to crawl, and select layout positions of the new surrogates as they are added to the temporal visual composition.

5.4.1 *Extracting Semantics through Metadata.* A semantic description of each container and clipping/surrogate takes the form of a set of metadata fields (see Figures 7 and 9). Each field has a name, a type, and a value. The extensible type system, which is based on ecologylab.xml [Kerne et al. 2008], already supports numbers, dates, and URLs, in addition to text. Metadata is acquired from sources such as the semantic Web, digital libraries, and HTML markup. For example, when information is acquired from semantic Web RSS feeds, like Yahoo News or Flickr, fields such as subject and description are utilized for each item. When HTML documents are processed, image surrogate candidates may acquire metadata from HTML markup in the form of the alt attribute of the img element, if this has been provided by the document author. We translate the alt

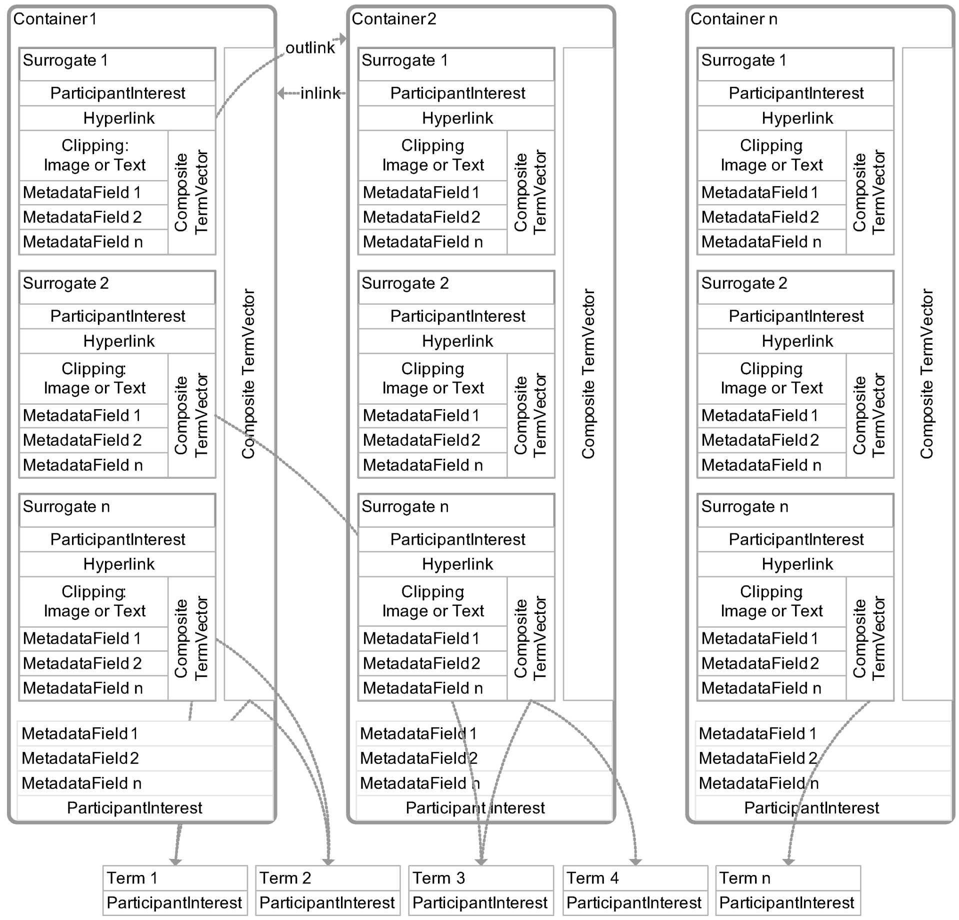

Fig. 9. The information collection models semantic entities and their referential relationships, through the hypermedia graph and term models.

attribute into caption, a more comprehensible nomenclature for the user. Future work will use the Document Object Model [W3C 2000] to associate a textual context with images that were not annotated by the Web page developer with alt.

When the user places her/his mouse over a surrogate within the composition, in-context metadata details on demand are displayed (Figure 6). 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 at 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.

5.4.2 *Hypermedia Graph Model.* The graph model is based on the hypermedia structure of World Wide Web documents (see Figure 9). This semantic model component continuously evolves from the processing of initial seeds and crawling. It defines relationships between surrogates and containers, which are utilized to spread activation when a user expresses interest, and by the generative agents, through weighting metrics. Each surrogate candidate that is clipped from a container document maintains a reference to its container. With each clipping and associated surrogate candidate, a hyperlink reference may, depending on the authored content and semantics, also be associated.

The container also refers to the set of surrogates which have been formed from it. Each surrogate hyperlink constitutes an *outlink* from its originating container. In turn, the hyperlink reference container destination functions as an *inlink* to the originating container. The inlink structure must be maintained in order to maintain the integrity of the graph when the crawler discovers server-side redirects. The graph model is propagated in this way through recursive chains, which are formed from the structure of the authored content encountered through processing the initial seeds, and through the focused Web crawler.

5.4.3 *Term Model.* The Hypermedia Graph Model supplements the Term Model. The Term Model enables associations to propagate across Web sites through textual, rather than hypermedia references. To form the Term Model, combinFormation uses information retrieval methods of the vector space model [Salton and McGill 1983; Rocchio 1971] to connect surrogate candidates (and visible surrogates) by common terms. These associations are utilized in weighting measures by the generative information collecting algorithm (Section 5.5.1), and in similarity measures by the generative temporal visual composition algorithm (Section 5.5.2).

Salton's vector space model assumes that each information resource is a document. Each document is represented as a vector, in which each dimension represents a stemmed [Porter 1980] term. The full set of documents is called the *corpus*. The next part of the vector space model involves the term frequency, TF, and the inverse document frequency, IDF. TF represents the number of instances of the term found within an information resource, which, originally, is assumed to be a document. This is an integer contribution to the scalar magnitude for the dimension associated with a term. A floating-point contribution to a term's scalar weight magnitude is IDF, which counts the number of documents in which a term appears in the corpus, and derives a normalized factor to represent this. IDF is high for terms that are found infrequently, that is, those which are good discriminators for search and relevance feedback, and low for frequently encountered terms. In the present research, we always set TF to 1. That is because our primary granularity for information resources is the surrogate, rather than the document. For a surrogate, the repetition of terms is not a significant indicator of semantic weight.

For each surrogate candidate and surrogate, a composite term vector is formed through the union of the associated metadata fields. Additionally, for text surrogate candidates and surrogates, terms from the text also become part of the composite term vector. The associated words are stemmed 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 *web*, *click*, and *e-mail*. In case there is no explicit metadata for a container or image reference, the system will attempt to form a mined keywords field from the URL. To reduce the occurrence of noisy noninformative term associations, which interfere with the operation of the semantic model, only terms found in the prebuilt term dictionary are added to the mined keywords.

An inverted index, which associates a set of surrogate candidates and surrogates with each term object, is formed, with entries for each surrogate candidate and container that refers to the term. Computationally, dynamically constructed term vectors are utilized in conjunction with a static prebuilt term dictionary, which contains frequency counts for each of the terms discovered by parsing a sample corpus 6000 random Web pages. We construct this dictionary based on sample corpus because we lack the resources to maintain a dictionary on each user's machine that represents every document on the Web, and that is available through file systems on their computer. We also do this, instead of using a dictionary from a source like WordNet [Miller 2006], because the terms and frequencies of words on the net differ from those found in various literatures. The dictionary is constructed during an offline process, and installed on the user's machine. Six thousand pages gives us 155K terms, and a resulting file size of 664 KB when zipped. This order of magnitude is similar to that of the code, itself, and so was deemed a reasonable level for imposing on users during installation. The term dictionary enables the computation of significance weights using inverse document frequency (IDF), the inverse of the ratio of pages that contain a term to the total number of pages:

$$IDF_{term} = \log(N/n_{term})/\log N,$$

where N:= sample corpus size, and n_{term} := number of documents that contain the term.

Of course, since the Web is not a fixed corpus, and the prebuilt term dictionary is only constructed from a sample of pages, it does not initially contain term entries, document occurrence counts, and IDF calculations for every possible term; some that were not encountered during the dictionary-building process will be discovered in the course of a session. Thus, the dictionary grows as the program operates and discovers new terms. However, the discovery of terms in the course of a session does not contribute to the IDF count. The reason for this is that it would penalize the agent's success in dynamically discovering relevant documents. In any given session, the agent collects some subset of the total set of documents available on the Web. The more successful the agent is in finding relevant documents, the more representative this collection becomes of the user's actual information needs, rather than the Web as a whole. Thus, relevant terms will be collected disproportionately. Inclusion of document frequency counts from this collection would be skewed. Also, unlike in prior systems [Balabanovic 1998], to account for the noisiness of collection information, newly discovered terms are not assigned the largest IDF value ($\log\frac{N}{1}/\log N = 1$). We followed that approach initially, but in practice, while processing documents on the open Web, too many noninformative terms were injected into the model, such as user ids and misspelled words. The current implementation assigns to newly encountered terms the mean weight of all terms in the dictionary. This policy allows the system to retrieve more relevant information by preventing

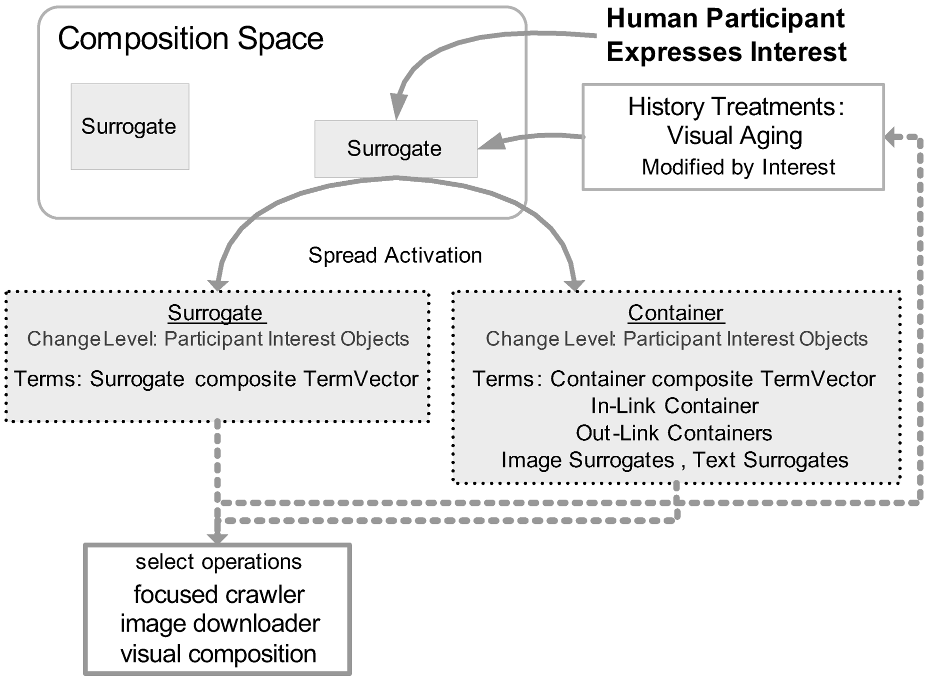

Fig. 10. Spreading activation in response to interest expression: The participant's expression of interest in a certain surrogate spreads through participant interest objects in the composite term vector, and in the surrogate and the container. These, in turn, impact the generative agent initiatives.

new terms from having too much affect on the retrieval model, while enabling them to make a balanced contribution.

5.4.4 *Modeling the Participant's Interests.* In order to enable the system to adapt to the user's spontaneously emergent sense of information needs, a participant interest object is associated with each surrogate candidate and surrogate, with each container, and with each term. The set of these forms a profile of the user's interests. Within the participant interest object, interest level is modeled as an integer value on $[-10, +10]$; these discrete levels are set through the participant interest interface.

When the participant explicitly expresses interest in a surrogate, this expression is propagated through the model into the appropriate participant objects by spreading activation [Pirolli and Card 1999] to semantically related nodes. The participant interest object of the interest expressed surrogate receives the interest activation (see Figures 10 and 9). The participant objects of terms in the composite TermVector of the surrogate also receive the interest activation. The interest activation spreads to the surrogate's hyperlink container, if there is one, which is the information resource that the surrogate most directly represents. If there is no hyperlink, a lesser amount of interest is propagated into the container. Candidate image and text surrogates formed from this container also receive spreading interest activation. For hyperlink containers, the propagation spreads recursively with a near half life damping factor (5/9). In addition to enabling the user's explicit expression, combinFormation utilizes implicit interest expression [Kelly and Teevan 2003], whenever the user submits a search query. The system automatically increases the participant interest level in each term that the user enters.

These propagated interest expressions contribute to the weighting metrics the agent utilizes in its generative collecting and visualizing initiatives. These measures govern the choices of containers and images downloaded by the generative information collecting initiatives (see Section 5.5.1). Interest expression can also affect the visual representation of surrogates. The temporal composition agent uses interest expression while computing weights to choose, size, and arrange the surrogates that are added to in the visual composition space over time, and those that are gradually removed. Likewise, interest expression impacts how surrogates age visually (see Section 5.5.2).

5.5 The Agent's Generative Initiatives

The semantic and participant interest model components drive decision making in the generative threads of execution that comprise the agent initiatives. The resulting experience is consistent with Amar and Stasko's [2004] recent call for information visualization systems that incorporate uncertainty and respond to change. Two threads perform generative information collecting, while two others generate the temporal visual composition of surrogates. When the user enables their operation, these threads run gradually over time, generating evolving state within the information collection and the composition. They utilize the candidate pools and weighting metrics of the model. The metrics themselves utilize statistics of IDF and the interest model. In the current implementation, all selection operations choose the maximum, given the weights of surrogates in a candidate pool. In case of a tie, an element is chosen randomly. A prior system [Kerne 2001] 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.

5.5.1 *Generative Information Collecting Initiatives.* The generative information-collecting initiatives identify documents that are relevant to the user, download them, and extract clippings and semantics. One of this initiative's agent threads is a focused Web crawler [Diligenti et al. 2000] (see Figure 11, left). The other thread periodically selects the maximum weight image reference and downloads the associated image (see Figure 11, right). The actual downloading is not directly processed by the agent threads. Instead, a download monitor queue manages document download threads and timeouts. The software objects that manage these threads have been engineered to robustly handle I/O errors and timeouts.

The focused crawler agent periodically wakes up and selects the highest weighted candidate container. The agent processes a document, extracting metadata, hyperlinks, image references, and text chunks. Metadata fields are

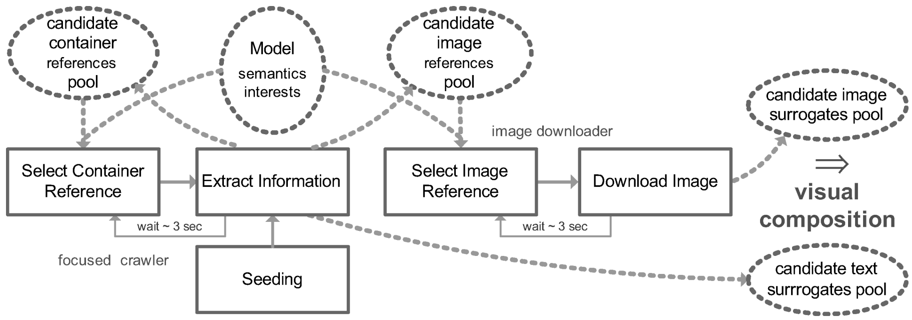

Fig. 11. Generative information-collecting agent initiatives.

filtered for noise, and, if suitable, added to the container object. Each hyperlink reference of a supported mime type, which refers to a document not previously encountered in the session, is used to form a container object. These are added to the candidate container references pool. All discovered images, except those whose dimensions resemble advertisements or spacers, are added to the candidate image references pool. Sometimes, the HTML for the document will specify image dimensions, enabling this filtering to be carried out in this early stage. Because so many text chunks are found, these are prefiltered; only those above a heuristically determined minimum weight threshold are added into the text surrogate candidate pools.

The time period for the agent's selection of candidate containers is typically once every few seconds. However, if the queue backs up, this can be increased, and, conversely, if the text surrogate candidate and image reference pools are too small, and the candidate containers pool is not empty, the wait time between downloads can be dynamically reduced. The operation of the second selection thread is simply to select the highest weight image references to download. Once downloaded, the image will be placed into the candidate image surrogates pool, unless, again, its dimensions suggest an advertisement or spacer.

In both of these selection operations, the agent operations choose the maximum weighed candidate. In case of a tie, an element is chosen randomly. Here we define the form for the surrogate weighting metric, $W_{surrogate}(S)$:

$$W_{surrogate}(S) = P(S) \times \underset{term \in S}{Avg}(W_{term}(term)) \times \underset{term \in C}{Avg}(W_{term}(term)),$$

where *P(S)* is participant interest measure for the surrogate S. The first average, over S, is computed over the terms in the composite term vector of the surrogate S, while the second average, over C, likewise represents the terms in the composite term vector of the container C.

We define the participant interest measure generally for χ, where χ can be a surrogate S, a container C, or a term:

$$P(\chi) = f(I_\chi, R_\chi),$$

where I_χ: is the interest level that the human participant has expressed in the surrogate, and R_χ: is a time-based boosting value that can increase the weight based on the recency of the participant's interactive interest expression. f is a weighting function. $W_{term}(term)$, the weight of each term is derived from the IDF and participant interest measure for the term:

$$W_{term}(term) = P(term) \times IDF_{term}.$$

The metric for the container, $W_{container}(C)$, is similar to that for the surrogate, except that the container weight only uses the container's composite TermVector weight, while the surrogate weight incorporates its own composite TermVector weight as a factor with the container's TermVector weight contribution.

$$W_{container}(C) = P(C) \times \underset{term \in C}{Avg}(W_{term}(term)).$$

5.5.2 *Generative Visual Composition Initiatives.* Through the information collection, the most significant surrogate candidates formed by the generative information-collecting initiatives serve as source material for 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 and image processing uses time as a continuous dimension for presenting the continuously evolving collection. In this way, *temporal visual composition* is a time-based visual medium, like video. There are two components of generative visual composition, each of which is accomplished by a separate thread: composition building, which brings surrogates into and out of the composition space and performs layout, and visual aging, which evolves layers of emphasis and deemphasis that promote readability.

Composition building is accomplished through an algorithm that prioritizes surrogates, based on weights in the model, and connects those that are related in the visual composition space, based on their term model semantics. This primary thread of visual composition agent initiative (see Figure 13) iteratively selects surrogate candidates for placement in the composition, and performs layout to cover that which is least important and cluster-related surrogates to visualize relationships. Next, the state of each surrogate already in the composition, which the user has not already expressed interest in (see Section 5.4.4, 5.3.2) is aged, to gradually reduce its visual prominence. This thread conducts a cycle of further steps to generate the layout. Through a series of such cycles, the layout emerges. Iterations through the cycle are 1 s apart by default; the user can change the rate or pause the process, using the tape recorder metaphor interface (see Section 5.3.5).

The composition space in combinFormation is (internally) divided into a matrix of rectangular cells (see Figure 12) that forms the basis of the $X{-}Y$ layout algorithm. Each cell keeps track 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.

During each iteration of the generative temporal composition algorithm, a new surrogate is added to the composition space. First, the highest weight surrogate candidate is selected from either the image surrogate candidates pool, or the text surrogates candidates pool. The choice of pools is determined by the

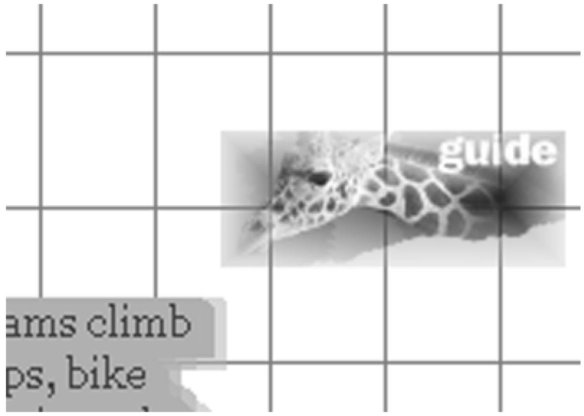

Fig. 12. The composition space is divided into a matrix of cells, with each cell keeping track of the surrogates that are placed upon it.

Image-Text Surrogates Mixer interactive control (see Section 5.3.6), through which the user selects a desired ratio of image and text surrogates, and the current ratio in the composition space. Before the new surrogate is added to the composition space, the surrogates already in the space are sorted based on their significance weights. Before the new element is added, the *Z* order is adjusted, stacking the most important surrogates on top.

Once a surrogate has been selected, a significance-based size, in grid cells, is assigned to the new surrogate based on its weight relative to those already in the composition. A range of possible sizes is established, in the form of a minimum and maximum number of grid cells, based on the grid granularity and size of the hot space (see Section 5.3.2), which is the subregion of the composition space that the agent is allowed to use. A linear interpolation is taken, in which the weight of the current element, as compared to the minimum and maximum weights of surrogates currently present in the composition space, is used as a proportional selector between the minimum and maximum number of grid cells.

5.5.3 *Semantic Clustering Layout: Relatedness Potential.* Each time the algorithm adds a new surrogate to the composition space, as per the loop to the right of Figure 13, the *relatedness potential* measure is computed for each cellular region of the visual workspace. The fundamentals of the relatedness potential algorithm were introduced in Kerne et al. [2005]. Subsequently, after several iterative design processes of deployment, testing, and user studies, the algorithm has been improved to reduce overlap in the placement of relevant surrogates and promote the better utilization of adjacent empty spaces. Here we present our improved algorithm, extending the previous research.

Relatedness potential measures how similar the information surrogates already present in each cellular region are to the new surrogate. At the root of the relatedness potential placement algorithm is the pairwise comparison of

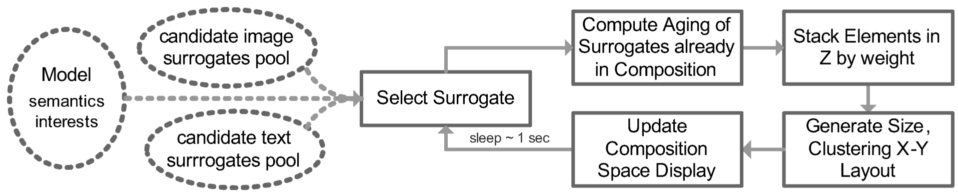

Fig. 13. The temporal composition thread generates the visual composition representation over time.

surrogates. It finds the set of cells in the space with the maximum relatedness potential in relationship to the new surrogate being placed into the space by the generative agent. Typically, in any vector space, the (cosine) similarity of two vectors is formed by their dot product. The problem with this calculation is that it assigns inverse magnitude to dissimilarities between vectors, because it divides by the product of the magnitudes of the vectors themselves. In order to be more sensitive to the presence of similarity, and to avoid giving inverse contribution to dissimilar dimensions, that is, terms that appear in one vector and not the other, we define the similarity computation between two surrogates λ and γ, only utilizing the common components:

$$Sim_{\lambda,\gamma} = \sum IDF(term_i),$$

where $term_i \in TermVector_\lambda \cap TermVector_\gamma$.

Each cell in the visual space has a list of the information surrogates that visually overlap with the cell's region. The relatedness potential for each cell α in relation to the new surrogate λ to be placed in the space is computed by iterating over each surrogate γ already in the composition space and each cell α that λ could overlap when it is placed in the space:

$$relatedPotential_\alpha = \sum_{\gamma \,\in\, \alpha} Sim_{\lambda,\,\gamma}.$$

The relatedness potential is a measure of how attractive this cell is for the next surrogate to be placed, based purely on semantics of the existing set of surrogates that overlap with this cell's region. The potential is computed for every cell in the visual space. We could select the cell with maximum potential to place the new surrogate. However, by simply selecting the cell with maximal potential, the result will be a layout in which similar surrogates are made to overlap and obscure each other. Instead, the current clustering algorithm is designed to position the similar surrogate adjacent to the most attractive cell, inasmuch as this is possible, rather than piling related surrogates on top of each other. That is, empty cells neighboring the maximally related cells should have greater relatedness potential for the new surrogate.

Additionally, the center and peripherae are treated differently. One reason for this is to disperse the positions of emerging clusters, encouraging their formation in distinct areas across the space. Another reason, for visual design purposes, is to reduce the amount of clutter in the center, so that the eye can rest there more easily. Thus, we make cells that are of greater distance from the center of the visual space be more attractive. This pushes the new surrogate that has no similarity with surrogates already in the space away from the similarity clusters that are already present.

In order to incorporate our design goals into the clustering algorithm, a spreading activation [Pirolli and Card 1999] scheme is again used. The cell activation potential of each cell α is spread to neighboring cells within a specified spread radius. In order to promote the use of empty space, and reduce the depth of piling, empty neighboring cells receive a raised activation, whereas non-empty neighboring cells receive a damped activation. The overall relatedness measure of each cell towards the next surrogate will be utilized to calculate the potential. The spreading activation schema is described by the equations below:

When the cell α is empty, the greater distance from the center distance of the visual space will have a bigger relatedness potential, $P_{\alpha,t+1}$. The clustering algorithm is invoked iteratively. One state serves as the basis for the next. t refers to the state of the composition space in the prior iteration of the visual composition algorithm, $t+1$ refers to the iteration we are now calculating. Let $P_{\alpha,t}$ represent the present relatedness potential and $P_{\alpha,t+1}$ represent the new relatedness potential we must calculate, with activation spreading from the previous state.

$$P_{\alpha,t+1} = P_{\alpha,t} + (Dist_{\alpha,center}/MaxDist) \times DistWeight.$$

When the cell α has a positive relatedness potential and the neighboring cell β is empty, the greater distance between the cells α and β (within a defined spread activation radius) and the smaller the number of surrogates overlapped across the cells $O_{\alpha,\beta}$, the higher raised activation the cell β will receive. This raising factor (rF) is a constant value, which is used to raise the relatedness potential of neighboring cells; m exponentially controls the magnitude of the impact of the raising factor based on the distance and surrogate overlaps.

$$\begin{aligned} P_{\beta,t+1} &= P_{\beta,t} + \{(1+rF) - rF^m\} \times P_{\alpha,t}, \\ m &= (Dist_{\alpha,\beta} \times O_{\alpha,\beta})^{-1}. \end{aligned}$$

When the cell α has a positive relatedness potential and the neighboring cell β is not empty, we define the relatedness potential for β such that the greater the distance between cells α and β (within a defined spread activation radius) and the smaller the number of overlapping surrogates in cells α and β ($O_{\alpha,\beta}$), the less damped activation the cell β will receive. The damping factor (dF) is a constant value less than one, which is used to gradually reduce propagation of relatedness potential to neighboring cells normalized by the above controlling factor, m. N_β refers to the value indicating the number of surrogates in the current cell, β.

$$P_{\beta,t+1} = P_{\beta,t} - \{dF^m + N_\beta - 1\} \times P_{\alpha,t}.$$

When the cell α has a negative or zero relatedness potential ($P_{\alpha,t}$), the relatedness potential of the neighboring cell β will be decreased by the number of surrogates in the cell β (N_β) and the overlapping degree with the cell α ($O_{\alpha,\beta}$).

The greater the distance between the cell α and β (within a defined spread activation radius), the smaller the decrease in relatedness potential for the cell β:

$$P_{\beta,t+1} = P_{\beta,t} + (N_{\beta} \times O_{\alpha,\beta}/Dist_{\alpha,\beta}) \times P_{\alpha,t}.$$

The aggregate potential for each cell is computed as the summation of the potential of all the cells covered by placing the upper left-hand corner of the new surrogate, given the significance-base size, in the cell under consideration. Thus, aggregate potential is computed for all candidate upper left-hand corner positions for the new surrogate. In case there is a tie, the final cell selected in randomly picked from the set of cells having the maximum aggregate potential.

Through this method, a new surrogate is periodically incorporated into the layout of the composition. To increase the legibility of the information composition as it develops over time, we have developed visual aging techniques.

5.5.4 *Visual Aging.* As the number of surrogates grows, the representation will tend to grow difficult to read. Surrogates will compete with each other visually for human attention. Tufte [1990] prescribed *layering* as a method for the presentation of complex information. Layering stratifies the presentation of information by using features such as color intensity and size to convey significance. As noted above, sizes are allocated based on weight. Visual aging is a technique in which the prominence of elements is diminished gradually, as they are present over time in the composition. This is a form of the read wear formulated by Hill et al. [1992]. Aging is better, for example, than moving elements or reducing their sizes, because those actions would take more of the user's attention. There is already plenty to think about, and we want the user to be able to focus on her/his own initiatives, rather than what the agent is doing. Thus, one thread serves to periodically impart visual aging onto those surrogates in the composition space, in which the user has not expressed positive interest. For text surrogates, visual aging is constituted by reducing opacity (alpha). For image surrogates, it is accomplished through gradual desaturation, in which colors seem to drain to gray.

Aged elements remain legible, so they can still catch the user's attention when they are particularly relevant. They are simply less prominent in the overall composition. The resulting method functions as an even more "generalized fisheye view" [Furnas 1986, 2006], in which saturation and opacity augment size as visual characteristics to which degree of interest is mapped. Full color and opacity are restored for the surrogate when it moves to the user's focus via brushing, that is, on mouse over. (This also results in the display of in-context details on demand.) Since expression of positive interest in a surrogate is a sign that it is more important to the user, such expression suspends visual aging for a given surrogate. This, again, is consistent with Furnas' degree of interest formulation.

The rate of aging is related to the rate of the visual composition agent, and the mean density of elements in the composition. Both of these can be set by the user. By default, the mean density is set to 0.7 elements/pixel, meaning that if there was no overlap, 30% of the composition would be white space. This default value is based on our intuitive sense of composition. The default rate is one new element every 1.2 s.

Eventually, there will be too many surrogates in the two-dimensional (2D) composition space. Even with visual aging, the density and extent of overlap will grow too great for legible readings. Thus, for each surrogate added by the visual composition agent, one will be removed. The removed surrogate takes age into account, but also takes into account semantics and the user interest profile. Future work will use zoomable interfaces to further increase the number of surrogates that the user can see, think about, and manipulate in the composition space at one time, before some are lost. Being able to scrub time will also contribute to the potential loss of important surrogates. However, the fact remains that the limits of human attention are inevitable. Even when interactive zooming and time travel are enabled, the user will often lack the time and attention to use them. There are guarantees neither about which surrogates will stimulate information discovery, nor about when. The most the program can do is effectively create opportunities for serendipity.

6. FIELD STUDY: THE DESIGN PROCESS IN UNDERGRADUATE EDUCATION

We conducted a field study to validate the use of the mixed-initiative information composition system for supporting creative processes in education. 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 to support their work on two invention-oriented information discovery assignments. The two mutually exclusive groups of students were both found to do better on the project when they used combinFormation to develop the prior work.

6.1 Experimental Method

In the field study, the mixed-initiative creativity support tool was utilized by students in two assignments in the interdisciplinary undergraduate environmental and design science course, The Design Process. The Design Process is an innovative interdisciplinary undergraduate course on creativity, innovation, and entrepreneurship. 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.

In The Design Process Course, the assignments engage students in invention-oriented 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.

We designed and conducted a comparative field study in The Design Process course. This is challenging, because unlike in a laboratory, where the experimenters can manipulate conditions, conditions in the course must be fair to all students. They also must meet the course's educational needs and negotiate established practices. We arranged an appropriate form of situated study. Students used either combinFormation or Google and Word to collect prior work for their Hybrid and Invention projects. They used the *re:mix* and *re:open* launch interfaces. For their prior work collection, those in the former condition turned in a saved composition space, in which each surrogate clearly refers to a source document. The others turned in a traditional bibliography. Half the class was assigned to use the mixed-initiative information composition system, combinFormation, for the prior work collection on The Hybrid, with the other half the class using Google to search and Word to assemble relevant results (Google+Word). For The Invention, the groups switched. Thus, each half of the class used combinFormation for one assignment, and Google+Word for the other. This was fair to students, while providing comparative conditions for study.

Each of the course's two sections has a Teaching Assistant. This TA assigns grades for all of the assignments performed by students in that section. The TA evaluates both components of the assignment—the prior work and the creative products—for both projects. The criteria and process for evaluating the creative products were established in The Design Process in prior years, before combinFormation's introduction there. For the creative invention products, the criteria involve originality, novelty, practicality, broad impact, and commercial transfer potential. We did not change this. For the prior work, The Design Process course and combinFormation research teams collaborated to establish criteria for evaluation: how informative, communicative, and expressive the collection is, and the extent of variety among the collected resources. For both components of both assignments, a new 1–5 scale was instituted for the study. This scale corresponds directly to the letter grades that are assigned in the course.

To further understand the experiment design, it is important to note that the relationship between The Design Process (TDP) course team of professors and teaching assistants, and the combinFormation research team, is one in which independent entities cooperate. Our sources of funding are entirely separate. While our goals overlap, they are established independently. TDP team's goals are based in creating an environment in which students learn to be creative; the combinFormation research team's goals are based in developing digital tools that promote creative experiences.

We continued the practice of having each course TA evaluate all assignments performed by students in his section. We discussed trying to mix up the grading, so that one TA would grade all of the prior work and the other would grade the creative products. This seemed to make sense from the standpoint of experimental validity, because it increases the independence of the evaluations. However, it does not make sense in the context of the course. Here, the TA assigned to a particular section gets to know many of the students. This knowledge becomes part of how assignments are graded in any course. It is a situated social practice. It does not make sense to separate evaluation of assignments from these grounded moorings. The course's TAs are responsible to the course, and its educational goals, and not to the combinFormation project, and its research goals. They have no stake in the success of our research, and a high level of accountability to their students. They are expert evaluators. Thus, this methodology is the only practical and most valid one for this research. Any process of evaluation by others would be more artificial and less grounded in truth. As part of the conditions of grounded field study, we are not able to control or even assess factors such as how much time students spend on their projects.

6.2 Quantitative Data and Results

We worked with the professor and teaching assistants for The Design Process to develop 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. A new 1–5 (5 corresponding to the highest grade) 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, and expressive the collection is, as well as the variety of the collected resources. Criteria for the actual inventions involve originality, novelty, practicality, broad impact, and commercial transfer ability. These measures are relatively objective, in that they are directly correlated and integrated with the evaluation process of the course. The TAs performed the evaluations as they were assigning grades based on the same criteria.

Approximately 81% of the students performed the Hybrid assignment. 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 Hybrid assignment creative product, and again, the result was statistically significant [$t(145) = 2.227$, $p = 0.028$] (see Figure 14, left).

The findings were similar for the Invention assignment. This time, 33.5% of the students used combinFormation, out of a total of 87.4% who did the assignment (see Figure 14, right). 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 versus 2.38 for Google+Word [$t(141) = 3.843$, $p < 0.001$]. For the actual Invention product, the scores were 3.41 versus 2.85 [$t(157) = 2.716$, $p = 0.007$]. The score differences both for the prior work and for the assignment creative products were statistically significant across both assignments.

From the field study, we found that combinFormation better supports students engaged in information discovery tasks in collecting and collecting prior work. According to the scores, the TAs found that representations of collections assembled in the medium of composition of image and text surrogates were better than textual lists for understanding, developing ideas, and the communication of meaning. Further, subsequent to developing prior work collections

Hybrid

	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
creative product	3.32	2.85		3.04

Invention

	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
creative product	3.41	2.85		3.06

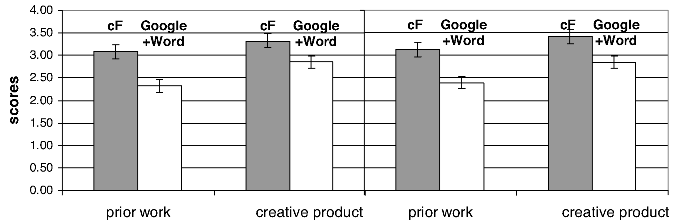

Fig. 14. Left: Student scores on the Hybrid assignment; right: Student scores on the Invention assignment. On both the prior work collection and the creative product, for both assignments, students performed significantly better when they developed their prior work using combinFormation's mixed-initiative composition space.

with combinFormation, students performed better on the actual Hybrid and Invention assignments than those who used Google+Word.

6.3 Qualitative Data

Students were asked to answer survey questions after they finished both of their assignments. The survey questions consisted of how many search queries they tried and what they were, whether students learned something new and interesting using the system, how they worked with the cool space and the hot space, what was easy and helpful in the more familiar and linear format, Google+Word in comparison with combinFormation, and what they would like combinFormation to do that it does not do now. This section presents an analysis of the qualitative data, students' answers to the survey questions.

6.3.1 *Combining Search Queries.* When launching combinFormation, the students combined an average of 3.25 search queries during each invocation of the *re:mix* interface. Only 7.37% of students utilized a single search. This result shows that a single search query was not adequate to address the information needs of the invention prior work collection information discovery task. Students saw a need to combine search queries. combinFormation's ability to mix the results of multiple searches was found to be useful.

6.3.2 *Learn from combinFormation.* Students were able to come up with new ideas by putting together found information from different sources in the composition space. The qualitative data shows how the combinFormation system helped students develop new and creative ideas by giving them space in which to put together found information.

User ID	Qualitative Data
P107	Your program is able to connect many websites to one html document. Then you can open and click on a picture or a set of words and it will take you to the website. *It is really cool to make one document that connects many Web sites.*
P116	I learned that computers have already been placed on grocery carts by accessing a link to newspaper article. However, the purpose of the computers was different, and thus my idea was still novel. *I liked how similar innovations also came up in the search, allowing me to work their components into my invention.*
P147	When using this program, I learned that *after you have all your information gathered in one area, it is easier to play around and come up with something new. It takes in all the different ideas in our brains, and just basically puts it on the table for us.*
P223	I just learned a little bit more about being able to *put two completely random thoughts together and coming up with a fun and crazy idea.*

A goal of the combinFormation agent is to retrieve relevant information for students and to generate compositions that are new and unexpected. The qualitative data demonstrates how the agent works for students. It helped them enjoy many different and fun ideas and stimulated their creation of new ideas. Aside from using combinFormation for assignments, P154 introduced the system to her brother, and enjoyed their moments of interacting together with combinFormation. Another student (P199) found the image and text surrogates representation in combinFormation helpful because it reduced the work of reading whole documents and allowed them to easily connect different ideas.

User ID	Qualitative Data
P176	It was helpful in stimulating thoughts that I hadn't previously considered. The constant access to household budgets was one of these subjects.
P154	Well the topic I did was one from the ideas challenge which I wasn't too interested in. But my brother did one on topics that he was interested in, and we had a lot of fun looking into all of the items.
P155	Yes! Top secret weaponry, creativity enhancement, cloaking, conspiracies, and also that it is actually used and patented was something that I would never have known by conventional means.
P166	I didn't realize how a few little words could generate so many ideas from so many different perspectives.
P179	For some reason I got a really funny link to selling underwear in a vending machine when I searched vending machine. I laughed for a while. I also learned that the vending machine industry is in the billions of dollars in America.
P199	Yes, like I said before, the current and future projected applications of ambient intelligence were very interesting, especially as it applied to the medical field and to our everyday lives. It was something I did not know very much about, but that there was a lot of info on. The really neat thing about the program was that instead of reading tons of long articles to find ideas or get concepts, it pulls quick quotes that allow you to connect lots of ideas instead of becoming lost or bogged down in just a few concepts.

6.3.3 *Cool Space and Hot Space.* The qualitative results showed that the cool space and the hot space helped them collect found information and formulate creative ideas. Students found it easy to understand and use the cool space and the hot space. They formulated and developed their new ideas by the iterative process of collecting information in the cool space, expanding the cool space, manipulating and connecting surrogates visually, controlling the agent, and collecting more interesting information.

User ID	Qualitative Data
P139	I started out with a small cool space and once it got filled I expanded it. I put the hot space settings on as fast as they would go, and would just systematically pause it to see if there was anything I needed or wanted to click on. It was very good.
P199	I found myself putting interesting quotes and pictures in the cool space or the center box. I liked layering the quotes and ideas and seeing how they kind of could connect or link. When you place them in groups and layers, then it is neat to see what new ideas you can have. I think a cool feature would be a way to visually connect ideas/concepts and explain the connection. Like the layering of images to show an idea be able to do it with words, and, word and pictures. The hot space, I just left things out there and liked to see how the things piled up, and then liked to sort through all the accumulated ideas with a pause. I liked to run the hot space very fast b/c the speed of the ideas helped me to generate random and crazy connections.
P241	I glanced over everything that appeared in the hot space. Anything remotely relevant to my project I moved into the cool space, where I organized my ideas and made the final decisions of what I wanted in my finished project.
P161	I brought stuff in and out of the cool space to emphasize what was important. It allowed me to corroborate my thoughts and see where I was going with certain ideas, depending on how they were arranged in the different spaces.
P122	Usually, I placed elements in the cool space to keep them there for awhile, as I observed elements popping up in the hot space. Once I formed a creative idea, I would pause combinFormation and use the items in the cool space to form a visual collage of what I wanted. I would resize them, and change the backgrounds. Once I had completed working on the cool space for awhile, I would start combinFormation again and look for more items in the hot space.
P209	I slowed down things and had it at a pretty low level. I continually looked around and if I saw anything that caught my eye I'd put it in the cool space and replaced if something better came along. Really easy to use.

6.3.4 *More Familiar with Linear Format and Google+Word.* Some participants said that they are familiar with Google+Word, so a new tool, combinFormation, didn't seem to give them good results initially. They liked Google+Word because they were more accustomed to using it. A participant mentioned that it takes longer to learn how to use something when it comes to computers. In addition, they preferred the linear format to spatial representation. They liked the static representation of information better than a transient visualization. Another participant found combinFormation a bit confusing and overwhelming because of the spatial, dynamic, and transient representation of information. One participant wanted combinFormation to provide more relevant information.

User ID	Qualitative Data
P112	My brain thinks linearly so a bunch of blurbs and pictures randomly popping up and then disappearing is a little frustrating. With Google I also get all of my results at once and can go through the ones I want. With combinFormation you have no idea what you're getting.
P100	I think combinFormation could be very helpful if it had more pertinent information instead of just bits and pieces that may have only mentioned the searched terms once as a minor detail. Google and Word allow me to quickly discard the unwanted stuff fast. Also, Google doesn't lose things. combinFormation has things disappear unless you click on them, and if I wanted something later, I didn't know where to find it.

6.3.5 *Compare Google+Word with combinFormation.* Participants compared Google+Word with combinFormation. They found that Google was somewhat faster for search, because it was easier finding the exact pictures on Google rather than waiting and hoping the needed pictures would arrive with combinFormation. However, with combinFormation they could spend more time being creative, combining ideas, and searching for what they had pictured in their minds. Also, the images and text presented on combinFormation sparked their intellectual interest. Using Google, they could direct searching to their specific information need. combinFormation could be used to search for an existing idea, but it may give new, fresh, and more unique ideas. Because combinFormation sometimes finds the unexpected, they liked it better for developing an idea, but Google better for research into an already formed idea.

6.3.6 *What Was Helpful in combinFormation.* Students were asked what was helpful in combinFormation to accomplish their assignments. Students were satisfied with the multiple search results generated by combinFormation and they liked image and text representation of search results. Students also mentioned that it was easy to control the agent and most of them agreed that the combinFormation system provided and afforded the means to develop new ideas by combining different perspectives of found information in the composition space.

User ID	Qualitative Data
P97	It generated ideas for me! I didn't need to form thoughts, because they came to me on screen. The link was user friendly and easy to use.
P101	The search functions were very powerful. The engine and the interface worked well preliminarily. The system quickly provided very good results.
P210	All you had to do was enter in the topic you wanted to research, and combinFormation did the rest. The ideas that the program searched out were very helpful and made the research easy. The pause and play button was easy to use, and it was also easy to drag items into the cool space.
P132	The agent looked through more sites than I would have been able to/would have had patience to look through.

Continued on next page.

P203	Finding the right Web sites for your information or idea and putting it all together.
P113	It was easy to see what ideas came up and it searched well. I like having pictures as well as words when searching through information. The command template was simple and worked well enough I knew how to do everything that I wanted to do with the program.
P157	What made combinFormation easy was the opportunity to use a large mix of search phrases. It gave you a wide vary of picture and text options to put in the cool space.
P116	I liked how it was easy to manipulate the elements. The pause and speed control features were very handy. The clear the hot space feature was also very handy because it allowed for the quick removal of all of the unnecessary elements.
P122	In combinFormation, it was easy to use the ”hand” tool to grab, drag, and resize elements. I liked the idea of combining different ideas to create a visual representation of something that is already in your head, or of something you discover while using the program. I also like the ability to speed up and slow down the program.
P139	It was very easy to pause and drag and change the appearance of the elements. It represented an array of information and was pretty self-explanatory if you didn't understand something.
P150	It is a relatively easy way to look at one idea from many different perspectives.
P227	It was nice to just sit back and watch the program tie together all of these concepts and ideas that no one would ever really put together on their own. I'd say that the program is great as a brainstorming tool, and it was interesting to see what kind of pictures and text would pop up.
P226	It was easy to see your technology take off in new directions as things popped up. It was also easy to focus on a certain idea or eliminate some by selecting things involving them as desired.

6.3.7 *Future combinFormation.* Students were asked to give us feedback about new features in combinFormation. Many students suggested having the means to reformulate their search queries in the middle of their combinFormation sessions. They also would have liked to control the sites and pages that the agent may crawl, and some of them wanted the agent to automatically reformulate search queries using their initial queries.

User ID	Qualitative Data
P132	Be able to change or alter searches half way through your combinFormation so that you can keep what you already have but improve on it as the agent triggers another idea.
P149	Maybe have a box somewhere that you could add new seeds on as your search goes on.
P144	Is it possible to limit the sites it is searching to only certain types, such as research sites? It might already do this, but I wasn't sure.
P237	I think combinFormation should give an even more refined search for exactly what the user specifies, and possibly, if the user already has an idea of a hybrid product, to allow the program to search for this combined idea as well as the individual components of the idea.
P243	A better way to control what the engine spits out in the form of results. It appeared that it simply brought up the results for the individual key words. I think a valuable option would be to create a way that the engine looks for results containing multiple seeds.

Students suggested advanced manipulation features such as sketching tools and cropping. One student (P113) wanted a way to increase the whole space size because, when she/he increased the cool space, the hot space would be eaten up.

User ID	Qualitative Data
P110	I wish I could trim exactly what I wanted. That I could take a pencil and draw around what I wanted and crop out the rest.
P113	I would like to be able to resize the entire window once combinFormation has already launched. Increasing the size of the cool space can eat up the entire hot space on half the screen and it's hard to process what all is going on when the images and words change every time because there's no space.
P155	Also if you could put word art like that of Microsoft Word to maybe add a little more visual art.

Some students suggested some interesting ways to socially interact with others using combinFormation. One student (P154) mentioned creating a collaborative composition space using combinFormation, and another student (P203) wanted a way to email interesting links to others while she was creating the combinFormation composition space.

User ID	Qualitative Data
P154	It would be cool if you could have review sessions (a saved file that was open to view by others and could be added on to). The new add-ons would be marked in a different color to show that they were add-ons. Just to allow for input by others. All in all, I really enjoyed the program. Good luck!

Students expressed interest in having multimedia files such as music and movies in combinFormation. They would like a way to mix and fuse those files in a space.

User ID	Qualitative Data
P155	Music would be a great function to have and maybe a mixer to show a fusing of music and sound.
P159	I think it would be neat if you could combine videos, so that, for example, you could have a cool space with Cinderella fighting Rambo. You could also use the combined videos for educational purposes.
P210	Maybe the program could find links that play video and audio clips from the Web that pertain to your topic or field of study. This would give users more options when they are doing research for a project.

Students wanted to have more visual feedback about their interest expression, information relevance, and relationships between links and information. They would like to see visual representation of connections and linkage of information relationships.

User ID	Qualitative Data
P160	Have a graded importance system of 1–10 so that there is a greater variance between how important certain items are.
P155	Perhaps those inquiries that I had shown interest in could be highlighted to show progress not so much to sway my creativity but just to let me know they were there easier.
P199	However, the implications of THIS technology is amazing and can really change the way people use the internet, esp. if it were to applied to other uses than just combinFormation, but using the Web crawler to simply find lists of sites; then using the crawler to cross-connect those lists to find the links/connections between lists of things that may or may not be unrelated, showing relationships between things that already exist easily.
P165	Show more linkages between the search queries—that would help me generate ideas.

7. CONCLUSION

The field study results in The Design Process Course demonstrate that combinFormation's visual, mixed-initiative composition space method for searching, organizing, and integrating information promotes creative processes of information discovery in education. From the quantitative results in tandem with qualitative data from questionnaires, we conclude 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. This is true despite their lack of familiarity with combinFormation, as compared to Google and Word. Adding the temporal and agent initiative dimensions to search increases the breadth of the set of information resources the human is exposed to.

The use of complementary image and text surrogate representations promotes cognition of this larger set of representations. The affordances of these surrogates in the composition space enable the manipulation of recombinant information, which contributes to participants' developing ideas. Further research is necessary to clearly measure the role of recombinant information in the creative process, by isolating it as a factor in controlled experiments. The affordances of the composition space surrogates also enable quick expression of interest, which tune the semantic model, and thus the agent's performance, to retrieve more information relevant to the user's emerging sense of invention. This is a solution to McNee et al.'s [2003] elicitation of the relevance feedback problem. Further, traversal of an emergent space of relevant possibilities wider than initial search results promotes information discovery, enabling students to create better hybrids and inventions.

The mixed initiatives of the combinFormation approach have proven valuable. The students made great use of the program's generative facilities. They combined searches together. They used the tape recorder transport and cool space to negotiate control, treating surrogates presented by the agent as suggestions, and building compositions of these suggestions in the cool space. Once elements were moved to the cool space, direct-manipulation composition-authoring functionalities were used extensively.

Beyond the research reported here, we continue to develop our relationship with The Design Process Course, investigating needs for learning and creative processes, 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 will 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. We will continue to develop direct-manipulation interactive mechanisms for authoring the composition form and directing the agent. As per users' requests, this involves making the agent's state visible, and enabling the iterative reformulation of search queries. We will also continue to develop the model and generative algorithms. We need to discover better ways to extract relevant surrogates, and make their individual and ensemble meaning visible. Further, we intend to investigate support for information discovery in other research task contexts, such as graduate students developing papers and theses.

One goal for future work is to more automatically model the user's context and interests. Many difficult problems arise in connecting information semantics with visualization and interaction. For example, we have tried to simply derive semantic intentions from composition authoring actions. In an earlier version of the software, the cut operation was assumed to mean a negative expression of interest. In practice, this did not work, because users sometimes ignore semantics and build their compositions for visual reasons. For the same reason, we cannot simply assume that when the user moves elements into proximity that they are semantically connected. Nonetheless, in future work we need to discover new methods for inferring the user's intent. We hypothesize that careful analysis of practice, followed by the contextualized development of sophisticated pattern recognition methods, will be sufficient to make significant headway. This is an exciting area for future research.

We are particularly interested in the relationships between support for information discovery, and for other related paradigms, such as information retrieval [Salton and McGill 1983], information foraging [Pirolli and Card 1999], information seeking [Marchionini 1995], and exploratory search [White et al. 2006]. While these paradigms overlap, they also frame research differently. Among these framings, the common goal is to find relevant information. Yet while foraging theory is useful, it does not strike us that humans engaged in creative tasks function to optimize the cost of information acquisition, as Pirolli and Card suggested [1999]. Information seeking and exploratory search go further by giving important attention to nonlinearities in the user's processes of specifying and fulfilling information needs. The present research is specifically framed to spotlight intellectual and creative tasks that also require putting together, that is, *composing*, the ideas embodied by relevant information clippings and their source documents. The human need to understand and create connections, to compare and abstract, as well as to choose, is emphasized. People, in general, and creativity, in particular, are not entirely rational and linear. Accounting for changes in the user's sense of what is relevant, in the course of work on a task, is part of supporting information discovery processes of comparison, abstraction, and ideation. Providing representations that promote information discovery processes is also essential. Further, we identify the value of expanding the space of encountered information, to overcome *fixation* [Smith 1994] and promote *cognitive restructuring*. We wonder what balances between finding exactly what the user has specified, and stimulating her/him with a somewhat larger space of related but perhaps unexpected work, will be most effective. This is expected to vary with task contexts. As the model and information collection methods grow more precise, it will be necessary to discover human-centered means for supporting adjustment of the breadth and range of information exploration. Scaling the representation further, with time scrubbing and zooming will be necessary to help the user make sense of even larger collections. Unlike the user's attention, computer power and network bandwidth continue to become less expensive. Thus, the incremental costs for locating a broader set of relevant information resources become less consequential. But the limits of cognitive attention remain essentially constant, dictating that we need to discover methods that balance factors to enable users to see, understand, and connect an optimal set of relevant and stimulating information resources, to think outside of the box while performing information discovery tasks.

ACKNOWLEDGMENTS

Thanks to the students, TAs, and Professor Rodney Hill of Texas A&M University ENDS 101/200 for working with us.

REFERENCES

AMAR, R. AND STASKO, J. 2004. A knowledge task-based framework for design and evaluation of information visualizations. In *Proceedings of the 8th conference on Information Visualization* (InfoViz), 143–150.

BADDELEY, A. D. 1992. Is working memory working?, *Quart. J. Exp. Psych. 44A*, 1–31.

BALABANOVIC, M. 1998. Learning to Surf: Multiagent systems for adaptive Web page recommendation. Ph.D. dissertation. Stanford University Department of Computer Science, Stanford, CA.

BALDONADO, M. Q. W. AND WINOGRAD, T. 1997. SenseMaker: An information-exploration interface supporting the contextual evolution of a user's interests. In *Proceedings of the ACM CHI.* 11–18.

BELKIN, N. 1982. ASK for Information Retrieval: Part I. Background and Theory. *J. Documentat. 38*, 2, 299–304.

BIER, E. AND PERER, A. 2005. Icon Abacus and Ghost Icons. In *Proceedings of the 5th ACM/IEEE Joint Conference on Digital Libraries* (JCDL). 404–404.

BREAZEAL, C. 2002. *Designing Sociable Robots*. MIT Press, Cambridge, MA.

BURKE, M. 1999. *Organization of Multimedia Resources: Principles and Practice of Information Retrieval*. Gower Publishing Company, Hampshire, U.K.

CUTRELL, E., DUMAIS, S. T., AND TEEVAN J. 2006. Searching to eliminate personal information management. *Commun. ACM 49*, 1, 58–64.

CUTTING, D. R., KARGER, D. R., PEDERSON, J. O., AND TUKEY, J. W. 1992. Scatter/gather: A cluster-based approach to browsing large document collections. In *Proceedings of ACM IR*. 318–329

DILIGENTI, M., COETZEE, F., LAWRENCE, S., GILES, C. L., AND GORI, M. 2000. Focused crawling using context graphs. In *Proceedings of the 26th International Conference on Very Large Databases* (VLDB). 527–534.

DING, W., MARCHIONINI, G., AND SOERGEL, D. 1999. Multimodal surrogates for video browsing. In *Proceedings of the 4th ACM Conference on Digital Libraries* (DL). 85–93.

DUMAIS, S. T., CUTRELL, E., CADIZ, J. J., JANCKE, G., SARIN, R., AND ROBBINS, D. C. 2003. Stuff I've Seen: A system for personal information retrieval and re-use. In *Proceedings of the 26th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 72–79.

EISENSTEIN, S. M. 1975. *The Film Sense*. Harcourt Brace Jovanovich, New York, NY.

FINKE, R., WARD, T., AND SMITH, S. M. 1992. *Creative Cognition*. MIT Press, Cambridge, MA.

FURNAS, G. W. 1986. Generalized fisheye views. In *Proceedings of the ACM Computer Human Interaction*. 16–23.

FURNAS, G. W. 2006. A fish-eye follow-up: Further reflections on focus + context. In *Proceedings of the ACM Computer Human Interaction*. 999–1008.

GLENBERG, A. M. 2002. The Indexical hypothesis: Meaning from language, world, and image. In *Working with Words and Images.* Ablex, Westport, CT.

GLENBERG, A. M. AND LANGSTON, W. E. 1992. Comprehension of illustrated text: Pictures help to build mental models. *J. Mem. Lang. 31*, 2, 129–151.

HEARST, M. A. AND PEDERSON, J. O. 1996. In *Proceedings of ACM IR*. 76–84.

HILL, W. C., HOLLAN, J. D., WROBLEWSKI, D., AND MCCANDLESS, T. 1992. Edit wear and read wear. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3–9.

HORVITZ, E. Uncertainty, action, and interaction. In *pursuit* of mixed-initiative computing. In *Intelligent Systems*. IEEE Computer Society Press, Los Alamitos, CA.

HUYNH, D., MAZZOCCHI, S., AND KARGER, D. 2005. Piggy Bank: Experience the semantic Web inside your Web browser. In *Proceedings of the 2005 International Semantic Web Conference*. 413–430.

INTERFACE ECOLOGY LAB. 2007. combinFormation. `http://ecologylab.cs.tamu.edu/combinFormation/`.

JONES, W., DUMAIS, S., AND BRUCE, H. 2002. Once found, what then?: A study of 'keeping' behaviors in personal use of Web information. In *Proceedings of the ASIST Annual Meeting*. Vol. 39. 391–402.

JONES, W., PHUWANARTNURAK, A. J., GILL, R., AND BRUCE, H. 2005. Don't take my folders away!: Organizing personal information to get things done. In *CHI'05 Extended Abstracts on Human Factors in Computing Systems*. ACM Press, New York, NY, 1505–1508.

KELLY D. AND TEEVAN, J. 2003. Implicit feedback for inferring user preference: a bibliography. *ACM SIGIR For. 37*, 2, 18–28.

KERNE, A. CollageMachine, a Model of Interface Ecology, Ph.D. dissertation. New York University, New York, NY.

KERNE, A. AND KHANDELWAL, M. 2003. Publishing evolving metadocuments on the Web. In *Proceedings of the 14th ACM Conference on Hypertext and Hypermedia*. 104–105.

KERNE, A., KOH, E., DWORACZYK, B., MISTROT, J. M., CHOI, H., SMITH, S. M., GRAEBER, R., CARUSO, D., WEBB, A., HILL, R., AND ALBEA, J. 2006. combinFormation: A mixed-initiative system for representing collections as compositions of image and text surrogates. In *Proceedings of the 6th ACM/IEEE Joint Conference on Digital Libraries* (JCDL). 11–20.

KERNE, A., KOH, E., SUNDARAM, V., AND MISTROT, J. M. 2005. Generative semantic clustering in spatial hypertext. In *Proceedings of the 2005 ACM Symposium on Document Engineering*. 84–93.

KERNE, A., MISTROT, J. M., KHANDELWAL, M., SUNDARAM, V., AND KOH, E. 2004a. Using composition to re-present personal collections of hypersigns. In *Proceedings of Computational Semiotics in Games and New Media* (CoSIGN).

KERNE, A. AND SMITH, S. 2004. The information discovery framework. In *Proceedings of the 2004 Conference on Designing Interactive Systems* (DIS). 357–360.

KERNE, A., SMITH, S. M., MISTROT, J. M., SUNDARAM, V., AND KHANDELWAL, M. 2004b. Mapping interest and design to facilitate creative process during mixed-initiative information composition. In *Proceedings of Interaction: Systems, Practice and Theory, Creativity & Cognition Symposium*. 1–25.

KERNE, A. AND SUNDARAM, V. 2003. A recombinant information space. In *Proceedings of the Computational Semiotics in Games and New Media* (CoSIGN). 48–57.

KERNE, A., TOUPS, Z., DWORACZYK, B., AND KHANDELWAL, M. 2008. Componentized XML-Java data binding framework connects software engineering with efficiency. Submitted to OOPSLA.

KHANDELWAL, M., KERNE, A., AND MISTROT, J. M. 2004. Manipulating history in generative hypermedia. In *Proceedings of the Fifteenth ACM Conference on Hypertext and Hypermedia*. 139–140.

KOH, E. AND KERNE, A. 2006. I keep collecting: College students build and utilize collections in spite of breakdowns. In *Proceedings of European Conf on Digital Libraries* (ECDL). 303–314.

LIPPARD, L., ED. 1971. *Dadas on Art*. Prentice-Hall, Englewood Cliffs, NJ.

MAIER, N. R. F. 1931. Reasoning in humans ü: The solution of a problem and its appearance in consciousness. *J. Compar. Physiol. Psych. 12*, 181–194.

MARCHIONINI, G. 1995. Information seeking in electronic environments. Cambridge Series on Human-Computer Interaction. vol. 9. Cambridge University Press, Cambridge, U.K.

MARSHALL, C. C. AND BLY, S. 2005. Saving and using encountered information: Implications for electronic periodicals. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (SIGCHI). 111–120.

MARSHALL, C. C. AND JONES, W. 2006. Keeping encountered information. *Commun. ACM 49*, 1, 66–67.

MARSHALL, C. C., AND SHIPMAN, F. M. 1994. VIKI: Spatial hypertext supporting emergent structure. In *Proceedings of the 1994 ACM European Conference on Hypermedia Technology* (ECHT). 13–23.

MAYER, R. E. AND MORENO, R. 2002. Aids to computer-based multimedia learning. *Learn. Instruct. 12*, 107–119.

MCNEE, S. M., LAM, S., KONSTAN, J. A., AND RIEDL, J. 2003. Interfaces for eliciting new user preferences in recommender systems. In *Proceedings of User Modeling*. 178–188.

MILLER, G. A. 2006. WordNet. Princeton University Cognitive Science Laboratory, Princeton, NJ. http://wordnet.princeton.edu/.

MORRISON, J. B., PIROLLI, P., AND CARD, S. K. 2001. A taxonomic analysis of what World Wide Web activities significantly impact people's decisions and actions. In *CHI' 01 Extended Abstracts on Human Factors in Computing Systems*. ACM Press, New York, NY, 163–164.

NARDI, B. AND BARREAU, D. K. 1997. "Finding and reminding" revisited: Appropriate metaphors for file organization at the desktop. *SIGCHI Bull. 29*, 1.

OXFORD ENGLISH DICTIONARY. 1992. 2nd ed., on compact disk. Oxford University Press, Oxford, U.K.

PIROLLI, P. AND CARD, S. K. 1999. Information foraging. *Psychol. Rev. 106*, 4, 643–675.

PORTER, M. F. 1980. An algorithm for suffix stripping. *Program*, *14*, 3, 130–137.

ROCCHIO, J. 1971. Relevance feedback in information retrieval. In *The Smart Retrieval System-Experiments in Automatic Document Processing*. Prentice-Hall, Englewood Cliffs, NJ, 313–323.

SALTON, G. AND MCGILL, M. J. 1983. *Introduction to Modern Information Retrieval*. McGraw-Hill, New York, NY.

SCHRAEFEL, M. C., ZHU, Y., MODJESKA, D., WIGDOR, D., AND ZHAO, S. 2002. Hunter gatherer: Interaction support for the creation and management of within-Web-page collections. In *Proceedings of the Eleventh International Conference on World Wide Web (WWW)*. 172–181.

SHIPMAN, F., HSIEH, H., MOORE, J. M., AND ZACCHI, A. 2004. Supporting personal collections across digital libraries in spatial hypertext. In *Proceedings of 6th ACM/IEEE Joint Conference on Digital Libraries* (JCDL). 358–367.

SHNEIDERMAN, B. AND MAES, P. 1997. Direct manipulation vs. interface agents. *ACM Interact. 4*, 6.

SIMON, H. A. 1971. Designing organizations for an information-rich world. In *Computer, Communications, and the Public Interest.* Johns Hopkins Press, Baltimore, MD, 37–72.

SMITH, S. M. 1994. Getting into and out of mental ruts: A theory of fixation, incubation, and insight. In R. J. Sternberg and J. Davidson, Eds. *The Nature of Insight*, MIT Press, Cambridge, MA, 121–149.

SHAH, J. J., SMITH, S. M., VARGAS-HERNANDEZ, N., GERKENS, R., AND WULAN, M. 2003. Empirical studies of design ideation: Alignment of design experiments with laboratory experiments. In *Proceedings of DETC 2003: ASME International Conference on Design Theory and Methodology*.

SPIES, W., ED. 2006. *Max Ernst: Life and Work: An Autobiographical Collage.* Thames & Hudson, London, U.K.

SPOOKY, DJ THAT SUBLIMINAL KID. 2004. *Rhythm Science*. Mediawork / MIT Press, Cambridge, MA.

SUN MICROSYSTEMS. 2006. Java. http://java.sun.com.

TUFTE, E. 1990. *Envisioning Information*. Graphics Press, Cheshire, CT.

WHITE, R. W., KULES, B., DRUCKER, S. M., AND SCHRAEFEL, M. C. 2006. Supporting exploratory search. *Commun. ACM 49*, 4, 36–39.

WILDEMUTH, B. M., MARCHIONINI, G., YANG, M., GEISLER, G., WILKENS, T., HUGHES, A., AND GRUSS, R. 2003. How fast is too fast?: Evaluating fast forward surrogates for digital video. In *Proceedings of the 3rd ACM/IEEE-CS Joint Conference on Digital Libraries* (JCDL). 221–230.

WOODRUFF, A., FAULRING, A., ROSENHOLTZ, R., MORRISON, J., AND PIROLLI, P. 2001. Using thumbnails to search the Web. In *Proceedings of ACM CHI.* 198–205.

WOODRUFF, A., ROSENHOLTZ, R., MORRISON, J., FAULRING, A., AND PIROLLI, P. 2002. A comparison of the use of text summaries, plain thumbnails, and enhanced thumbnails for Web search tasks. *J. Amer. Soc. Inform. Sci. Tech. 53*, 2, 172–185.

W3C. 2000. Document Object Model (DOM) Level 2 Core Specification. `http://www.w3.org/TR/2000/REC-DOM-Level-2-Core-20001113/`.

YAMAMOTO, Y., NAKAKOJI, K., NISHINAKA, Y., ASADA, M., AND MATSUDA, R. 2005. What is the space for?: The role of space in authoring hypertext representations. In *Proceedings of ACM Hypertext*. 117–125.

Received January 2007; revised January 2008; accepted March 2008