Making Visible the Invisible: A Data-Driven Media Artwork, in Continuous Operation for 15 Years

George Legrady and Rama Hoetzlein

ABSTRACT

Making Visible the Invisible is a six-screen, dynamic data visualization artwork at the Seattle Public Library. It visualizes patrons' library checkouts received by the hour through four different animations to give a sense of community interests. The artwork was activated in September 2005 for a 10-year operation and was extended until 2019. As one of the longest-running media arts projects based on live, continuous data, the work has presented significant challenges regarding robustness and durability over time. We review the motivation, design of the system, backup plan, maintenance and novel strategies for a decades-long media artwork.

George Legrady
Media Arts & Technology
University of California
Santa Barbara, CA 93105
U.S.A.
glegrady@ucsb.edu

Rama Hoetzlein
Sentient Artists LLC
Oakland, CA 94610
U.S.A.
ramahoetzlein@gmail.com

A Public Commission

Making Visible the Invisible is an installation commissioned by the Seattle Arts Commission for the new Seattle Central Library in response to an open call in March 2003 (Fig. 1). The competition advertising asked for "artists whose works deal with social structures, people and systems." Out of 286 respondents, four finalists were selected and invited to participate in a weeklong residency in September 2003 to meet representatives of all levels of the operational infrastructure, the intent being to provide insights to assist in the conceptual development of the art proposals. The artists were asked to specifically address the functions, patrons and ideals of the library. Following extensive prototyping, a concept was developed to create a dynamic visualization artwork that explored the library as an information exchange system [1].

Early Development

During the concept development phase, it became apparent that with hardware costs of computer purchases, LCD screens and electrical installation at around $40,000, the artwork could not be realized within the $50,000 budget given that the primary focus in such a computationally intensive process involved significant experimentation in software development for the data analysis, visualization and design of the distributed information flow. The Committee of 33, a Seattle women's organization that annually supports the production of an artwork, stepped in and contributed an additional $55,000 to
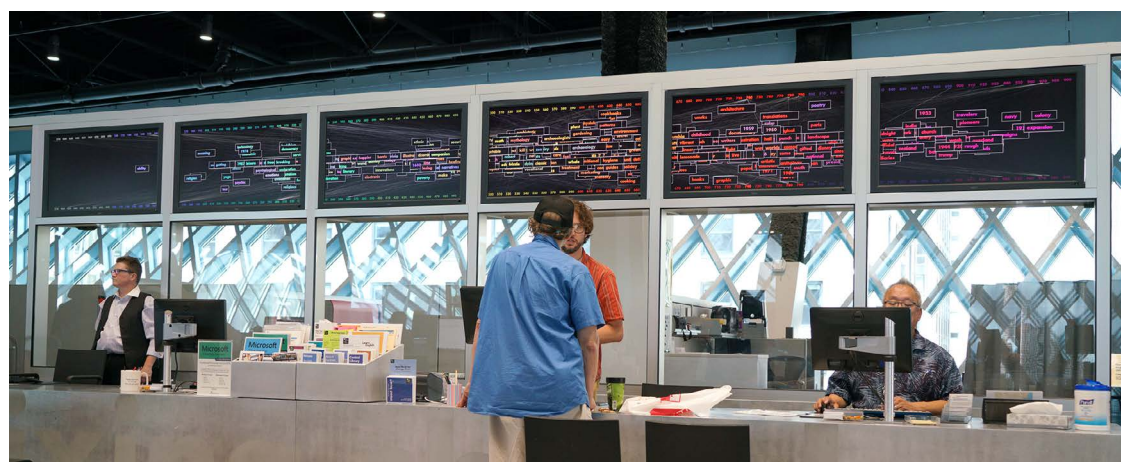

Fig. 1. The *Making Visible the Invisible* installation consisting of a distributed system, data processing with real-time visualization on six screens at the Seattle Public Library. (© George Legrady. Photo: George Legrady, 2018.)

make the project possible. The commission was approved in July 2004, and data analysis and visualization prototyping transpired over a 13-month period at the Media Arts and Technology program, UC Santa Barbara, with researchers August Black and Andreas Schlegel. The complete system hardware design, software production and stability testing, all with offline limited data samples, were ultimately realized by Rama Hoetzlein over a 10-week period in summer 2005. The onsite installation with actual real data testing happened in the first week of September 2005 [2].

Design Details

Making Visible the Invisible is an electronic, dynamic data visualization artwork. It consists of aggregating patrons' checkouts received by the hour and visualizing the results through four different animations as a way to make visible current topics of interests at any given time. The artwork receives its data from the library's IT server, which provides a precise, anonymized list of what items patrons have checked out during the previous hour. The item may be a book, a DVD, a CD, a VHS tape or any other item in the collection, for instance, sheet music. The database is refreshed every hour as new information is provided. Each item is represented by a rich set of multivariate data that includes checkout time (and later the return time stamp), title, media, acquisition date, bibliographic classification information, call numbers, associated keywords and unique IDs such as barcodes or collection codes that identify which library branch the item belongs to. The Seattle Central Library organizes its nonfiction items by the Dewey decimal classification system, an ideal system for visualization given its precise numerical categorization system. Items that are classified as fiction have a less precise labeling system, but nonetheless allow for analyses of media type, subject classification and time-based activities. Approximately 16,000 to 25,000 items are checked out per day starting at 10:00 am, with heaviest activity between 1:00 pm (13:00) and 5:00 pm (17:00) and slowing down at 8:00 pm (20:00) when the library closes.

The screens feature four animation visualizations through which the hourly data is presented (Fig. 2). The sequence of animations proceeds from the literal to the chronological, then 2D spatial and concludes with an expressive animated display of keywords from titles. "Vital Statistics," the first animation, presents through numerical statistics how many books, media, Dewey and non-Dewey items have been checked out in the last hour and since the morning. This is followed by "Floating Titles," a linear chronological
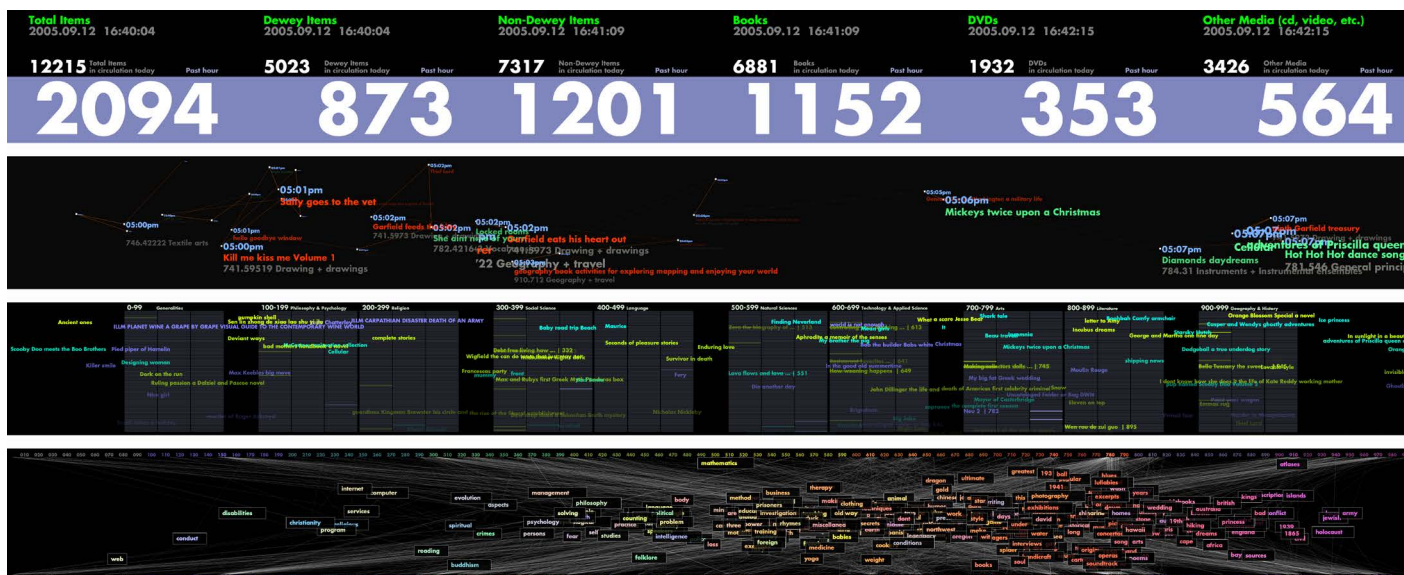

Fig. 2. The four visualizations, top to bottom: "Vital Statistics," "Floating Titles," "Dewey Matrix Rain," "KeywordMap Attack." (© George Legrady. Screen Capture: George Legrady, 2006.)

Legrady and Hoetzlein | Making Visible the Invisible: *A Data-Driven Media Artwork, in Continuous Operation for 15 Years*

sequence of the items checked out. "Dewey Matrix Rain" is a two-dimensional mapping of Dewey activities with non-Dewey items slowly moving vertically down and out of the screens. "Keyword Map Attack" is the fourth animation, an animated, colorful visualization of the most popular keywords collected from titles of checkouts, their color and position on the screen determined by the averaged Dewey affiliations.

As of winter 2019, the artwork has recorded over 92 million patron interactions with the library, providing major cultural insight into the interests of the downtown Seattle community and indirectly registering the library's transitional role in the second decade of the Internet. Due to concerns of technological longevity, the artwork was initially scheduled for a 10-year period (2005–2014) and was then extended to 2019. The library IT department had initial misgivings about the project's robustness given its origins as an artistic work, but the work has successfully operated over a 14-year period with minimal maintenance. It may be the longest-running data visualization art project in operation to date.

The system's operation consists of two components: a) hourly data processing of incoming records on a server, and b) the continuous visualization of the results viewed through four sequenced animations on six large adjoining electronic screens above the main information desk on the "Mixing Chamber" floor. The placement of the visualization was finalized after multiple discussions with the architects. This location fit with the conceptual and aesthetic intention of the project to create a mood of "information exchange" somewhat like the check-in counters at airports or stock exchange information visualization, which incidentally fit right in with the architects' concept sketch of the Mixing Chamber area as a "kind of trading floor" (Fig. 3) [3].

Context and Relevance

Making Visible the Invisible is a singular artwork in the field of public commissions as it dynamically taps into and visualizes the infrastructure of the library. Whereas nearly all public commissions are fixed, or motorized, this artwork is data-driven, its content continuously changing, a visual feedback system of what the community is interested in. The new Seattle library's innovative vision made this project possible given all the challenges that such a work may entail. Due to its early development in the mid-2000s and that it has been operational for such a long time, it is one of the most significant projects in the literature of data visualization artworks, a multidisciplinary field of interest to media arts, architecture, public art and academic research. It has also paved the way for new public artworks that explore data as visualization. It was followed by two other public art projects: a) *Kinetic Flow* commissioned by the Los Angeles Metro

Fig. 3. Installation location at the Mixing Chamber information booth on the third floor of the library. Left: "SPL Service Bar," architects' concept. (© Seattle Public Library) Right: art installation. (© George Legrady. Photo: George Legrady, 2007.)

rail for its Vermont/Santa Monica station, and b) in 2007, a one-year data visualization installation titled *Data Flow*, mounted in the executive spaces of CEB (The Corporate Executive Board Company), a global "best practice insights" company in Arlington, Virginia [4]. The artwork's reputation brings specialists from all over the world to experience it in its site-specific context. No other public artworks explore data in a public context in such detail.

The Artwork as a Feedback System

The artwork was designed as an onsite feedback system by which patrons and librarians would be able to perceive the flow of cultural content in or near real time. The four animations have been designed to provide a view as to the process by which the library could be considered as an interactive data center. Patrons have been motivated and inspired by the flow of titles, a form of suggested browsing, which has led many to retrieve items they would not have thought of otherwise. Librarians have appreciated the artwork driven by their interest in having a continuously updated overview of what categories, what media and to what degree the Dewey classified nonfiction items have circulated in contrast to the non-Dewey fiction titles. Part of the maintenance process has involved the need to oversee the data flow, leading to the creation of an online window at 128.111.26.109/parsing/index.php, where one can compare the total number of current checkouts and their breakdown into books, media, Dewey and non-Dewey classification items. The site also provides the chronologically sequenced set of titles of the top 20 Dewey categories. Comic books and cookbooks permanently retain the top circulation followed by shifting topics of poetry, health, folklore, general history, zoology and others depending on the time of the day.

Even though the animations only feature the previous hour's activities, the artwork's server has kept an hourly historical record of all checkouts since the beginning of its operation, thereby extending its function into a long-term live database, foreshadowing current interests in big data and data science. The stored data has in fact been used to teach a visualization course each year at UC Santa Barbara to explore how the data can reveal changes in trends over time, and to also develop new visualization studies (Fig. 4). Through this pedagogical context, assignments have generated a broad range of insights in both cultural trends and how such a database functions.
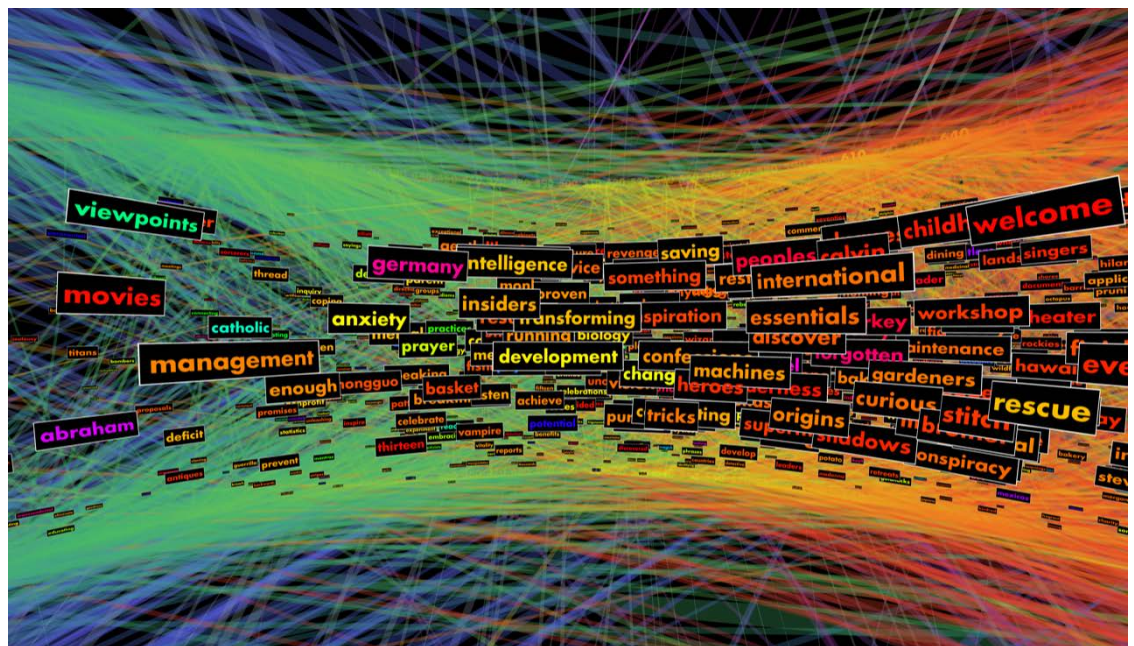

Fig. 4. Seattle library data keywords spatialized study for the Advanced Visualization and Interaction Environment's 360° stereoscopic system at the iCinema, University of New South Wales, Sydney, February 2018.

Technical Overview

Making Visible the Invisible consists of six plasma displays operated by three visualization computers with each running two displays (Fig. 1). These client systems are networked together for synchronization and connected to a data analysis server. The data server is an intermediary system that interfaces with the Seattle library circulation server on one side, then collects, processes and stores data in a compact format and forwards the data to the three client systems for display (Fig. 5). Each connection is made by a high-speed network internal to the library infrastructure and all placed behind a firewall. The software for both the data analysis server and the visualization clients was custom-made for this project and written in C++. The visualization uses the GamesX open-source graphics engine designed by Rama C. Hoetzlein [5]. Decisions about the operational design were influenced by imagining the longevity of the project.
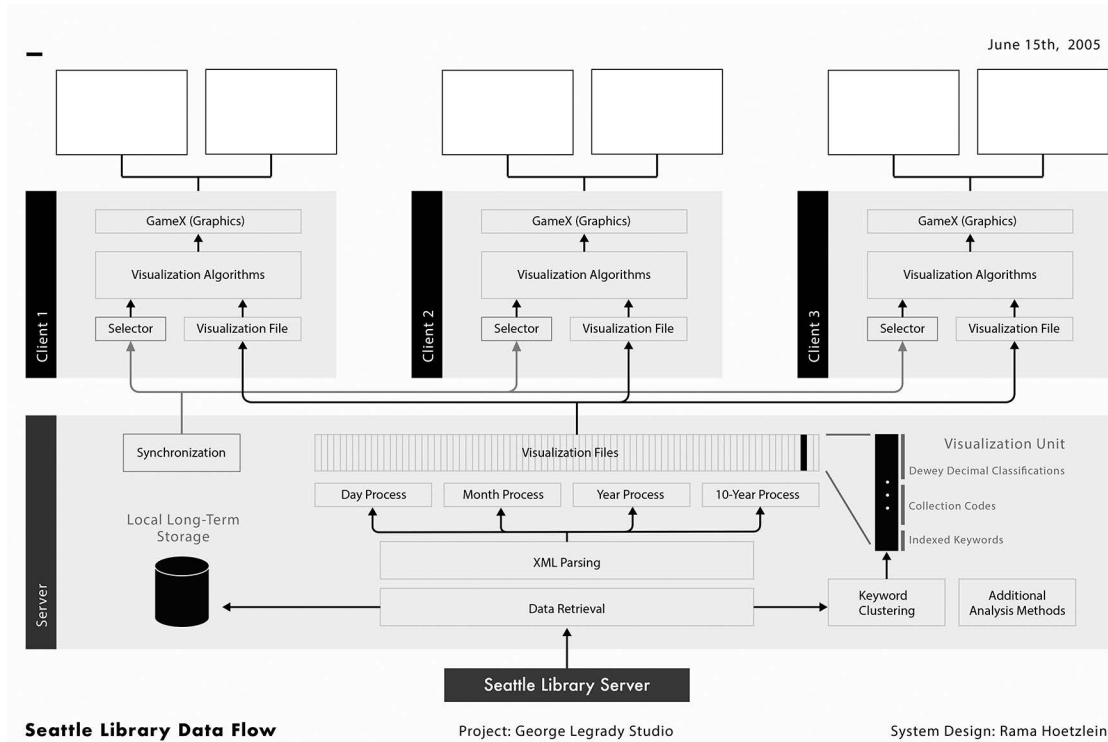

Fig. 5. Hardware and data flow layout of *Making Visible the Invisible*. (© Rama Hoetzlein. Sketch: Rama Hoetzlein, 2005.)

Planning for Longevity

Software that operates robustly over long periods of time is common in space science [6], aeronautics and other continuously operating, mission-critical programs [7]. Since our project is a media artwork, our goals are different from these yet share some common features. We consider *Making Visible the Invisible* to be a mission-critical project in the sense that it is highly visible and must operate continuously without interruption. However, it is not a life-critical project and is also not autonomous (like a satellite), with periodic maintenance permitted so long as it does not visibly affect the system. Design of the software was based on these constraints with a focus on redundancy similar to that found in other critical systems.

To plan for longevity we physically separated the data collection from visualization. The visualization client computers act autonomously and make requests to the data analysis server (Fig. 5). This pull operation is more robust than pushing the data from the server. If no data is available, for whatever reason, the visualization computers contain backups that can be shown in the interim. Thus, the visual clients are intended to always operate regardless of any momentary data problems.

The data analysis server is also self-sufficient, with original data from the Seattle Central Library's system placed on the data server roughly every hour. Since the exact time of "data drop" is unknown, the system continuously polls every 10 seconds to check for data. Once new data is received the analysis server collates it; processes it into month, day and year segmentations; and stores it compactly. With our customized compact format, all data over the 10-year period was designed to fit in under 200 GB, eliminating the need for cloud storage. The analysis server will typically serve up the most recent, current hour of data for display but is free to serve up past data if any momentary issues arise. With this redundant design at both the display and data server, the displays have been designed to be continuously operating.

A final layer of robustness against hardware issues protects from momentary outages. This software layer, like other critical systems, is capable of self-restarting. The display clients are constantly checking one another for crashes. Should a system crash occur on any one of them, they will automatically signal the others and initiate a restart procedure. Finally, if the system is unable to detect a problem itself, a single button press by library staff allows the entire system to restart.

Longevity of a new media artwork presents specific challenges. Multiple layers of robustness are needed to ensure the display continues even with momentary data outages. Adopting practices in mission-critical software for robustness has influenced our design for multiple backups and multiple autonomous and manual restart procedures.

Conclusion

Making Visible the Invisible explores cultural narrative through its continuously changing cultural data and its visibility at the Seattle Public Library. The fact that it uses real circulation data from the library, and has lasted over 10 years, demonstrates that new media artworks can have a lasting and continuous impact on a community that remains relevant and meaningful.

References and Notes

1. G. Legrady, "Culture, Data and Algorithmic Organization," *Leonardo* **45**, No. 3, 286 (2012).
2. G. Legrady, "Making Visible the Invisible: Seattle Library Data Flow Visualization," *Digital Culture and Heritage*, Proceedings of ICHIM05, Paris (2005).
3. G. Legrady and A. Forbes, "Data in Context: Conceptualizing Site-Specific Visualization Projects," *Leonardo* **50**, No. 2, 200–204 (2017).
4. G. Legrady and A. Forbes, "Staging Data Visualization Installations in Site-Specific Situations," *Proceedings of the IEEE VIS 2014 Arts Program*, 9–14 November 2014, 85 (2014).
5. R. Hoetzlein and D. Schwartz, "GameX: A Platform for Incremental Instruction in Computer Graphics and Game Design," *ACM SIGGRAPH Educators Program*, Los Angeles, CA (2005). Additional info: gdiac.cis.cornell.edu/resources/GameX/index.htm.
6. A. Shahrokni and R. Feldt, "A Systematic Review of Software Robustness," *Information and Software Technology* **55**, No. 1, 1–17 (2013).
7. A.K. Ghosh and J.M. Voas, "Inoculating Software for Survivability," *Communications of the ACM* **42** (1999) pp. 38–44.