Staging Data Visualization Installations in Site-Specific Situations

George Legrady
Media Arts & Technology Program
University of California, Santa Barbara

Angus Graeme Forbes
Department of Computer Science
University of Illinois at Chicago

ABSTRACT

Five public artworks are presented that feature data visualization situated in site-specific installations, each with distinct conditions of data collection, data analysis, audience interaction, and data archiving. This paper describes unique features of these projects and notes solutions to issues related especially to the staging of visual representation. Based on an examination of these issues, observed in different settings throughout the last decade, the paper concludes with an overview of key considerations specific to presenting data visualization in site-specific locations.

1 INTRODUCTION

The field of data visualization explores the creation and analysis of visual representations of information. In most cases, data visualization projects translate abstract data into a visual form that makes it easier to perceive relationships in the data [20]. Representing data in visual form is a complex practice that is similar in approach to the structuring of language through rules of syntax and grammar where discrete units such as characters and words are combined to create meaning. Through the ordering of visual primitives, such as form, space, color, line, dimensions, scale, balance, texture, direction, and motion, visual representation similarly involves a construction of meaning [1]. Kosara’s discussion of visualization criticism underscores that visualization is both context and goal dependent, and that intention determines a visualization’s emphasis on communication clarity and/or visual impact [6]. Data visualization, which began as a critical tool in scientific inquiry and other disciplines necessitating the statistical mapping of data, has in the past decade crossed-over into the arts, becoming a new hybridized artform [4, 15]. This new genre explores visual form through computational means, and has led to more abstracted outcomes with an emphasis on aesthetic exploration.

Standard delivery of data visualization results are normally presented in print or on screens. Interaction is usually limited to conventional keyboard and mouse use mediated by a graphical user interface. In most cases, the visualization is independent of the context within which it is presented. Public presentation of data visualization can take either of two approaches. In the first approach, the content has no connection to the site where it is presented. For example, the location of a billboard is usually predetermined and the lifespan of each particular advertisement it displays is usually brief. In the second approach, the visual display has specific functionality or relationship that creates an engagement with the site where it is located, perhaps featuring visualized content that originates or reflects back onto the site. Signage that overtly announces a building’s function and that has a direct indexical relation to its location is one example of this. A site-specific location, such as Times Square, may impose visual elements that, by virtue of it’s location, function symbolically to signify spectacle.

In the early 1970s, the architects Robert Venturi and Denise Scott Brown published “Learning from Las Vegas,” a challenging, seminal book about the role of spectacle signage, commercial billboards, and electronic signs that illuminate the Las Vegas urban landscape [18]. Their analysis acknowledged that the billboards, generally considered as an aesthetic detriment to the landscape, were in fact significant, integrated architectural elements, and that it was necessary to acknowledge their impact in the scenery.

In the same time period, the artworld began to expand its activities beyond exhibiting artworks in the enclosed spaces characteristic of galleries and museums to creations in the environment and other sites considered public or outside the traditional context of artistic and visual experiences [8]. This new direction in art, to some degree stimulated by public commissions, set in motion a research-style approach to analyzing the information layers of site-specificity. It implied that artworks needed to be created so as to take into consideration the impact of where they were situated, and how an artistic intervention would transform the site. Such an approach is well exemplified by the artist John Roloff’s statement describing what his planning methodology involves: “strategies employing inversions, intrusions, displacements, assemblages and extended analogies/metaphors, often in geologic parlance, of existing, often predictable, ecological beliefs and systems in order to disrupt, re-cast and extrapolate their epistemological, ontological and associative potential” [14].

Figure 1: A photo showing a gallery visitor contributing data to the Pockets Full of Memories installation.

The conceptual development of temporary and permanent site-specific data visualization projects involves extensive study of the contexts within which the work is to exist. This is similar to the preliminary studies necessary for the planning of an architectural work. These include: a review of the function of the space; what information is generated within the space and the use of the space; what kinds of human circulation take place during what time periods; what are the lighting and acoustic conditions; and what may be best locations where to situate the space, among other considerations.
2 PROJECT DESCRIPTIONS

In this section, we present five data visualization installations, each of which was created principally by the first author within the past 13 years. Each work will be described with a brief summary of what it is, the nature of the team involved in its realization, what data source was used, and how this data was collected, analyzed, and visualized. These projects are introduced as a means to discuss the site-specific conditions of public installations and issues of staging data visualization within these locations.

2.1 Pockets Full of Memories (2001-2007)

*Pockets Full of Memories* is an art installation commissioned by a major European museum with the intent to activate the general public to create an archive of cultural data. Production included an international team: Multimedia engineering with software and hardware production was realized in Budapest; graphic and interface design was realized in Stuttgart; an industrial psychologist, also from Stuttgart, contributed a semantic differential questionnaire and data analysis; a computer science artificial language expert from Helsinki implemented an artificial neural-net algorithm by which to organize the collected data. Software development took place in Helsinki, Budapest, and Santa Barbara [9].

The project is a database installation where the audience contributes a digitized image of an object in their possession and inputs textual descriptions about that object through an interactive questionnaire. The sum of the archive of objects is continuously re-organized by a self-learning training algorithm with its result dynamically visualized on a large projection in the museum, the positioning objects of similar descriptions in precise proximity to each other, based on the computational optimization that the algorithm produces over time. The collection is empty at the start of the exhibition, increasing over time with each new contribution. In the process of building the collection at each site, each results in a specific historical record and expression of the venues’ specific contextual conditions. The contributions can then be statistically analyzed to compare subject matter and explore the variations in the semantic descriptions. During each exhibition, the specific collection was accessible both in the museum and also on the Internet where the public was able to contribute further by adding comments about any of the objects.

Each contribution consists of a scanned object with associated semantic data. The object is classified via an Osgood semantic differential slider [16] which requires contributors to provide a scale value for the object’s properties, such as age, hardness, and to what degree it is natural, personal, fashionable, useful, functional or symbolic. Each object is also described by three keywords and a description. Additionally, demographic information is requested regarding the name, age, gender, occupation, and origin of the user, all of which are used by the algorithm to calculate each object’s semantic value in relation to every other object in the collection.

The visual component of *Pockets Full of Memories* is featured on a large cinematic projection that displays a 2D map divided into a 12x32 matrix of 384 cells. Each object in the collection is positioned at a specific location in the two-dimensional grid. Items with similar descriptive features are placed closer to each other, and the overall order of the map is an outcome of all the inputs; when all cells are filled, older contributors are placed behind recent ones, becoming invisible. The algorithm continuously recalculates the relationships arriving over time to a final ordered state, which is then disrupted when a new object enters the collection. Four animations feature the data in different ways, with the most recent contributions highlighted by bright yellow frames. One animation shows the objects and their relocation to new positions as the algorithm recalculates. This is followed by a visualization that shows lines indicating displacement for all objects. The greater the number of objects in the archive the more complicated the visualization becomes, due to the increased crisscrossing of lines. There is also an animation that gives a textual description of the objects along with the name of the contributor. Finally, the last visualization consists of each cell colored by a shade of grey to indicate its relative similarity to neighborhood cells. The darker the coloring the greater the value distance. Light colored cells represent clusters with similar values, whereas darker cells are cluster separators, denoting semantic distance, somewhat similar to mountain ranges which separate valleys.

The staging of the installation follows the traditional cinematic layout with a large screen at the front of the installation. A data input kiosk box is positioned in the middle of the space, and two internet accessible computers are placed near the rear of the space to allow viewers to browse the previously input objects’ metadata and to send e-messages to each object. This format emphasizes an effective spatial distribution that allows data contributors to group around the kiosk while other viewers can instead study the large visualization in clusters at both ends of the space. The staging is further enhanced by lighting so that the front space displaying the projection is darkened to better perceive the image while the back area is lit to allow individuals to easily position themselves. Figure 1 shows an picture of the *Pockets Full of Memories* installation, including the interactive kiosk for contributing objects and the internet-enabled computers. Figure 2 visualizes a statistical study of how cellphones where described at each of the eight exhibitions. Because of the diversity of visitors and the variations of each exhibition’s context and duration, one can see a diversity in approaches as each venue features variances, except that all the venues have the same highest value for the eight semantic differential scales—The cell phones, according to this study, are universally considered to be new, hard, synthetic, personal, fashionable, useful, and functional.

Figure 2: A data visualization showing a comparison of how cellphones have been rated in the eight exhibition venues of the *Pockets Full of Memories* installation.

The exhibition took two years to plan and produce. It opened at the Centre Pompidou, Paris in the summer of 2001 to a large international general audience. Approximately 20000 people visited the installation, and approximately 3300 contributed objects. It proceeded to the Dutch Electronic Arts Festival, Rotterdam (2003), to Ars Electronica, Linz (2003), and to Aura, C3 Center for Culture & Communication, Budapest (2003), each of which were short exhibitions attended primarily by media arts-savvy audiences. The installation then travelled to the Museum of Contemporary Art Kiasma, Helsinki (2004), Cornerhouse Gallery, Manchester (2005), Frankfurt Museum of Communication (2006), and ultimately concluded at the Museum of Contemporary Art, Taipei (2007). These various locations provided a broad international sampling of over
2.2 Making Visible the Invisible (2005-2019)

Making Visible the Invisible was commissioned by the Seattle Arts Council for the Seattle Central Library. It is housed in a building designed by Rem Koolhaas. The installation consists of 6 large LCD screens, positioned side-by-side, and located on a glass wall that stands behind the librarians’ main information desk on the 2nd floor of the library, in an area labeled “Mixing Chamber.” The screens feature animated algorithmic visualizations generated via custom statistical analyses of real-time data. Prototyping and production took over two years to complete and included the efforts a small team number of media artists and engineers [10].

The artwork receives its data from the library’s IT server, which provides a direct list of what items patrons check-out at any given time. The database is refreshed every hour as new information is provided. This multivariate data is organized in chronological order. The item may be a book, a DVD, a CD, or a VHS tape, and a rich set of metadata is also sent to the project, including check-out and check-in times, titles, keywords, call numbers, and unique IDs (such as barcodes or collection codes).

Long term discussions with IT over what may be appropriate metadata for the library to share while protecting patrons’ privacy and the system’s integrity was at the core of the negotiations. The understanding of what the data is and does, and how it serves the institution which collects it, is the greatest challenge for any data visualization project that uses live data belonging to an organization. The Seattle Central Library organizes its non-fiction items by the Dewey classification system, an ideal system for visualization due to its precise numerically labeled categorization. All subject topics are indexed in a decimal structure between 0 to 999,999 to allow inclusion of any publication in a relative location so that it can be placed in the library collection. 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 25000 items are checked-out per day, adding up to about 900,000 per year and totaling over 75 million since the artwork’s inauguration in September 2005.

Four animations repeatedly cycle every few minutes, each presenting the newly received library data. The sequence of animations begins with numeric statistics, then proceeds to a linear chronological sequence of the items checked-out, followed by two-dimensional mapping to give an overview of how all the data in the hour have performed. The fourth animation consists of an animated, colorful visualization of the most popular keywords in the titles check-outs, their color and position on the screen determined by the averaged Dewey affiliations.

The placement of the visualization was finalized after multiple discussions with the architects. Ultimately is decided that it would be best positioned above and behind the main information desk situated on the second floor. This location fit well 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 info visualization, which incidentally fit right in with Koolhaas’ concept of the “Mixing Chamber” area as a “kind of trading floor” [5]. Figure 3 shows a photo of the project installed in the Seattle Public Library.

A specific challenge for this project was its longevity. Initially intended to function as a ten-year project, its lifespan has been extended for another 5 years. Indeed, it may be the longest, continuously existing dynamic data visualization project in existence. Key issues have included planning the data storage and archiving and the software and hardware maintenance. A specific built-in challenge for a long-running project is the potential for technological obsolescence [13]. The software has so far managed to be fully operational but at some point it may be necessary to migrate or reconstruct its operations, since future hardware might not continue to be compatible with the current operating system in use.

2.3 We Are Stardust (2008)

We Are Stardust is a two-screen projection installation that replays the sequence of 36064 sky observations by the NASA-launched Spitzer Satellite telescope1 from the start of its mission in 2003 to the summer of 2008 (which is when the artwork was commissioned by the Spitzer Space Center and the Art College of Design, Pasadena). The Spitzer Space Telescope is an infrared temperature sensing instrument that is orbiting the sun, trailing in the earth’s orbit. The intent of the work was to create a project based on data produced by the satellite. The initial idea was to chronologically map the sequence of observations, to visually reveal what the sequence of nodes on a flattened universe map would reveal. It later became compelling to also integrate metadata about each observation such as the name of the star, the distance, location, time/date, scientist/lab, and which of the 4 instruments was used. The production team consisted of a small group of artists and engineers, each of who helped with aspects of the engineering and 3D visualization2. The installation was featured at the Art Center College

---

1http://tinyurl.com/k58l4ez
2http://tinyurl.com/kcvexcyc
of Design (2008) as part of a group exhibition titled “Observe”\(^3\). It was also featured in the CODE LIVE exhibition at the 2010 Vancouver Olympics [12].

The installation consists of two projections on opposite walls of a large exhibition space. Each of the projections chronologically sequence through the history of the satellites’ observations. A full view of the universe with constellations which is overlaid with the traces of the Spitzer telescope’s observations is at one end of the gallery space. The full sequence requires five hours to cycle through the five-year activities, a length of time beyond what a gallery visitor is expected to experience, so the full sequence is subdivided into five minute segments, each of which begins with an animation of the “birth” of the sky from random to the ordered location of constellations as we know it. At the end of the five minute sequence, all of the 36064 observations are briefly featured.

At the opposite end of the gallery space, another projection, also consisting of the sequence of the same Spitzer observations, provides a radically different perspective. With the use of a military grade, heat-sensing, mobile infrared camera, the inside of the gallery is visually scanned according to the adjusted angle-of-view of the original Spitzer telescope coordinates. This visualization provides a colorful, engaging heat-based view of gallery spectators who can position themselves under the gaze of this searching, moving optical instrument that maps their own thermal presence and actions onto the large screen. Figure 5 shows a photo from the We Are Stardust installation.

The staging of this project was dictated by the necessity to provide two points of views, one focused on mapping the observation data onto the universe, and the other engaging the audience in the local, enclosed space of the gallery. This polarized differentiation dictated the positioning of the two projections at opposite ends with the heat-sensing camera mounted on the ceiling in the middle of the gallery space.

2.4 Cell Tango (2006-2009)

Cell Tango is a dynamic artwork presented as an interactive installation in fine arts museums, in galleries, and at special events [3]. The public visiting the exhibition is invited to participate as a contributor to the project through the submission of cell phone images with tags sent from their phones to a database. These images, and the accompanying tags that categorize and describe them, are projected large-scale in the gallery, continuously shifting as new contributions are added. The images are organized by a database which uses the images tags to search and retrieve other associative images from the online Flickr\(^4\) photo sharing website.

The installation features different visualizations of the images in the four different animations. The animation ‘Cell,Bin’ selects the most recent images in the database on the screen using an algorithm that scales the images, then places large images first and gradually fills in blank spaces with smaller images until all of the empty screen spaces are filled. ‘Cell,Clusters’ consists of thematic clusters of found images based on contributors’ tags placed around each incoming contributed images that are marked with a yellow frame. The ‘Cell,Burst’ animation throws images on the screen that then open like bursting fireworks, placing the tags once the image appears, followed by the retrieved Flickr images that are associated with each tag. ‘Cell,Finale’ concludes the visualization sequence by positioning the full collection of submitted images randomly on the screen.

Cell Tango was presented at multiple venues, each time being enhanced as different aesthetic ideas were more thoroughly developed. It premiered at the International Society of Electronic Arts (ISEA 2006) conference, San Jose, in 2006. It was then included in a broad range of venues, including a solo exhibition in a commercial gallery (Pari Nadimi Gallery, Toronto, 2006) and in the group exhibition “Speculative Data and the Creative Imaginary” (National Academy of Sciences Rotunda Gallery, Washington, DC, 2007). It also participated in the inauguration of the National Theatre Poitiers, Poitiers, France (2008), and was exhibited at a public university, (Eastern Michigan University, Ypsilanti, 2008-2009), in a media arts exhibition (Beall Center for Art & Technology, UC Irvine 2009), and at Wellesley College (2009). Sonification was added\(^5\) and premiered at the Lawrence Hall of Science, UC Berkeley (2010), and was incorporated again during the last presentation at the Poznan Biennale, Poland (2010).

The project’s unique strategy is to integrate a curated set of data (the cellphone images submitted by the public) with an open-ended selection from the public Flickr photography database. The selection from Flickr is based on tags culled from contributors’ submissions. The conceptual premise has been to explore the potential of unexpected juxtapositions where common semantic labels function to generate visual relationships. Gallery viewers and participants were able to submit images from anywhere, not necessitating their presence in the gallery. The site-specific conditions for this work

\(^3\)http://tinyurl.com/lj8qgle

\(^4\)http://flickr.com

\(^5\)http://www.cj.lovelyweather.com
has been very flexible as it only required a large screen surface on which to project the animations. Figure 6 shows a photo of the installation at the Poznan Biennale.

2.5 Data Flow (2009)

Data Flow was commissioned by Gensler Design for the Corporate Executive Board (CEB) Corporation in Arlington, Virginia. The project consists of three visualizations that map time-based communication exchanges between “members” (that is, users of the CEB website) and the CEB document repository. The artwork location was predetermined in advance, and was situated on the corporation’s “Feature Wall” in the staircase areas extending from the 22nd to 24th floor. Following a study of the data and the architectural space, three locations were selected to position the three sets of screen. Depending on where the viewer was positioned, most locations on those floors allowed for viewing of at least 2 screens simultaneously. The project was installed for just over a year and took approximately six months to develop.

The flow of information from the CEB database to the installation consists of the following: CEB IT produces appropriately formatted and anonymized data which is retrieved every ten minutes by the custom Data Flow project server which is then stored in a local database, where it is kept for 24 hours. The project server also retrieves the longitude and latitude of each user’s (approximate) location and discards any data that does not correlate with the requirements of the visualizations. This data is then forwarded to three visualization computers that each process the received data according to their individual animation requirements.

Each of the three animations analyze incoming members’ messages from different perspectives. The top level screen, titled “BIGRAMS”, retrieves the most recent 20000 events and analyzes them for the most popular 12 keywords. Links show which industry members’ words come from, and to which “Programs” (subsections of the CEB website) they are primarily associated with. (Each member is classified according to their practice, which may be in the areas of finance, government, health, human resources, information technology, innovation strategy, legal, marketing, or operations and sales.) The middle screen titled, “GEOMAP”, animates and deforms a grid-based map of the world according to the location where members are checking-in from [2]. The animation receives the 500 most recent events, and then sequentially expands cells that represent the active members’ locations. Each grid increases in size based on the length and number of actions. Every three minutes the distorted map collapses back to its original form. The animation indicates the sweep of activity as it moves westward over the course of each day. The lower screen, titled “PROGRAM TO CONTENT”, proportionately maps how various content choices and downloads are used by members from different membership categories. First, the animation selects the most recent 1000 events. Of these most recent events, the animation visualizes the top 15 most active categories.

Staging data visualization art within a corporate environment presents its own opportunities and challenges, including involved negotiations with administrators and in-house designers, and heightened awareness of the need to anonymize and limit the data being used so as to respect customer privacy. Additionally, there was an ongoing concern regarding the possibility of inappropriate words appearing in the “BIGRAMS” visualization. A data visualization in this type of environment thus must, for better or worse, navigate a range of potentially competing interests. Figure 7 shows the “GEOMAP” and “PROGRAM TO CONTENT” data visualization from the vantage point of the 24th floor of the CEB headquarters.

3 Conclusion

This paper has addressed the conditions by which large-scale site-specific data visualizations require extensive formulation in regards to their staging and presentation. Five artistic commissions engaged with visualizing data have been presented to address how each venue’s contextual circumstances, including location, positioning, lighting, scaling, and function, impact user perception and cognition. While some recent work has explored aesthetic issues, such as the effects of style on visualization [17], or how artistic elements can augment a data visualization [19], the examination of substantial public data visualization installations could still provide insight into how visualization can most effectively be staged in site-specific locations. For most of the projects described above, an additional analysis step took place after the conclusion of the life of the installation. In these analyses, the collected data provides an insight into which specific stagings were most effective and which encouraged a richer user interaction with the installation. In the case of the Making Visible the Invisible installation, the data collected from the library is being used to create further data visualization projects. Figures 2, 4, and 8 each show an example of these post-installation data visualizations.

Data visualization projects created as site-specific installations
additionally may integrate within its process the sampling of data from the site itself. This iterative approach is similar to architectural design methodologies where the condition of a building’s location and function impacts on its design [7]. This is further enhanced through technological sampling of data (such as demographics, noise, light levels, or traffic flow). The success of site-specific visualizations therefore requires extensive pre-production planning, in-depth negotiation with administrative and technical site managers, and consideration for what the long-term impact of the work may be as well as to what degree efforts are to be invested into the storage, archiving, and analysis of data collected from the installation.

REFERENCES