Integrating Layout into Multimedia Data Retrieval

Isabel F. Cruz

Kimberly M. James

David C. Brown

Computer Science Department
Worcester Polytechnic Institute
Worcester, MA 01609, USA
{ ifc | beez | dcb }@cs.wpi.edu

Abstract

The Delaunay^{MM} project is concerned with creating a customizable layout-driven approach to querying multimedia digital libraries. Delaunay^{MM} layouts can be fully specified from scratch by a layout designer to accommodate different classes of users and tasks. In this paper, we overview the main aspects of the Delaunay^{MM} project and present our approach to the layout specification and enhancement of virtual documents.

Introduction

With an increase in the number of users daily, the World Wide Web has become an indispensable technology. With increasingly diverse user populations and available technologies, the value of reliable searching and information navigation mechanisms is becoming more significant. The average user spends between five and fifteen minutes searching on the web before finding his or her first piece of useful information, with expert users completing the task in less than five minutes (Georgia Tech Research Corporation, 1998).

The gap between expert and novice capabilities is staggering. To reduce the overall time for searching, in both the novice and expert population, search interfaces must combine powerful search and refinement capabilities with intelligible views so as not to overwhelm the user.

Delaunay^{MM} presents users with a unique search interface, supporting both pre- and post-query refinement for novice and expert users, as well as user-specified layout of query results. These capabilities are coupled with a framework that supports multiple information repositories and media types (image, video, text, and audio). The primary application of this work has been with digital libraries, but the principles and interface methods can be extended to other disparate information repositories such as the World Wide Web. There are three concepts that make our approach unique:

Pre- and user-defined virtual document templates for defining spatial layout. To address the needs of different users and a variety of applications, the spatial layout of the multimedia objects can be visually specified on a template, the *virtual document template*, containing "space-holders" for these objects. Users may either select a predefined template or design one to match their needs to create documents, called *virtual documents*, assembled using multimedia objects from distributed sources. Documents are *virtual* in the sense that are created to meet a particular user needs and possibly cannot be found in that form elsewhere. Users can, however, save them for later perusal.

Pre- and post-query refinement. Through interactive interfaces both expert and novice users are able to refine their queries easily. Experts may refine their queries based on the attributes of the information repository and novices may use the Search Assistant to narrow their queries to relevant information repositories for pre-query refinement. Both novice and expert users can use the Refinement Assistant to resubmit queries using a representative set of results for post-query refinement.

Nested query boxes for specifying browsable groupings. In the generated virtual document, browsing of the retrieved objects associated with the visual representations is determined by the semantics associated with *query boxes*, which are areas of the display that hold sets of images (or other multimedia objects) that can be visualized as a group. Associated with a query box there is a query that specifies the objects to retrieve. Query boxes can be *nested*. Roughly, if a query box is nested at a more inner level than another "outer" object it is associated with, browsing of that query box will display a group at a time without changing the display of the outer object (like displaying the parent in a subtree and subsets of its children, one at a time). The user can therefore specify what information to retrieve, as well as how that information is to be displayed. The end result is an automatically generated, browsable virtual document of multimedia objects that has been customized to the exact specifications of the user.

The flexibility provided by this framework effectively supports a multitude of users, as well as information repositories, thus providing a global interface for information retrieval and presentation.

In the next section, we describe the architecture and interaction with Delaunay^{MM}. Then, we present in greater detail the customizable layout feature of the system. The fourth section covers the usability studies that we performed. The following section discusses techniques for enhancing layouts in Delaunay^{MM}. We conclude the paper mentioning ideas for future work.

Architecture of Delaunay^{MM}

Figure 1 provides an overview of the Delaunay^{MM} architecture for querying and presenting multimedia data retrieved from digital libraries (Cruz & James, 1999b). This new architecture builds on the visual querying and layout design methods previously developed (Cruz & Lucas, 1997) (Cruz & Lucas, 1999) .

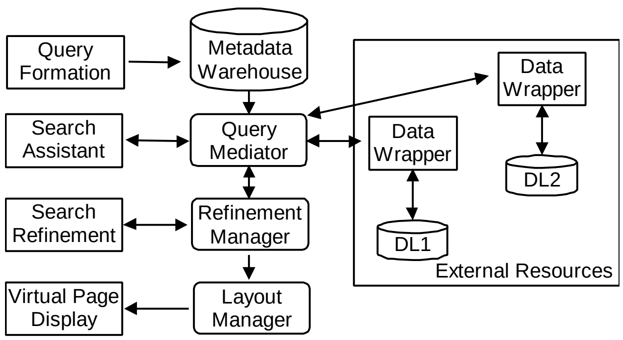

Figure 1: The Delaunay^{MM} Architecture

Users begin by submitting keywords for their query. They may also select a layout style for result display, information repositories for querying, and the type of media objects to return (image, text, audio, or video) at this point (see Figure 2). These parameters are recorded and used by the Refinement and Layout Managers to support query refinement and provide virtual document layout, respectively.

Before the query is submitted to the individual information repositories, the keywords are submitted to a resident resource, the Metadata Warehouse. This data warehouse (Widom, 1995) contains the keyword structure, a tree-like structure depicting the relationship between keywords, of all available repositories. Querying the Metadata Warehouse prior to sending the query to the different resources allows users to refine their query based on the general content of the different repositories.

After selecting the query's context, the user's keywords are sent using data wrappers to the various information repositories, via the Query Mediator. Results are returned to the Query Mediator and then passed to the Refinement Manager. Using the Refinement Manager, users may refine their queries by selecting a group of representative objects (see Figure 3). The resulting objects are passed to the Layout Manager, which creates the pages of the virtual documents.

Delaunay^{MM} and Customizable Layout

Using the optional page format specification, users are able to customize the layout of their query results. These layouts are predefined by a layout designer. The designer, drawing upon his/her knowledge of human-computer interaction and layout principles, creates and saves layout templates for a specific user group. These templates are created using the visual specification tools available through Delaunay^{MM} (Cruz & Lucas, 1997) (Cruz & Lucas, 1999). When query results are returned, these templates are populated with the results. Notably, these templates are static. Thus, the final layout is not dependent upon the actual query results; rather it depends solely on which template is selected.

In the current version of Delaunay^{MM} we have experimented with three model templates: Student/Researcher, Curator, and "Compare two at a time". Each template is designed to support a particular task or user group. The Student/Researcher template is designed to provide complete sets of information. For example, images and their supporting text appear together in this template, as shown in Figure 4. The Curator template shows an object with its associated metadata. For example, an image would be displayed with file location, size and type, as well as other information about the image. Finally, the "Compare two at a time" template presents the query results in two traversable windows side-by-side. The only purpose of this format is to support comparison on query results, as shown in Figure 5.

We refer to the instantiation of a template as the virtual page. The layout of each page is the same, because the templates are static. Each object placeholder is given its dimensions at layout specification time. Due to this static layout type, the selection of the template is of the utmost importance. Assuming that hundreds or thousands of these templates may be defined, selection is not trivial. Selection must take into account the attributes of the user's selected template group (e.g., Student/Researcher may be associated with a group of one-hundred templates) and the objects returned from the query (e.g., if only images are returned, the template should not reserve space for text). As at present there are only a few templates available, selection is not an issue. Our ideas for the non-trivial situation presented above are discussed later.

Here, predefined templates represent the base cases and the target uses the object types returned from the query. Thus, the target template must match the specific attributes of the query results. This process ensures that an appropriate template is selected based on the objects returned from the query.

Our virtual page metaphor also supports the preservation of context while traversing query results, unlike most traditional query engines. By *context* we mean the ability to gather and link together only the multimedia objects (e.g., text, images) that are retrieved as a result of the users' queries. Optionally, users may leave the Delaunay^{MM} interface to get from a virtual page to the original web pages containing their query results: the "Go to Web Page" button opens a new browser to support this option. Ultimately this option does not alter our ability to preserve context, just provides a link to the more traditional hyperlink traversal. We assert that the preservation of context in navigation offered through Delaunay^{MM} enhances a user's ability to understand his/her results.

An example of traditional hyperlink navigation is depicted in Figure 6. Here, users must jump between the list of results and the site containing the content. Each site has its own context, navigation style, and design through users must maneuver. As users are moving between the result list and the various sites, the initial "context" may be easily lost. Users may proceed to continue hyperlink browsing or to search or query to gather another list of "related" results.

In Delaunay^{MM}, we have closed the gap between result list and content, by providing everything in one interface. The customizability of the interface also supports user- or task-specific information gathering. Figure 7 shows the navigational model provided in Delaunay^{MM}. As previously stated, users may optionally leave Delaunay^{MM}'s site to explore the original web sites. This option affords users with the flexibility to find out more information (Cruz & James, 1999a).

Usability Studies

In the first prototype of the Delaunay^{MM} system, end users created the virtual documents by assembling the templates from scratch (Cruz & Lucas, 1997) (Cruz & Lucas, 1999). The first prototype was implemented for the Perseus Digital Library (Crane, 1996). While the digital librarians were enthusiastic about this facility, the end users (students from the Classics Department) had difficulty in defining the templates. However, once the templates were defined for them, they appreciated the new capabilities provided by Delaunay^{MM}, e.g., of displaying more than one object at a time for comparison. Previously, in order to achieve this capability, the end user would have had to open two browser windows side by side and by careful navigation of the web site retrieve two objects for comparison. Therefore, iterating through two sets of related objects comparing two at a time used to be too laborious a task to be performed in practice.

Having this preliminary user feedback, the new interface to Delaunay^{MM} (as presented in this paper), was designed and implemented. After the first extensive user study, improvements were introduced to the interface and second usability studies performed. In the next subsections, we summarize the results of the two studies.

First Usability Study In December 1998, 34 students participated in the first usability study of the new Delaunay^{MM} interface and architecture. This study focused on two issues: the intuitiveness of the new interface, and defining the flaws in the approach and interface. We created a digital library, Digital Dinosaurs, for use with this survey to provide a more engaging subject that would interest users of all types. The overall results of the first study supported the approach and interface, noting minor changes that would significantly increase overall usability.

For this study, users were asked to use Delaunay^{MM} to submit two queries to various digital resources, and to create their own virtual page format using the drawing tools. Each query used different Delaunay^{MM} attributes to tailor the interface to the specific task. Once finished, users were asked to complete a 40-question survey (Shneiderman, 1998) in which they either affirmed or negated statements about each aspect of the interface. Two users were monitored to confirm that the study's usage scenarios and survey questions were fair. All other users performed the study over the course of one week for 30 to 45 minutes at their leisure.

The overall results of the study were very positive, but the user population was biased thus diminishing its capacity to prove general usability. With 59% of respondents being expert web searchers having high confidence in their searching capabilities, and all others being average, this study did not have a representative user sampling (Georgia Tech Research Corporation, 1998). All respondents noted that they are typically successful in finding information on the Internet, with most (88%) finding the necessary information within submitting a few searches. This bias was most heavily noted in response to the statement: "I would prefer this interface to existing search interfaces."

Although there was a close split between positive and negative responses (36% and 40%, respectively) many of the users commented that the response time was too slow. This slow response time is a result of the need to transfer multimedia elements, as opposed to text-based HTML, over the Internet, and is an area for further technical exploration. Ultimately, for expert web searchers response time is a key usability factor, thus the interface was not regarded as highly by these users (Georgia Tech Research Corporation, 1998).

Many users were pleased to see a new approach to information display and retrieval, but others found the uniqueness of the approach unsettling. One user stated, "My general idea on the whole thing is that people aren't used to it. Everybody expects to see textual links and descriptions when you give them pictures, audio/video, that scares them." This concern may be attributed to the high number of expert users who are successful with existing interfaces and reluctant to change. This concern also highlights the uniqueness in this approach, showing that many users have not been previously exposed to this type of query interface.

Second Usability Study In April of 1999 a second usability study was conducted. The goals of this study were to reaffirm the findings of the first study, assess the benefit of the subsequent changes made to the interface, and determine universal usability. Unfortunately, this study could not achieve all of these goals because of the small number of respondents. With only thirteen participants, accessing overall usability and improvements, reaching a convincing conclusion is difficult.

For this study, participants were presented with two use case scenarios to use with Delaunay^{MM}. In each scenario participants were asked to submit keywords for searching, as well as optional fields to specify search criteria. After completing the scenarios, participants were asked to complete a 25-question survey. The anticipated time for the study was 15-20 minutes. No participants were monitored during this study, mass emails were sent out to faculty, staff, and students to solicit participants. This method allowed us to gain a more diverse user population but it is limiting in that the only participants are those that desire to participate. The responses from the survey were received anonymously.

Of the thirteen participants, there were nine males and three females. The age distribution was more diverse than the first study, with seven participants between 18-24, three between 25-38, two between 39-50 and one over fifty. This distribution approximates more closely the average Internet population. Most of the participants, 70%, use the computer at home and work, and 100% use the Internet daily. Fifty-eight percent also search on the Internet daily, with 34% searching a few times a week and eight percent less than once a week. One hundred percent of the participants are successful in finding information on the web within a few searches. When participants were asked to rate their Internet searching skills seven designated themselves as expert and six as average searchers.

With respect to the interface, the overall results were promising but not exceptional. Fifty-five percent agreed that they would prefer Delaunay^{MM} to a traditional search interface, assuming that the response time was the same. This result shows a 19% increase over the previous results, but is still 11% shy of a positive result.[1]

Participants who had completed in the first study were also asked if the new interface was an improvement over the previous one. Two participants replied in the affirmative to this statement. Because the responses were anonymous we have no way of confirming that these two were previous participants, but we trust that the participants were honest in all their responses.

Many of the comments from this study revolved around Delaunay^{MM}'s approach and not the interface attributes itself. Because the respondents in this case were not members of a human-computer interaction course, as in the first study, they probably did not articulate their observations in these terms. Although the survey was somewhat limited in the number of participants, the results reinforce those of the first study and affirm the approach in general.

[1] A positive result is achieved when at least 66% of respondents reply in the anticipated manner.

Enhancing Layouts in Delaunay^{MM}

Templates can be instantiated with retrieved objects to generate actual displays (i.e., "virtual pages"). A template can be viewed as an abstraction of every page that it can generate. The templates supplied with the current system were configured by considering standard users and tasks, and their information display needs (e.g., compare two at a time).

In this section we present ideas for the enhancement of the layout that can be incorporated at two levels. One level has to do with extending the current capabilities of the system to automatically generate (or help in generating) layouts suitable to the user and to the task. At this level, information contained in the metadata (Arms, Blanchi & Overly, 1997) (Daniel, Lagoze & Payette, 1998) (Weinstein, 1998) could be used, e.g., to determine appropriate objects to be compared. At another level, new kinds of functionality can be added, for example to take the templates and generalize across them to produce families of templates with shared characteristics. Such more general templates might be used to describe classes of queries and displays.

In our current system the elements of the templates are fixed and are in fixed relationships. However, variation in the data to be displayed should provoke variations in the display. The idea of "dynamically resizing" the templates using a constraint solver, so as to preserve the size or proportions of the multimedia object, is a natural extension (Cruz & Lambe, 1998) (Graf, 1996).

If an existing template does not provide the layout of objects that best matches the user, task, the objects retrieved or the relationships to be preserved among those objects, then a new template needs to be generated.

The least constricting approach is to dynamically configure the template from scratch from appropriate template elements. Configuration is an active research area in Artificial Intelligence (Faltings et al., 1999). There's also a history of research on the knowledge-based generation of displays (see references in Balazs et al. (1997)).

Probably the most common approach to configuration is Dynamic Constraint Satisfaction (Mittal & Falkenhainer, 1999) — also more recently called Conditional Constraint Satisfaction. For Delaunay^{MM} templates, constraints arise naturally, controlling sizes and proximity for example. In addition, as new elements are included in the configuration, additional constraints may be needed.

Normal Constraint Satisfaction Problems (CSPs) have all the variables and constraints known, and fixed, at the start. Dynamic CSPs (DCSPs) relax that requirement, allowing constraints and new variables to be introduced.

For example, given a particular kind of retrieved multimedia object, an area to display a picture might be required. But if the retrieved pictures have associated captions, then another element, a text area, will be required in the configuration, as well as a new constraint that requires the caption be close to the picture.

Other possible use for template configuration is to build them from task-oriented template pieces in response to the user's task(s). For example, to compare images of two coins found at two different but related archaeological sites, it would be convenient to have two picture areas side-by-side. If the comparison is between a single coin and a set, we would want one picture area to be able to scroll through the set. So the task of comparing two items leads to a different display than the task of finding whether a single item is in a set. Task-dependent display configuration is described by Balazs et al. (1997), but their approach did not use DCSPs.

As the system is extended to include more templates, perhaps including templates at different levels of abstraction and for different tasks, the selection of a template to use will become more difficult. If the set of templates is fairly heterogeneous then the system help by using index-based retrieval techniques from Case Based Reasoning (Kolodner, 1993). For more diverse templates with rich structure and different types of template elements, a more powerful mechanism would be needed.

Use of the Structure-Mapping Engine (SME) (Falkenhainer, Forbus & Gentner, 1986) (Falkenhainer, Forbus & Gentner, 1989) might be appropriate. Characteristics of the task, user, objects and object relationships would be given as the "problem", and, after some pre-filtering of stored templates based on a limited number of primary features, SME could be used to provide an indication of the matches, and their quality, between the problem and the candidate templates.

However, as the SME tends to consider relationships between things as more indicative of a match than the things themselves, this would require explicit information about relationships to be stored with the templates. This could perhaps be the constraints that they satisfy, or the relationships that they can convey/preserve between retrieved objects (such as the order of dates when text samples were written).

Future Work

Clearly, many of the above enhancements constitute "Future work". However, there are several research ideas that may deserve our attention. First, we would like to develop a scheme for evaluating the resulting layouts, which should preserve the intentions of the web site developer and actually enhance understanding.

Second, as in other constraint and design problems, there could be several ways to meet the specifications, or none. In the presence of multiple solutions, priorities among the layout criteria should be established. In the absence of a solution, the priorities can be used to selectively relax the constraints until a solution is reached.

References

W. Arms, C. Blanchi, E. Overly. An Architecture for Information in Digital Libraries. D-Lib Magazine, 1997.

M. E. Balazs, D. C. Brown, P. Bastien & C. E. Wills. Graphical Presentation of Designs: A Knowledge Intensive Design Approach. Knowledge Intensive CAD, Volume II, (Eds) M. Mantyla, S. Finger & T. Tomiyama, Chapman & Hall, 1997, pp. 173-188.

G. R. Crane, ed. Building a Digital Library: the Perseus Project as a Case Study in the Humanities. First ACM Int. Conf. on Digital Libraries, 3-10, 1996.

I. F. Cruz and K. M. James. Preserving Contextual Navigation in Hypermedia Querying. [Online] Available http://www.cs.wpi.edu/~beez/, 1999a.

I. F. Cruz, and K. M. James. A User-Centered Interface for Querying Distributed Multimedia Databases. ACM SIGMOD 99, 1999b.

I. F. Cruz and D. I. Lambe, Implementation of an Efficient Constraint Solver for the Layout of Graphs in Delaunay. Sixth Symp. on Graph Drawing Springer-Verlag, 1998.

I. F. Cruz and W. T. Lucas. A Visual Approach to Multimedia Querying and Presentation. Fifth ACM Int. Multimedia Conf., 109-120, 1997.

I. F. Cruz and W. T. Lucas. A Customizable Layout-Drive Approach to Querying Digital Libraries. SPIE Multimedia Computing and Networking, 122 - 134, 1999.

R. Daniel, C. Lagoze, S. Payette. A Metadata Architecture for Digital Libraries. IEEE Int. Forum on Advances in Digital Libraries, 276-288, 1998.

B. Falkenhainer, K. D. Forbus & D. Gentner. The structure-mapping engine. Proceedings of the American Association for Artificial Intelligence, 1986, pp. 272-277.

B. Falkenhainer, K. D. Forbus & D. Gentner. The Structure-Mapping Engine: algorithm and examples. Artificial Intelligence, 1989, Vol. 41, No. 1, pp. 1-63.

B. Faltings, E. C. Freuder, G. Friedrich & A. Felfernig (Eds.). Proc. AAAI Workshop on Configuration, 1999.

Georgia Tech Research Corporation. GVU's 9th WWW User Survey. http://www.cc.gatech.edu/gvu/user_surveys/survey-1998-04/, 1998.

W. H. Graf. Towards a Reference Model for Intelligent Multimedia Layout. ECAI 96 Workshop Towards a Standard Reference Model for Intelligent Multimedia Systems, 1996.

J. Kolodner, Case-Based Reasoning. Morgan Kaufmann, 1993.

S. Mittal & B. Falkenhainer, Dynamic Constraint Satisfaction Problems. Proc. 8th Nat. Conf. on AI, AAAI-90, 1990, pp. 25-32.

B. Shneiderman. Designing the User Interface: Strategies for Effective Human-Computer Interaction. Addison Wesley Longman, Inc., 1998.

P. Weinstein. Ontology-Based Metadata: Transforming the MARC Legacy. Proceedings of the Third ACM Conf. on Digital libraries, 254-263, 1998.

J. Widom. Research Problems in Data Warehousing. In Proceedings of 4th Conf. on Information and Knowledge Management, 25-30, 1995.

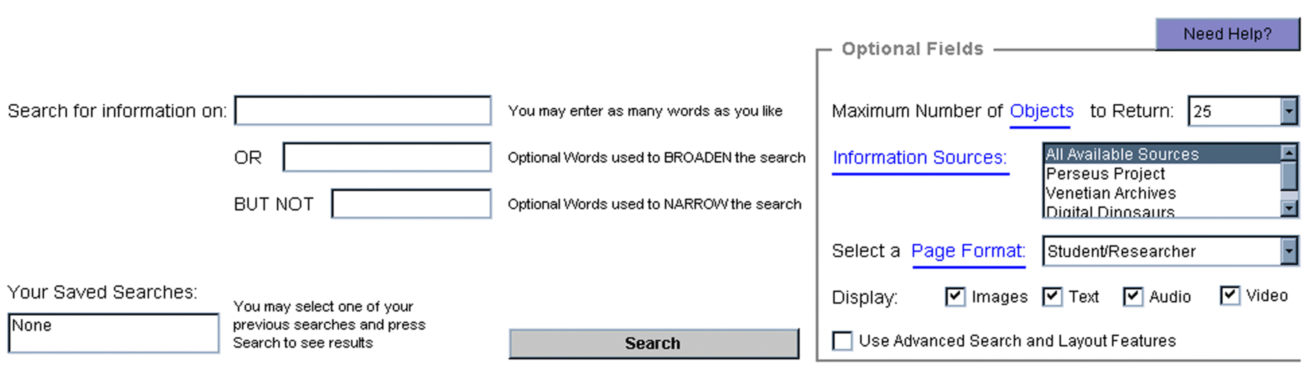

Figure 2: Delaunay^{MM} query specification

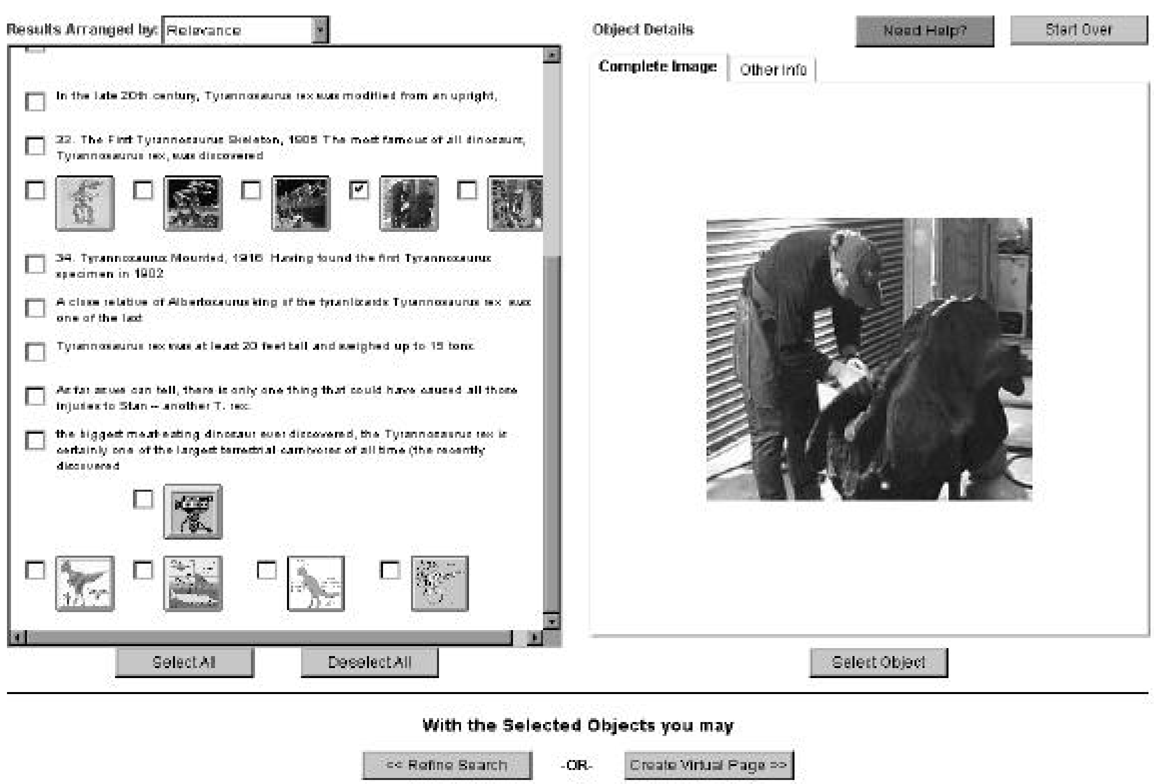

Figure 3: Post-query refinement (Refinement Assistant)

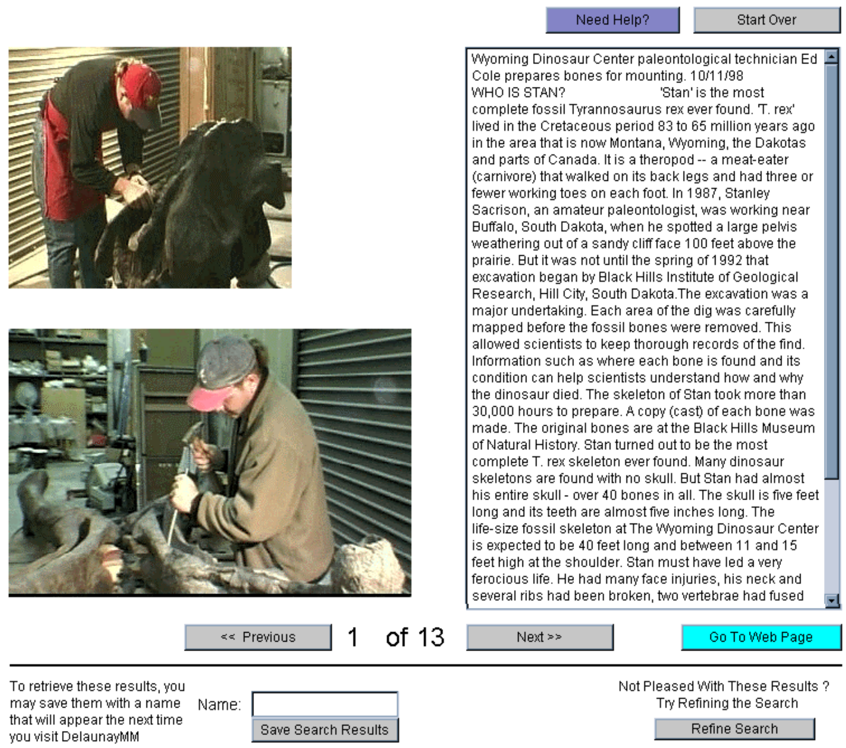

Figure 4: The Student/Researcher virtual page

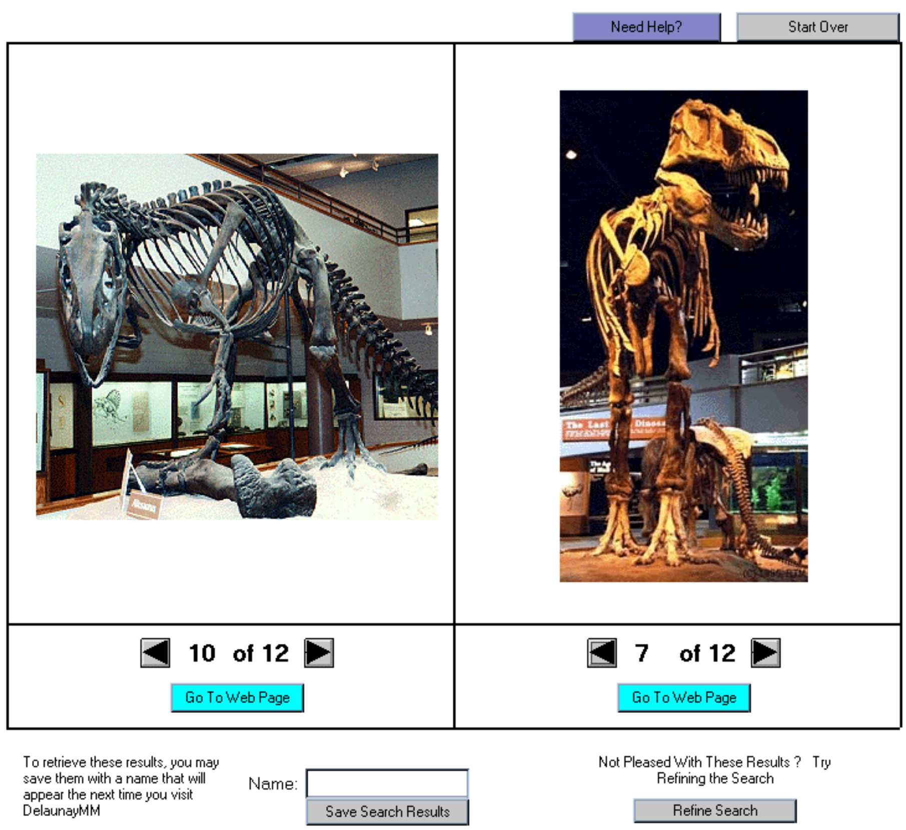

Figure 5: The "Compare Two at a Time" virtual page

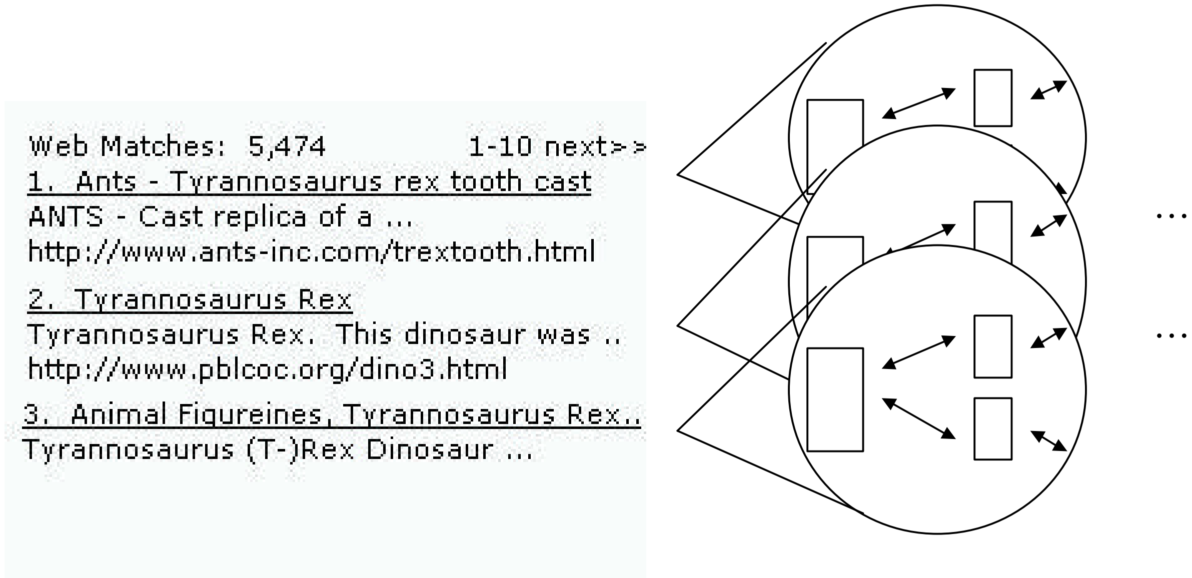

Figure 6: Traditional hypertext search results

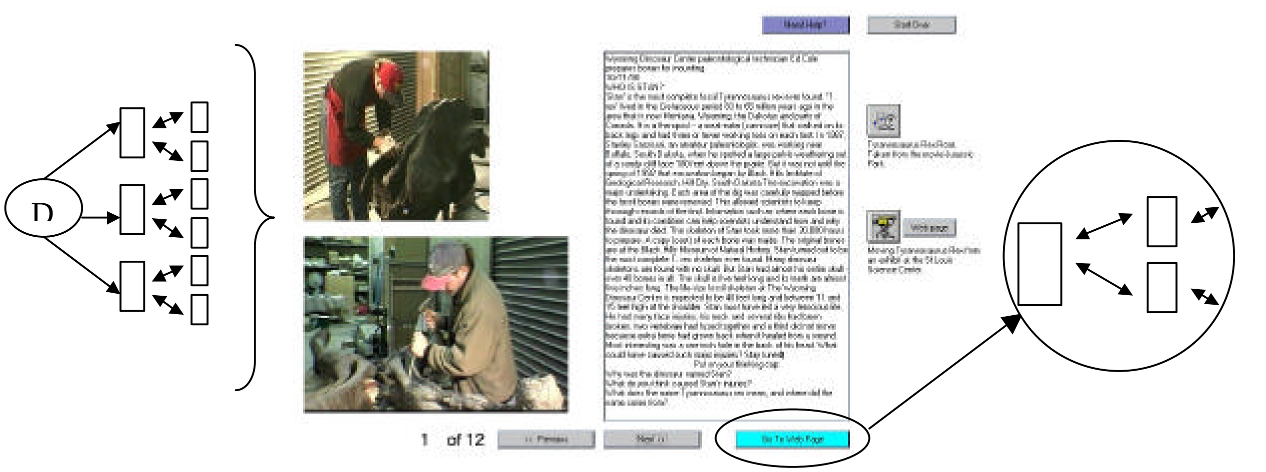

Figure 7: Delaunay^{MM} navigational model (Student/Researcher Layout)