ID2 – A Scalable and Flexible Mixed-Media Information Visualization System for Public Learning Exhibits

Eric A. Wernert, Jagannathan Lakshmipathy, Michael J. Boyles
Advanced Visualization Laboratory
Indiana University, Bloomington, Indiana, USA
ewernert@indiana.edu, jlakshmi@iupui.edu, mjboyles@iupui.edu

Katy Börner
School of Library and Information Science
Indiana University, Bloomington, Indiana, USA
katy@indiana.edu

Abstract. This paper describes an innovative, mixed-media information visualization display method and system architecture for public learning exhibits. The system is scalable in the number and types of displays and flexible in its applicability to a range of data domains and interaction models. This work extends and improves the original *illuminated diagram* paradigm which merges static, ultra-resolution, hardcopy prints with dynamic, lower-resolution, softcopy graphics. In this paradigm, projected graphics are used to highlight relationships among the detailed information in the hardcopy prints. Multiple coordinated views enable the comparison of data between reference systems, and interactive queries employ brushing and linking methods to synchronize highlighted information between views. This system is applicable to any information visualization display where multiple, static reference maps can be used in unison to improve the exploration and understanding of very large data sets.

Introduction

Many of today's most important scientific, social, business, and policy decisions rely on the effective analysis of massive amounts of data. This task is made challenging not only by the size of the data sets, but also by the fact that the data is aggregated from many different sources, each benefiting from different representations and reference systems. In addition, lay people are being exposed to more advanced information displays and frequently want or need to deal with more complex data. Technology-enabled public information and learning exhibits, such as those found in modern museums, libraries, and science centers, can play an important role in helping to educate the broader public on the benefits and methodologies of information visualization techniques.

A variety of methods from the fields of information visualization, human perception studies, and human computer interaction can help address the challenges of the data deluge, including: ultra-high resolution display methods, multiple coordinated views with brushing and linking interfaces, focus plus context techniques, intuitive and efficient interaction methods, and the design of perceptually effective information representations. An innovative paradigm that draws from these techniques was developed by Paley and is dubbed the "illuminated diagram" (Paley, 2002). The illuminated diagram (ID) technique merges static, ultra-high resolution, hardcopy information with dynamic, lower-resolution, softcopy graphics. The interactively-driven softcopy graphics are projected from digital projectors onto detailed, hardcopy prints and serve as a "smart spotlight" to highlight subsets or illustrate relationships among the detailed information. The ID method was extended to a multiple view configuration in (Boyack et al. 2007).

In this paper, we present a major re-interpretation, improvement, and generalization of this multiple-view ID system. Our goal was to develop an architecture that provides flexibility, scalability, and easy deployment for public exhibits or dedicated learning stations. The resulting system is called the ID2 for "improved and distributed illuminated diagram". The novel contributions of the ID2 project are: (1) a new technique for illuminating printed diagrams using large, flat panel displays and translucent overlays that enable improved brightness, resolution, and alignment, while simplifying setup and maintenance, (2) a distributed software architecture that can scale to an arbitrary number of systems, allowing flexibility in the number of displays and overall resolution while keeping compute and graphics hardware requirements relatively modest, and (3) a distributed interaction model that enables local or remote collaboration by multiple users over the same data while also permitting the introduction of novel interface devices and interaction methods.

Currently, this system is being used in conjunction with several deployments of an international traveling exhibit on the art and science of mapping science. However, this technique and software framework are broadly applicable to any data-intensive information visualization application where static frames of reference (printed base maps) exist and multiple coordinated views are required. We detail the motivating science mapping application in the next section and illustrate the generality of the ID2 system in a later section by describing several new examples that are under development.

Figure 1. Left: the original science mapping ID using art prints and front projection. Center: the new ID2 implementation using translucent overlays and rear illumination from LCD screens. Right: the touch screen interface where viewers may touch regions in the map of science, the map of the world, or select scientist or meta-topic buttons.

Motivating Application

The primary driving application for the ID2 architecture is an interactive information display that was designed as a component of a traveling exhibit on mapping science (Hardy et al. 2009, Boyack et al. 2007). This display pairs a "topic map" of science with a geographic map of science, each based on the analysis of approximately 800,000 scientific articles from a scholarly publication database. Articles in this database are linked to each other based on reference citations. To create the topic map, the articles were clustered into scientific paradigms or fields based on the highest levels of co-citation. These resulting scientific fields or paradigms (776 in total) were arranged using a force-directed layout algorithm based on the frequency of cross-citations between papers in respective fields. For the geographic map, articles were assigned to over 12,000 different locations based on the country and postal codes of the primary author.

The printed topic map contains a node for each scientific paradigm and text labels for the most highly populated nodes. Clustered around each paradigm are commonly occurring keywords selected from papers within the domain. The printed geographic map contains standard geographic and political boundaries as well as nodes for all cities represented in the data set and labels for the most frequently referenced cities. For both maps, the aggregate number of articles in each node is reflected by relative node size. This data analysis and map generation work was done by another team and is described in (Paley 2008).

The original 800,000 data records were aggregated to keep the query times low and the interaction speeds high. In this aggregated data set, each scientific paradigm links to a set of geographic locations. Conversely, each geographic location links to a set of scientific paradigms. Additional, pre-defined queries based on well-known scientists or timely scientific paradigms were extracted from the original data and stored in separate tables. In total, four distinct types of data exploration are supported to allow for multi-faceted exploration of the complex relationships contained within the data. (1) For any given geographic location or range of locations (selected by dragging a region), the corresponding nodes in the map of science are illuminated, illustrating the types of scientific research conducted there. (2) For any selected scientific paradigm, the corresponding nodes on the world map are illuminated, illustrating institutions were a lead author has published on that topic. (3) For several well-known, "meta" paradigms such as nanotechnology and sustainability (paradigms that span many disciplines on the topic map), users may select a button to highlight all of the contributing paradigms and geographic locations. (4) For several well-known scientists, such as Albert Einstein or Francis Crick, users may select a button to highlight a series of article sets (and hence their geographic and topic locations): those authored directly by that scientist (first generation articles), those articles that reference the first generation articles (second generation articles), and those that reference second generation articles. In this way, users can explore the scientific and global impact of a given scientist's intellectual contributions.

While very successful, the initial implementation of this system had some major support challenges, and the hardware and software design was not extensible, flexible, or sustainable. Therefore, we set about a full redesign of the underlying hardware and software infrastructure to address the shortcomings and to make a system that was more widely applicable.

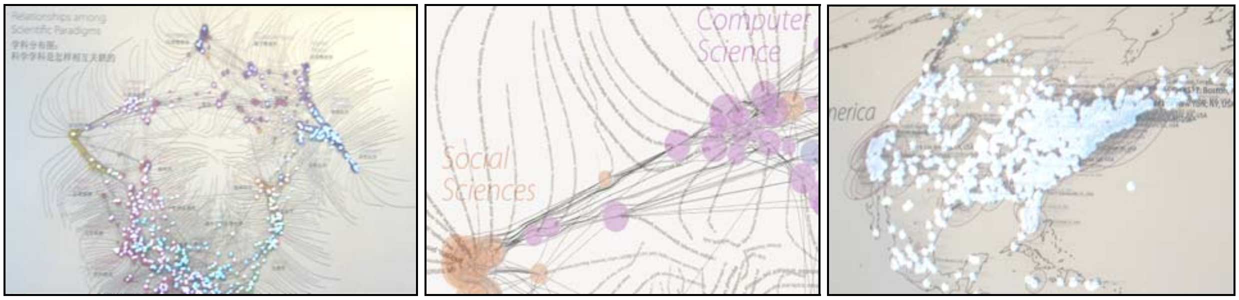

Figure 2. Images from the ID2 for science mapping. Left: photograph of illuminated science topic map. Center: detail of science topic map. Right: detail of illuminated geographic map. (Topic map by W. Bradford Paley; geographic map by John Burgoon.)

Related Work

Computer displays have increased significantly in size and resolution in the past several years. Despite these improvements, large-format, hardcopy prints still have considerable advantages over large-format, softcopy displays in terms of resolution, size, and cost. The illuminated diagram is a mixed-media method that seeks to combine the benefits of these two complementary technologies.

The illuminated diagram technique embodies several proven techniques from the field of information visualization; chief among these are the techniques of *focus plus context* and *brushing and linking*. The premises of *focus plus context* techniques are that the viewer needs or can benefit from both overview (context) and detail (focus) information simultaneously, and that both types of information can be effectively combined in a single display (Card et al. 1999). *Brushing and linking* techniques combine different visualization methods to overcome the shortcomings of any individual technique. Interactive changes made in one visualization window are automatically reflected in the other visualization views, providing more information than either of the component visualizations could independently (Keim 2002).

Other techniques that combine large format, ultra-high resolution printed information with dynamic, digital displays are beginning to emerge. Notable among these are the efforts of the GIGAprints group that utilize electronic pens as the basis of the interaction model (Yeh et al. 2006).

New Technical Approaches

Display Innovations

The original ID concept utilizes large-format, art-quality prints mounted on form core panels that are hung on a gallery wall. Digital projectors mounted on stands in the gallery space in front of the print are used project interactive graphics to highlight and complement the printed information. The major shortcomings of this arrangement include the challenge of aligning the projected information with the printed map (requiring precise positioning of the projector and advanced features such as keystone correction) and the problem of viewers occluding the projection when they approach the printed diagram to study the details.

Our improved display method utilizes a rear-projection configuration of printed translucent overlays on top of LCD panels[1]. This not only solves the problem of projection occlusion, but provides greater addressable resolution for the interactive illuminating graphics.[2] In addition, with proper stands, LCD screens permit for arbitrary positioning of the displays in the gallery space (e.g., away from walls) along with the ability to provide more visible and comfortable viewing by angling the displays in the vertical and horizontal axis.

For the printed overlays we use a translucent, self-adhesive material designed for use in retail environments such as large signs or graphics on store windows. This material provides precise adhesion that releases cleanly when the overlays need to be replaced. It is also designed for rear illumination, allowing for deep saturation of ink colors into the print surface to produce vibrant results. This technique can also be applied to laptop and standard desktop LCD screens for testing and development purposes or deployment with smaller, less detailed prints. The effective resolving power of the viewer's eyes is the only practical lower limit on the size of the prints.

The physical application of the prints directly onto the screens alleviates the major alignment problems (position, focus, scaling, keystone correction) introduced by the use of projectors and eliminates the need for frequent adjustments caused by drift or viewers bumping into projector stands. The maps are scaled to the size and aspect ratio of the screen. Once affixed, setup features in the software facilitate the fine-tuning of horizontal and vertical offset and scaling factors to insure precise alignment between the softcopy and hardcopy. This data is then saved in a configuration file that is accessed whenever the software starts in normal display mode.

Scalable Software Architecture

The original illuminated diagram used a single computer with multiple graphics cards to drive three separate displays. This represented the single biggest limitation of the original system since all of the data management, querying, graphical rendering, and user interaction monitoring took place on the same system. Even with the prevalence of workstations with multiple, multi-core CPUs and multiple graphics cards, the computational power, rendering capabilities, number of displays, and aggregate resolution of any single system is still finite. We designed the ID2 system around a new, fundamentally distributed architecture. The components and interaction of this architecture are described below and are illustrated in Figure 3 as they are configured in the science mapping application. The system was developed for Microsoft Windows XP or Vista using the .NET framework and OpenGL for the graphical rendering.

The *data server object* provides an encapsulation of the data file readers and the overall system data structure. On system startup, it parses system configuration files and loads the data files into an internal data structure. It handles startup data requests from each frame data object, passing each object the information needed to display its representation of the data. Currently, system data files are stored in XML formats. Future iterations will utilize a relational database to enable more general, viewer-specified queries. For every installation, there is a single data server object.

Each unique base map in the system represents a different reference system or a distinct frame of reference into a shared set of data. These reference frames are referred to as *views* in our architecture. The *data view object* encapsulates the portion of the overall data structure that is needed to illuminate data within its frame of reference. As such, it contains code for rendering its data into an appropriate visual representation. The data view object also stores references to linked data points in other views within the system so that it can respond to user selections within its interface by passing messages to its peer views. In the science mapping example there are two view data objects: one for the science topic map and one for the geographic map.

A convenient feature of the .NET framework is the capability to create shared network objects. If a view data object is required to appear on multiple displays (as it is in the science mapping example where the control screen contains a small, interactive instance of each of the larger, full-screen passive instances, see Figure 1), it only needs to be instantiated by one process. When the object needs to be referenced in another process running on the same system or on a remote system, .NET provides for a proxy object that transparently handles the networking

---

[1] Initial investigations looked at other flat-screen technologies, including plasma and DLP rear-projection displays. Plasma screens were rejected because the thick front glass causes the softcopy content to appear out of focus. Rear-projected DLP monitors were rejected because they lack the wide viewing angle of LCD displays.

[2] At the time of our initial investigation, LCD screens were the only available technology offering true 1080p resolution. At the time of this writing, projectors with 1080p resolution are becoming more widely available, but they are still many times the price of an LCD monitor with equivalent resolution.

details. Updates made to the shared network object directly or through its proxy object are automatically propagated to all other systems sharing the object through a TCP/IP protocol.

Each display in the system requires an *interface object*, whether it is passive (output only) or active (capable of capturing user input in addition to displaying output.) Each interface object contains an instance (singular or shared) of one or more data view objects. The interface object uses configuration data to enable it to map the visual representations generated by its data view objects onto its display. This mapping also allows the interface object to pass user selection events to the data view objects in order to perform brushing (highlighting) and linking (message passing) operations. In the science mapping example there are three distinct interfaces: one for the large topic map that creates the instance of the topic data view object; one for the large geographic map that creates the instance of the geographic data view object; and one smaller, touch-screen interface that shares references to the existing geographic and topic data views, and that also creates two additional data views for the pre-defined meta-topic and scientist selections. (See center node in Figure 3 and interface photograph in Figure 1, right.)

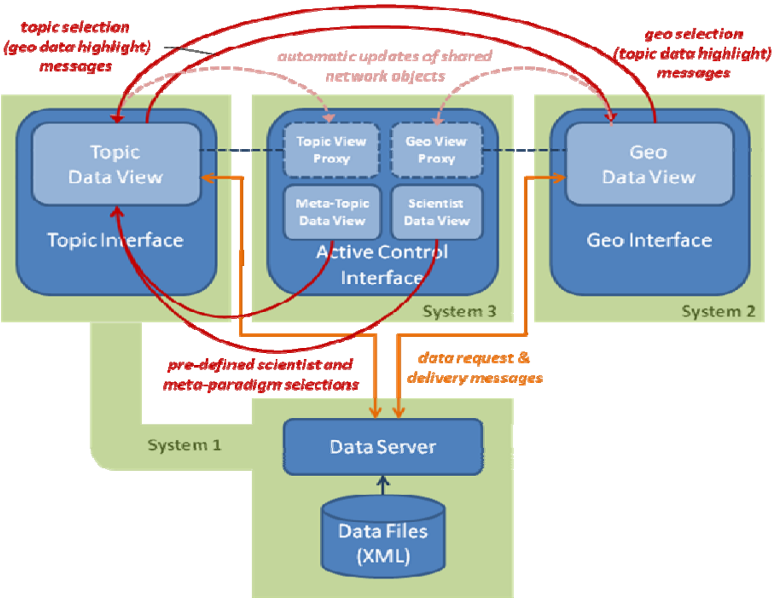

Figure 3. Software components of the ID2 architecture as instantiated for the science mapping project. Components reside on three different networked systems. Arcs indicate messages sent as a result of user selections and interactions.

There are two general types of communications between components of the system; all are accomplished through standard TCP/IP sockets. (A third type, the shared network object update, is implicitly handled through the .NET framework and keeps all shared data views synchronized.) Each component accesses a configuration file which provides the network addresses and port numbers where it should send and/or receive messages. *Data requests* occur at startup time as each data view requests the portion of the overall data structure that it needs to generate its view. The data server object receives and fulfills those requests. The system relies on an explicit startup order of: data server, then interfaces instantiating data objects, followed by interfaces linking to shared data objects. *Data selections* occur any time a user touches a region of one of the active maps or chooses a predefined query button. In the science mapping example, selection of a set of topics in the topic map will highlight those topics and trigger a message to the geographic map to highlight its corresponding nodes. The reverse is also true. Because of the nature of the pre-defined meta-topics and scientist selections (recorded as sets of topic nodes), user selection of those buttons routes a message to the topic data view which propagates corresponding changes to the geographic data view.

While the architecture provides the flexibility for all the software components to be run on a single system or each to be run on a separate system, some practical considerations influence the actual hardware configuration and distribution of components. To simplify the hardware requirements and setup time, and to minimize cabling in exhibit spaces, we tend to use one computer for each display system. We also examine the memory, computing, graphics processing, and network traffic requirements of each system to ensure an equitable load balance. In the science mapping setup, we found that the topic map had a much smaller data management and rendering load (776 nodes) when compared to the geographic map (over 12,000 nodes). Even though the topic map had a noticeably higher communication requirement (since most selections are queried through its data structure), there was still sufficient resources remaining on that system to also run the data server there. The green boxes in Figure 3 show the distribution of processes across the three systems that comprise the installation. The science mapping exhibit can be run adequately from three ultra-compact computers (Apple Mac Minis running Windows XP) which allows the computers to be attached to the back of the displays or otherwise be hidden from view.

Distributed Interaction

Another useful and powerful resultant of this distributed architecture is the ability to distribute the interaction model across multiple computer systems, and hence, among multiple simultaneous viewers. This allows for the creation of a variety of local and remote collaboration scenarios, and also opens up possibilities for deploying novel interface hardware and techniques.

The original ID implementation was based on a single compute system, and therefore restricted the interaction to a single user interacting with a single screen designated as the point of control. In the ID2 architecture, the common hardware and software structure of the displays and their peer-to-peer relationship allows each display to serve as an interaction point. For example, in the science mapping application, selections on the geographic display are propagated to the science topic map and vice versa. A similar relationship exists between the topic and geographic displays and the consolidated interface display. Rather than using an explicit flow-of-control or control token model as do some distributed methods, we employ a simple FIFO event model and rely on verbal communication and basic social interaction norms to resolve any potential control conflicts. This approach was recognized as being the most effective and natural in much of the early work in collaborative tele-immersion environments (Leigh et al. 1999). As part of our ongoing work, we are investigating an explicit timing model where concurrent selections on different displays would be interpreted by the data server as an ‘AND’ query. For example, concurrent selections on the scientific topic and geographic map could be interpreted as a query for the number or titles of papers that were published on that topic by researchers at that geographic location.

Another benefit of this distributed interaction model is the relative ease with which we can test new input devices and methods. When working with a single compute system, it is often difficult or impossible to incorporate multiple, similar input devices (e.g., multiple pointing devices) or multiples of the same input device (e.g., multiple touch screens.) With the ID2 architecture, each display system can easily pair with its own input device, such as a touch screen, trackball, wireless device, or camera-based tracking system. These technologies are part of our ongoing investigation as we look to make the viewer’s interaction and experience even more intuitive and engaging.

Current Results

The ID2 system has been successfully deployed with the Places & Spaces exhibit to several locations in North America in 2008. In addition, the system was deployed at the National Science Library of the Chinese Academy of Sciences in Beijing, China starting in May 2008 and the “Expedition Zukunft” train (“expedition into the future” exhibit, traveling to 62 cities in 7 months) in Germany in April 2009. (See Figure 4.) When displayed in non-English-speaking countries, the touch-screen interface is translated into the local language and the printed overlays are supplemented with local language subtitles. One of the convenient features of the illuminated diagram concept is that if the interactive illumination graphics are designed without text (using only strokes, filled regions, or images to illuminate text on the print), then language translations can be made with straightforward modifications of the prints and images requiring no changes to the software. We made extensive use of this feature in the preparing the exhibit for China and Germany. These ID2 deployments also benefitted from the fact that, unlike the original ID, no specialized equipment or stands needed to be shipped or sent through international customs. All necessary technology was borrowed or rented on site, and only the printed overlays (which are reprinted to the dimensions of the rented screens) needed to be shipped.

In other local test installations, we have demonstrated some of the innovative capabilities of the ID2 system. We have verified the remote interaction capability by running duplicate sets of displays linked across a wide area network. We have also demonstrated basic wireless interaction at each display by using WiiMotes to emulate mouse interactions.

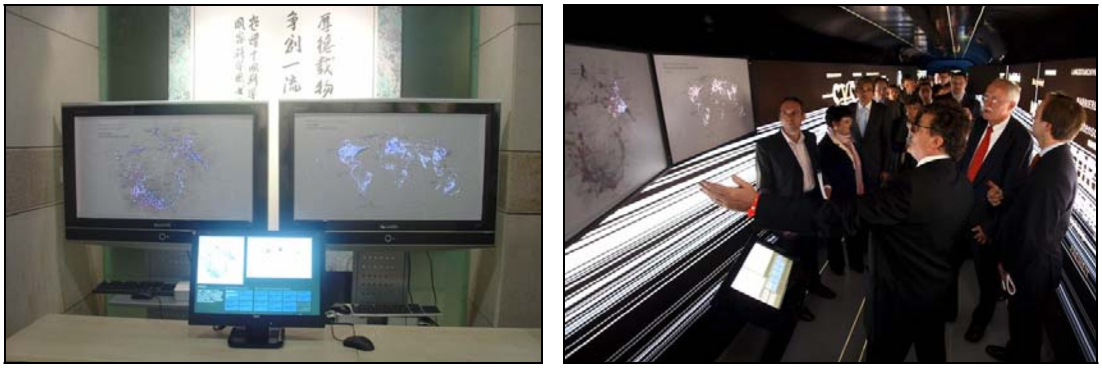

Figure 4. Left: The ID2 deployed as part of the Places & Spaces exhibit at the National Science Library of the Chinese Academy of Sciences in Beijing, China. (Photo courtesy of Weixia (Bonnie) Huang.) Right: The ID2 installation on the "Science Train" in Germany. (Photo © www.archi-me-des.com, Oliver Wia.)

Additional Applications

In addition to the science mapping application, we are actively developing additional uses of this infrastructure. These applications have the fundamental characteristics that make the ID2 system appropriate and effective: they involve large amounts of data that are interesting in aggregate as well as in detail, and this data can be understood best by parallel presentation in complementary reference systems, be they standard reference systems or newly established ones. While not a requirement, it is also the case that these data sets describe networks that also illustrate relationships or similarities of data points *within* each of these reference frames.

The first of these applications visualizes the results of a project in music information retrieval (MIR). MIR is a subfield of information science focused on the automated analysis of audio files for classification, recommender systems, or organization and retrieval from digital libraries. The objective is to rely on feature extraction and similarity clustering based on analysis of the digital audio files themselves, rather than on specific, manually-applied metadata (McCaulay and Sheppard 2008). The initial data for this application consists of a set of over 3,000 recordings from 142 musical artists or groups. Analysis is performed on each pair of songs, resulting in a similarity score for the pair. For each pair of artists, the similarity scores of their respective song pairs are aggregated to compute a similarity measure between the two artists. We anticipate that this project will employ three parallel reference systems (see Figure 5): (1) a force-directed network layout of songs based on their pair-wise similarity scores, (2) a force-directed network layout of artists based on their aggregate similarity measures, and (3) a simple, standard reference system consisting of an alphabetical listing of the artists with alphabetical sub-listing of their songs.

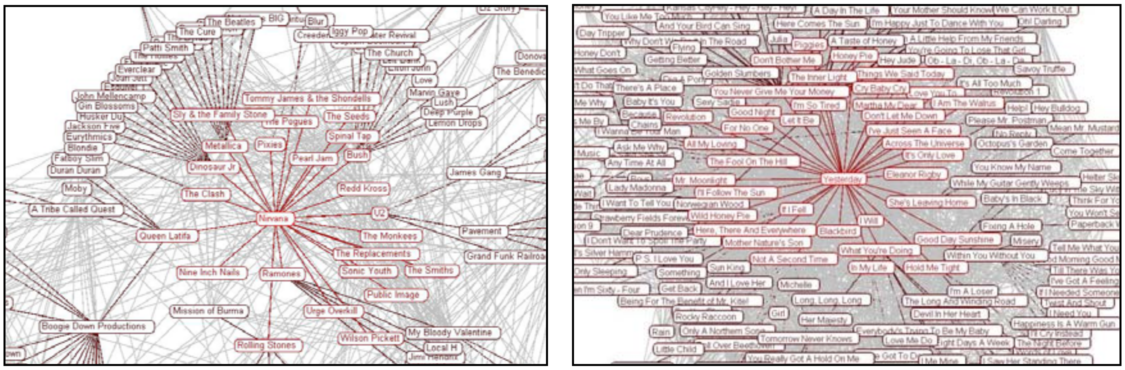

Figure 5. Prototype images for the ID2 exhibit for music similarity. Left: force-directed layout of artists and groups with Nirvana at center. Right: force-directed layout of songs from the Beatles representing similarity to the song "Yesterday". (Images courtesy of D. Scott McCaulay.)

A second planned application involves mapping the allocation and impact of cyberinfrastructure (CI) resources, such as those made available by the NSF and DOE via the TeraGrid and the Open Science Grid. The data we are working to collect and analyze comes from multiple sources and includes the location, types, and utilization statistics for CI resources, as well as the location, scientific domain, resource allocations, scholarly output and grant funding of CI users. This project would utilize two geographic maps – one for the location and details of CI resources, and one for the location of CI users – along with the existing map of science. Related work in mapping distributed CI systems and applications was described in (Allcock et al. 2002).

In addition to these new applications, we are also working on a method to enable supplementary focus maps for the existing science mapping application. This would allow more detailed analysis of geographic locations or areas of science that are likely to be of greater interest to the local viewing audience. Currently, this type of information can be difficult to observe due to the very high density of information on the printed maps.

Future Work

As we have gained experience with the ID2 architecture and display method, we have come to recognize its rich potential and have planned a series of extensions and improvements that will further enhance its usability, effectiveness, and generality.

One task is to fully generalize the current architecture by transitioning all current peer-to-peer communications to route through the central data server in a star topology. While such a topology does create a potential single point of failure or performance bottleneck, it is easier for developers to conceptualize, and it keeps the network management requirements for the client display processes simple. This change will then allow us to conduct tests to better quantify the practical scalability of the number of displays. We anticipate that this result will be highly dependent on network loads and the size of the data results being transferred. We also plan to implement a full relational database underneath the data server abstraction. Since many of the data sets we are working with are currently expressed in some type of linked table form, this would eliminate the data extraction and consolidation steps, and would allow viewers to express arbitrary complex queries instead of being limited to general single-variable queries or predefined complex queries.

While we have successfully demonstrated the capability of the system to work across geographically distributed sites, we would like to more fully study and evaluate remote collaboration scenarios. These scenarios include co-investigation by two or more peer-level users at distributed, concurrent deployments of an exhibit, as well as remotely guided “tours” of the data exhibits by topic experts from different locations. These scenarios would be supplemented with text, audio, and video conferencing capabilities. Likewise, we have also demonstrated the capability of this system to work with multiple copies of non-standard input devices, including integrated touch screens and WiiMotes serving as wireless mouse emulators. However, we would like to more systematically investigate other interface technologies, including: larger format touch screens, multi-touch screens, camera-based touch-free input, and low cost tracking systems.

Finally, we need to develop a better understanding of the possibilities of color and light interaction between the printed overlay, the LCD illumination, and the ambient and directed lighting in the environment. In order to more systematically study the interaction between the print and the rear illumination, we are designing test patterns that will allow comparison of different color combinations as well as the relative advantages of light-on-dark versus dark-on-light prints.

Summary

We have described a scalable and flexible information visualization display method and system architecture that builds upon the existing mixed-media technique of the illuminated diagram. This system has shown itself to be effective and reliable in a number of public information exhibits. We have described new methods for merging hardcopy and softcopy media which offer significant improvements in practicality and effectiveness over the original technique. In addition, we have developed a new, distributed software architecture to ensure scalability in the number of displays, resolution, and data size. This architecture also yields significant benefits for testing advanced interface techniques and collaborative scenarios. We have documented the utility and effectiveness of this system through successful international deployments, and have illustrated the generality of the system by describing new applications that are under development. Finally, we have outlined a set of future investigations which seek to further improve the generality of the system, quantify its scalability, and expand its interface flexibility.

Acknowledgements

The authors would like to acknowledge the inspiration and contributions of the creators and developers of the first Places & Spaces illuminated diagram: W. Bradford Paley (typography, graphics, and interaction design), Kevin W. Boyack (scientometrics and data shaping), John Burgoon (geographic mapmaking), Peter Kennard (system design and programming), and Richard Klavans (scientometrics and node layout). The data underlying this exhibit was courtesy of Thomson ISI. The authors would also like to acknowledge the contributions of the Places & Spaces exhibit team, including Elisha F. Hardy, Weixia (Bonnie) Huang, Bruce W. Herr II, Julie M. Smith, Bryan J. Hook, and Mark Price.

The music similarity examples and data are courtesy of D. Scott McCaulay. Visualizations were created with the Prefuse information visualization toolkit via a plug-in module for the Network Workbench, a large-scale network analysis, modeling, and visualization toolkit (NWB Team 2006).

Portions of this work were supported by the National Science Foundation under grant no. IIS-0238261, CHE-0524661, IIS-0513650, IIS-0737783 and IIS-0715303. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Allcock , W., Bester, J. , Bresnahan , J., Foster, I., Cawor, J., Insley, J. A., Link, J. M., and Papka, M. E. (2002). GridMapper: A Tool for Visualizing the Behavior of Large-Scale Distributed Systems. *Proceedings of the 11 th IEEE International Symposium on High Performance Distributed Computing HPDC-11 2002 (HPDC'02).*

Boyack, K. W., Klavans, R., Paley, W.B., and Borner, K. (2007). Mapping, Illuminating, and Interacting with Science. *In ACM SIGGRAPH 2007 Sketches*, San Diego, CA.

Card, S.K., Mackinlay, J.D., and Shneiderman, B. (Eds.) (1999). *Readings in Information Visualization*: Using Vision to Think, pp. 1-34, Morgan Kaufmann Publishers, San Francisco, CA.

Hardy, E., Stevens, A., Courtney, C., Ambre, S., Neirynck, T., Hook, B., and Smith, J. (2009). *Places & Spaces: Mapping Science*. Accessed on 4-2-2009 at http://scimaps.org.

Keim, D. A. (2002). Information Visualization and Visual Data Mining, *IEEE Transactions on Visualization and Computer Graphics*, 8 (4), 395-6.

Leigh, J., Johnson, A., DeFanti, T., et. al. (1999). A Review of Tele-Immersive Applications in the CAVE Research Network. *Proceedings of IEEE VR `99*, 180-187, Houston, TX

McCaulay, D. S. and Sheppard, R. (2008). PFASC – a parallel framework for audio similarity clustering. *Proceedings of TeraGrid '08*, Las Vegas, NV.

NWB Team (2006). *Network Workbench Tool*. Indiana University, Northeastern University, and University of Michigan. Accessed on 4-28-2009 at http://nwb.slis.indiana.edu.

Paley, W. B. (2002). Illuminated diagrams: using light and print to comparative advantage. *Proceedings of IEEE Symposium on Information Visualization (InfoVis 2002).*

Paley, W.B. (2008) *Map of science in the journal Nature, SEED and Discover Magazines.* Accessed on 4-2-2009 at http://didi.com/brad/mapOfScience/index.html.

Yeh, R.B., Brandt, J., Klemmer, S. R., Boli, J., Su, E., Paepcke, A. (2006). *Interactive Gigapixel Prints: Large Paper Interfaces for Visual Context, Mobility, and Collaboration.* Stanford University Computer Science Department Technical Report.