Creation and Collaboration: Engaging New Audiences for Information Visualization

Jeffrey Heer[1], Frank van Ham[2], Sheelagh Carpendale[3], Chris Weaver[4], and Petra Isenberg[3]

[1] Electrical Engineering and Computer Sciences,
University of California, Berkeley,
360 Hearst Memorial Mining Building, Berkeley, CA 94720-1776, USA,
jheer@cs.berkeley.edu

[2] IBM Research, Visual Communications Lab,
1 Rogers Street, Cambridge, MA 02142, USA,
fvanham@us.ibm.com

[3] Department of Computer Science, University of Calgary,
2500 University Dr. NW, Calgary, AB, Canada T2N 1N4,
{sheelagh, petra.isenberg}@ucalgary.ca

[4] GeoVISTA Center and the North-East Visualization and Analytics Center,
Department of Geography, Penn State University,
302 Walker Building, University Park, PA 16802, USA,
cew15@psu.edu

1 Introduction

In recent years we have seen information visualization technology move from an advanced research topic to mainstream adoption in both commercial and personal use. This move is in part due to many businesses recognizing the need for more effective tools for extracting knowledge from the data warehouses they are gathering. Increased mainstream interest is also a result of more exposure to advanced interfaces in contemporary online media. The adoption of information visualization technologies by lay users – as opposed to the traditional information visualization audience of scientists and analysts – has important implications for visualization research, design and development. Since we cannot expect each of these lay users to design their own visualizations, we have to provide them tools that make it easy to create and deploy visualizations of their datasets.

Concurrent with this trend, collaborative technologies are garnering increased attention. The wide adoption of the Internet allows people to communicate across space and time, and social software has attained a prominent position in contemporary thinking about the Web. For example, one can think of software teams distributed over different time zones or multiple people collaborating to build an online encyclopedia. Furthermore, collaborative issues are not limited to the web: novel display and interaction technologies, including wall-sized and tabletop interfaces, introduce new possibilities and challenges for co-located collaborators. An increased need for specialization means that we can no longer rely on a single person to perform deep analyses of complex phenomena. These

A. Kerren et al. (Eds.): Information Visualization, LNCS 4950, pp. 92–133, 2008.

developments signify an increased desire for collaboration around complex data, yet, information visualization tools are still primarily designed according to a single user model. To meet the demands of an increasingly diverse audience, the design of information visualization technologies will have to incorporate features for sharing and collaboration.

In this paper, we discuss creation and collaboration tools for interactive visualization. Our goal is to begin to characterize the increasingly diverse audience for visualization technology and map out the design space for new creative and collaborative tools to support these users. In section 2 we classify the expanding user base for visualization technologies by looking at their skills, goals and the data they are trying to analyze. We then take a look at existing information visualization tools and classify them along these dimensions. In sections 3 and 4 we examine the new collaborative trends. Section 3 discusses co-located collaboration, while section 4 explores the area of distributed, asynchronous collaboration on the Web. Finally, we conclude by considering the ways the research community should respond to these developments.

2 End-User Creation of Visualizations

The term "end-user visualization" encompasses a broad range of visualization users and use-cases. For example, a marketing executive might create an overview of the sales in different product segments to show to his manager, a scientist may create a coordinated visualization application to study a biomedical dataset, or a Facebook user may present her social network in a visualization on the site. All these use cases involve different types of users employing information visualization to tackle different types of problems. If we wish to provide end-users with the ability to construct and deploy custom information visualizations of their own data, we need an understanding of these users, their goals, and their data. In the following sections, we will broadly classify each of these dimensions. Note that we do not intend to construct a formal taxonomy of users. Instead, our goal is to broaden the discussion on who our users are and how visualization can help them.

2.1 Data

Scientific, geographic, economic, demographic, and other domains of human knowledge produce vast amounts of wildly different forms of information, varied in terms of both individual interest and broad social importance. Visualization seeks to provide perceptually and cognitively effective tools to display and interact with these different kinds of data. Data is commonly categorized by inherent complexity (e. g., data homogeneity, number of dimensions) or size. In this section, however, we consider data from the perspective of users by categorizing three different kinds of data in terms of potential audience.

Personal Data: Personal data encompass all types of organized information collections that are of personal interest to a particular user, but less interesting to a broader community. This may involve data on user-owned media (such as DVD collections or playlists), data on life organization (financial data or address books) or data related to hobbies and general interests (photo collections, fitness schedules or coin collections). Visualizing personal data might not always lead to deep new insights about the data itself. In such cases, the visualization instead may serve more as a compact visual artifact that can be used to remember certain events in ones life and serve as a visual representation of self [95]. These visual representations of self may then be used as online avatars, or simply as catalysts for storytelling, much like photo albums.

Community Data: By community data we mean data that might be relevant to a broad community of users due to similar interests or general appeal. Examples of community data include the content of political speeches, the number of users online in a World of Warcraft realm, or voting results per county. Often this type of data has a social component associated with it: data might be related to a social application such as Facebook or MySpace [45], contain statistics on a large population as with census data [43], or may be related to current events [104]. Precisely because community data has a lot of general appeal it will often generate a lot of discussion.

Scientific Data: Scientific data is data that is of interest to a (relatively) small number of specialists. Traditionally, information visualization has focused on the sciences, because they generate a wealth of structured and often numerical data in ready need of analysis. This makes them very suitable to mathematical analysis techniques and visual mapping. In the humanities, however, most information comes in unstructured raw text format. If we want visualization to be applied in domains such as literature and political science, we will need to define suitable pre-processing techniques that can extract meaningful information from a body of text. This will often require some amount of natural language processing or expert input. While there are a few applications of information visualization to data from the humanities (e. g., [101]), the area remains largely untapped despite substantial promise to yield many useful techniques with applicability to many different areas of everyday life.

Interplay of Data Types: Note that the distinction between types of data is not always clear cut and many data sets could fall into different categories depending on their use. For example, a community data set on World of Warcraft users and their interactions might be considered a scientific data set by social scientists, while the personal data of celebrities might have a broad general appeal. Visualizations of all these types of data can be shared, albeit for different purposes. Personal data might be shared with other users as a means of personal expression. Community data is often shared to spark broad discussion, while scientific data often needs to be shared because it is too complex for one person to analyze on their own or because it requires multiple specialized skills to analyze.

The recent trend toward visual analytics [91] is driven by the increasing need to support open-ended management and exploration of large, loosely-connected, and often unstructured information sources as well as the smaller, isolated, structured data sets typical of information visualization applications. Information collection often involves assembling "shoeboxes" of loosely related nuggets and data sets [107]. Visual analysis of information occurs by following chains of evidence, evaluating formal hypotheses [27], testing competing explanations [86], or telling stories [37] using visual metaphors to convey relationships and dynamics. These activities are particularly challenging in intelligence analysis, emergency management, epidemiology, and other critical areas that involve high-dimensional abstract information [83] and large geospatial datastores [36]. However, the heterogeneous and idiosyncratic nature of the data sets and analysis activities in these endeavors are similar to those in everyday domains, making it likely that the outcomes of visual analytics research will translate readily into visualization approaches that will help to engage broad audiences.

2.2 Skills

Novice Users: By novice users we mean users who have experience operating a computer, but no experience with programming in general, let alone programming visualization techniques. The vast majority of novice visualization users act as consumers: they will interact with the visualization within the possibilities offered but will rarely extend existing functionality to suit their analysis needs. If we want these users to be able to produce visualizations, we have to take care to make this process as easy as possible. Some points of consideration when designing visualizations for novice users are:

Data Input: We cannot expect a novice user to write their own data parser, write database queries that export data to a particular format or understand the file formats for more complex data types. Most novice users seem to take to using spreadsheet programs such as Microsoft Excel to store and analyze their data. One useful input format then, is a simple tab delimited input file, as this format is both human readable and can be directly copied from the spreadsheet editor.

Automatic Selection of Visualization Type: Novice users have no experience designing visual mappings and may even choose mappings that produce nonsensical visualizations. Recurring examples include the use of line charts over categorical data dimensions, for which a bar chart would be a better choice, and using a pie chart for data that do not form part of a whole. For this reason, visualization techniques geared towards novice users should at least partly automate the selection of visual metaphors. This may involve analyzing the data dimensions to see if there are any ordinal attributes, check for aggregated variables and totals, and examine values in dimensions for possible hierarchical structure [59,60].

Useful Defaults: Novice users likely will not spend time tuning an ugly looking visualization to fit their needs. It is therefore important to provide a set of sensible defaults for data and view parameters (such as scales, colors, item sizes and viewpoints) to help constrain the parameter space that users have to explore. Multiple combinations of these parameters can be offered by providing a preset list. As an added bonus, a good set of presets can show users what is possible and educate them on what is sensible.

Contextual Information: With contextual information we mean visual items that explain to the user what data is being mapped to the screen and what encodings are being applied. This involves legends, scales, labels, pop-ups, titles and explanations of visual mappings. Although visual graphics in print media take great care to provide contextual information, interactive visualizations are often lacking in this respect because most of the design attention is focused on the visual mapping itself.

Savvy Users: By savvy users we mean people who have experience performing relatively sophisticated data organization and manipulation, using a combination of manual processing and limited amounts of programming or scripting. Because savvy users are a small but non-trivial part of the population of visualization consumers, they are a critical bridge between experts and novices. As such, savvy visualization users may act variously as:

– *experts* who train or guide novice users in the use of particular visualizations by clarifying exploratory and analytic functionality in terms of interface appearance and behavior,
– *designers* who plan, construct, debug, test, and deploy new visualizations for ongoing evaluation and routine operation by novice users,
– *end-users* who can bring more extensive experience to bear when using existing visualizations to analyze data from their own knowledge domains, to browse data with which they are less familiar, and to share their results with others, and
– *explorers* (or *user-designers*) who combine the roles of designer and end-user by extending and redesigning visualizations on the fly during open-ended exploration of their data.

Expert Users: By expert users we mean people who have extensive experience with interactive graphical software development and the theory and application of data modeling, data processing, and visual data representation. As such, visualization experts may act both as:

– *researchers* who invent, specify, and evaluate methods for accessing, querying, rendering, and interacting with data, often with an eye toward extending and enhancing the functionality of existing visualization systems and tools, and

– *developers* who design and implement visualization modules, toolkits, systems, and tools of various sizes and scopes, often adapting and integrating existing functionality from other visualization toolkits and systems.

In particular, visualization research frequently involves the development of prototypes for evaluating the correctness, flexibility, and performance of new data processing algorithms and the usability and utility of new interaction techniques.

Facilitating the interdependent needs of novice, savvy, and expert users is a key part of supporting broader audiences for information visualization. The number of people who can act as visualization designers or visualization developers – let alone the core visualization researchers who by necessity often fill these roles – is rapidly becoming overwhelmed by demand for visual tools brought on by blossoming public awareness of the power and accessibility of information visualization techniques. It will become increasingly necessary to provide users of all skill levels, including novices, with the capability to explore and analyze data sets of personal and professional interest without direct assistance from traditional visualization practitioners. However, understanding how to design accessible yet flexible software artifacts for individual visual exploration and analysis is only half of this equation. Social organization of visualization roles through collaboration and other means, as described later in this paper, is the critical second half.

2.3 Goals

One of the traditional rationales for information visualization is that the human visual system has high input bandwidth and has evolved as an excellent tool for spotting patterns and outliers in our surroundings. If we then map large amounts of data into visual form, we can use these innate human abilities to explore the data to find patterns that would have been exceedingly difficult to identify through purely automated techniques. A current prominent example is bioinformatics research that visually explores gigabytes of gene experiments to investigate the mechanisms that drive a particular disease. Such "explorative" use-cases have dominated most of the research in visualization over the past two decades. Explorative use can either be open-ended, where the user wants to browse their data without having a predefined question in mind, or analytically driven, in which the user has a particular question in mind and uses the visualization to answer it. Often times these two types of exploration will be intertwined: a user will explore a previously unknown data set without a particular question in mind, stumble on an interesting data point and then use the analytic features in the visualization to either answer the question or redirect their open-ended exploration.

Exploration and Analysis: Recent visualization environments have begun to offer users various degrees of interactive control over different parts of the entire information interface design process, thereby opening up possibilities for much deeper exploration of data. Such environments allow computer-savvy user-designers to interactively access data, create, layout, and coordinate views, and connect data to views. Design typically occurs directly within the interface that contains data views, and often take effect immediately without the need for a separate compilation or build stage. This live, amodal approach to interface design allows users to switch rapidly between building and browsing tasks during exploration and analysis. The result is a form of exploration that is free form and open-ended, particularly during initial inspection of newly encountered data sets.

IVEE [2], DEVise [58], DataSplash [71], Snap-Together Visualization [69], GeoVISTA Studio [88], Improvise [102], and Tableau/Show Me [60] are a few of many well-known visualization environments that support open-ended data exploration to various degrees. Such environments typically consist of a graphic user interface on top of a library of visualization components which may or may not be exposed as a visualization programming toolkit in its own right. This combination of user interface and underlying library can enable open-ended exploration in a very broad sense if it bridges the activities of visualization users performing various roles with different levels of expertise, whether as individuals or in collaborative groups.

To connect developers and designers, a key advantage for open-ended exploration is an extensible library that provides an application programming interface (API) for adding new software modules for various visualization components (including data access, queries and other data transformation algorithms, views, and visual data encodings). In particular, the most useful APIs support the definition of new data transformation operators—including appropriate input and output data object types—that give designers the ability to express rich relationships between data, queries, and views. This requirement is essential for applying newly discovered visualization techniques to emerging sources and forms of information, without needing to constantly architect and implement new toolkits (and retrain visualization designers in their use).

To connect designers with users, the user interface must support the ability to access data sets (and metadata) from local or remote sources in various formats, create and position views on the screen, specify how navigation and selection affects views, specify queries on data, parameterize queries in terms of interaction, and attach data sets and queries to views. In particular, designers should be able to specify the appearance and behavior of their visualizations directly within the user interface, without resorting to programming or other workarounds for interface limitations. To do otherwise would effectively require that designers be trained as developers.

User interfaces that truly support open-ended exploration would exceed the requirements of basic visualization design and operation by: supporting live building of complete browser interfaces, including immediate designing, debugging, and testing of intended functionality; facilitating collaboration between end-users and designers to turn analytical questions into structural changes (through remote, nearby, or side-by-side efforts to communicate and effect rapid visualization prototyping and polishing); and enabling rapid switching between building and browsing to perform more extensive exploratory visualization by modifying visualization views and queries on the fly. In particular, it is highly desirable for explorers to be able to see all raw data quickly to make decisions about how to visualize it, rapidly create and lay out views, rapidly attach data and queries to views, rapidly modify queries, store, copy, and reuse views, copy-and-paste/drag-and-drop visualization components, and use macros to build common multiple view constructions. Many of these capabilities are also desirable for non-exploring designers who prepare visualizations for domain analysts.

In all of this, availability of common and familiar interface functionality is essential to broad adoption. The user interface should run in the user's normal working environment, require no programming or design activities, and provide a way to disseminate analytical results. For communication and collaboration, it is highly desirable for the user interface to run easily on any platform, allow visualizations to be opened and saved as normal documents for sharing between users, and provide the ability to bookmark or screen capture visualizations in different graphical states.

Communication: At the other end of the spectrum of information visualization goals is the "communicative" use-case, where the main user goal is simply to convey a message to others. This use-case is already present in many traditional media: think of diagrams explaining the numbers behind a news story or a bar chart that has been included in a slideshow presentation. Although these particular representations are fairly static because of the affordances of the media used, this does not mean that communicative information visualization is limited to static visualizations. Interactivity is a very useful means of engaging users and may make them more receptive to a particular message. However, interactivity also poses some problems when communicating visualizations, because it's hard to reproduce interactive features in a static medium. Many information visualizations use tooltips and mouse-overs to provide contextual information, offer the user different viewpoints on the data, and allow for dynamic analysis of data. Videos alleviate this problem only in part, because it is often hard to follow what is going on and much of the context in the exploration process is missing. Simply sharing findings using static representations of interactive visualization is therefore not the optimal solution, and we would do better to consider these issues beforehand when designing visualizations for communicative use.

Apart from traditional mass communication, the communicative use-case also plays a pivotal role in collaborative applications, especially ones that are non-collocated and/or asynchronous. If the analysts do not share the same time or space it is important for them to be able to communicate findings and bring each other up to speed on the current state of the process quickly. Furthermore, each of these types of collaboration (collocated or distributed) has its own type of requirements, which we discuss in-depth in sections 3 and 4. In general, communicative use of information visualization usually involves a small investment of time on the user end, with a small but guaranteed payoff. On the other hand, explorative use involves a large amount of investment in tools, training and time, while the (potentially high) payoff is not always guaranteed. (See also [93] for a discussion of these tradeoffs.)

In the next section, we consider a number of representative tools that help us meet these differing goals of communication and exploration.

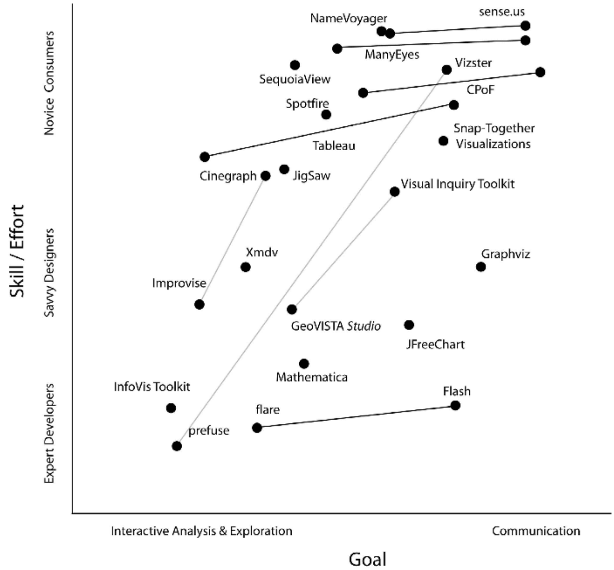


Fig. 1. A few of the many available information visualization tools, roughly mapped according to targeted end-user and targeted goal. Light lines connect toolkits and development environments to examples of visualizations created in them. Dark lines roughly capture similar ranges of user/goal targets for relevant tools.

2.4 Tools

Note that most real-world uses of information visualization will form a combination of the use-cases and roles described in the preceding sections. A researcher might program a new visualization technique to explore his complex data and then present findings to a manager by sending a screenshot. In this case the researcher takes on the roles of both consumer and developer and performs both exploration and communication. Most current information visualization tools and toolkits are geared towards one particular user skill and goal, although a recent trend towards more flexible tools can be observed. To illustrate the rough classification outlined in the previous subsections, in this section we give an indicative sample of an end-user visualization tool for each user skill and goal combination. Figure Fig. 1 illustrates a number of available visualization tools categorized according to the skill level of the target user base and the degree to which the tools support analytic and communicative tasks. Systems that span a range of tasks or skills are presented as line segments indicating the range of users and usage.

Expert Communication: In the bottom right corner of the matrix we find information visualization tools and toolkits that are geared towards communicative use, but assume a significant amount of knowledge on suitable visual techniques and their implementation. One such toolkit is Adobe's Flash development environment. Flash is an browser based interactive graphical design tool. Because of its ease of online deployment, it is particularly suited to communicating messages in a graphical way. In fact, many of the interactive graphics on the Internet today are Flash-based. However, Flash does not offer the developer tools that would support structured data exploration. (In fact, it offers only the most basic of data structures.) Moreover, its timeline-based design environment is not particularly well-suited to interactive visualization development.

This situation has improved with the release of Actionscript 3 and Flex, offering a more advanced programming model and a full-fledged user interface development package. The Flare toolkit [41] implements basic visualization capabilities within Flash, making it easier to develop interactive information visualizations, while still retaining the benefits of Flash, such as its relatively lightweight means of online deployment.

Savvy Communication: Information visualization tools in this category (middle right) allow users to create and share complex information visualizations, but require a base level understanding of computer programming and information visualization. A concrete example of such a tool is AT&T's GraphViz [33] library, which allows users to generate static images of graphs but requires some programming effort to integrate it with existing applications because it uses a proprietary data format.

Use-cases for GraphViz often involve reporting engines that need to be able to display networked data of some sort. Many of the features in GraphViz are geared towards presentation instead of exploration. For example, it is possible to heavily customize node rendering. Special care has been taken to avoid label overlapping, as this would make static images completely unreadable. Both of these issues are less of a problem in interactive systems in which users can use tooltips to get more information, or zoom into a dense cloud of labels to remove overlap.

Novice Communication: Until recently, if novice users wanted to share information visualizations with others they would be limited to taking screenshots of information graphics for sending by e-mail, etc. This mode is often sufficient if the goal is one-way information dissemination. For example, a pie chart may be included in a presentation, or an advanced information graphic may be printed in a newspaper. However, this mode of publication fails if users want to collaboratively analyze a complex data set.

Recent tools like Many Eyes [97], Sense.us [46], Swivel.com, and Spotfire Decision Site Posters make this process much easier, allowing easy sharing of interactive visualizations. As discussed in greater detail in section 4, users of these systems can share a particular state of a visualization encoded as a URL and add custom annotations and comments while still having access to the interactive features of the visualization. This makes it possible to quickly switch between analysis and communication, a necessity for successful collaboration.

Expert Exploration and Analysis: The bottom left corner of the matrix contains visualization software that supports deep and broad exploration of the space of visualization techniques, as well as more focused exploration and analysis of particular data sets. Such flexibility in the overall process of visualization almost always requires substantial expertise, typically requiring programming skills. As a result, visualization software for exploration by experts often takes the form of toolkits that are written in a popular programming language but that encapsulate well-known visualization components and techniques.

One such toolkit, prefuse [44], provides a Java-based library of visualization building blocks including views, visual encodings, processing algorithms, multi-view coordinations, and a common data model that supports tables, trees, and graphs. Graphs, hyperbolic trees, treemaps, and scatter plots support accessing, filtering, rendering, and displaying data using a variety of layout and distortion algorithms.

Similarly, the InfoVis Toolkit [34] is a set of Java visualization components designed around OpenGL and a data model that represents tables, trees, graphs, and metadata in column format for efficient selection, filtering, visual encoding, and coordination. Views include scatter plots, parallel coordinate plots, treemaps, and a variety of node-edge tree and graph displays that can incorporate fisheye lenses and dynamic labeling of items. Visualizations created in the toolkit display textboxes, sliders, and other controls alongside views for dynamic editing of visual encodings.

The extensible programming interfaces of both toolkits and those like them provide a means to incorporate new components and techniques, in essence expanding the scope of exploration, considered broadly, to include the results of future visualization research.

Savvy Exploration and Analysis: As described in the previous section, visualization in the expert exploration category revolves more around programming rather than around interaction in integrated user interfaces intended for designing and building tools. Research on integrated visualization environments focuses on packaging the exploratory capabilities of toolkits in ways that are accessible to users who are visualization savvy but not necessarily visualization experts.

For instance, Improvise [102] is a self-contained Java application that appears and behaves like other office productivity applications based on the multiple document desktop metaphor. Users build Improvise visualizations by interactively constructing the data, queries, views, and coordinations of tools that can be saved, opened, copied, and shared as self-contained Extensible Markup Language (XML) documents. Users browse visualizations using the mouse and keyboard to navigate and select data items in multiple coordinated views.

Similarly, GeoVISTA Studio [88] is an integrated visualization development environment for building geovisualizations interactively using a graph-based visual coordination editor. Any component that conforms to the JavaBeans specification can be a view. Development of new views by the community of GeoVISTA Studio users has resulting in a large library of views utilized in numerous visualizations. A particular strength of GeoVISTA Studio is its extensive functionality for representing and displaying geospatial information (based on the GeoTools [57] open source Java GIS toolkit).

The combination of browsing with rapid, iterative building in a single application (much like in spreadsheet programs) enables improvisational visualization, in which it is possible to design and evaluate different ways of analyzing particular data sets in the form of rapid prototypes having more concrete and stable collections of analytic functionality.

Novice Exploration and Analysis: As far as we know, there are no tools that truly allow novice users to interact with their data in the broadest sense of exploration. This may result from an apparent fundamental tradeoff between flexibility and accessibility in visual analysis, in that increased expressiveness necessitates greater expertise when it comes to data manipulation and visual representation. Even in savvy exploration-analysis tools like Improvise that strive for a balance between these factors, reproducing many common visual components and techniques currently requires a high degree of visual language expressiveness that necessitates a corresponding high level of expertise beyond that of most novice users. Conversely, novice analysis-communication tools like Many Eyes seek to increase visualization flexibility for broad audiences keenly interested in modest analytic expressiveness as a means to better communicate ideas about information. In between, analysis tools like Tableau/Show Me successfully occupy analysis niches that provide bounded but particularly useful forms of data interaction to relatively broad audiences who are sufficiently motivated to devote time and effort to modest training. It may well be that open-ended exploration tools for novices will evolve from future research into ways of combining these three seemingly complementary directions.

2.5 Directions

Current end user visualization tools are becoming more and more flexible in the types of scenarios and goals they can handle. Tools like Many Eyes allow novice users to create advanced visualizations with very little effort and also support communicative use-cases by allowing flexible sharing of visualization states. Tools like Improvise allow tight integration of many different types of visualizations, but require some programming skills on the side of the end-user, an expectation that is not always reasonable of domain experts dealing with the visualization. Tableau allows end users to set up and pivot different types of basic visualizations in a fairly intuitive manner and the recent addition of Tableau Server allows sharing of and commenting on these visualizations in an online environment, making it also suitable for communicative purposes. Although flexible, the only visualization types allowed are 2-dimensional small-multiple displays, which limits the visualization and analysis types to basic business graphics.

In our opinion, the ultimate goal of letting novice users flexibly specify their visualization needs and couple different types of views together has not been fully realized yet. We expect that users' visual literacy will increase as information visualization becomes more mainstream, and will start demanding advanced visualizations beyond the trusted bar chart. Integrating advanced visualizations in an flexible, collaborative and easy to understand framework for open-ended exploration and analysis is an important and solvable problem. We expect this solution will have important implications for many areas of human endeavor that necessitate the handling of complex data.

3 Co-located Collaborative Visualization

Given the choice, it is common and natural for people to work together. This is not a new phenomenon. Small groups of people gather for all kinds of reasons including many that are work related; such as to get a job done faster, to share expertise for a complex task, and to benefit from different insights from different people. Also, when one considers the rapid growth in size and complexity of datasets, it is not surprising that increasingly the practicality of an individual analyzing an entire data-set is becoming unrealistic. Instead, the expertise to analyze and make informed decisions about these information-rich datasets is often best accomplished by a team [91]. For instance, imagine a team of medical practitioners examining a patient's medical record to plan an operation, a team of biologists looking at test results to find causes for a disease, or a team of businessmen planning next year's budgets based on a large financial dataset. All of these situations involve a group of people making use of visual information to proceed with their work. Research towards supporting these team-based information processes will expand the situations in which information visualization can be used and is part of considering how to best support people in their normal everyday information work practices.

This section draws from a wide variety of literature to shed light on questions and issues that need to be considered during the development of co-located collaborative information visualizations. We do not consider this discussion to be exhaustive; rather it is our intention that the discussion will form the beginning of design guidelines and considerations that will be modified and extended through future research in collaborative information visualization.

Research in information visualization draws from the intellectual history of several traditions, including computer graphics, human-computer interaction, cognitive psychology, semiotics, graphic design, statistical graphics, cartography, and art [64]. The synthesis of relevant ideas from these fields is critical for the design and evaluation of information visualization in general and it is only sensible to think that fields concerned with collaborative work also add valuable information to our understand-ing of requirements for collaborative information visualization systems. Our sources include work in co-located collaboration in computer supported cooperative work [39,53,75,73,76,77,80,81,82,90], information visualization [85,105,109,110,111], and empirical work investigating collaborative visualization use [61,68,72].

The organization of this section is as follows. A brief overview of existing research that relates to co-located collaborative information analysis is given in section 3.1. Next, section 3.2 discusses the impact of recent advances in hardware configurations and section 3.3 focuses on more general human computer interaction issues important for the support of the co-located collaborative process, primarily drawing upon computer supported co-located collaborative literature. Then section 3.4 presents information visualization specific issues that may need re-consideration in light of co-located collaborative applications.

3.1 Related Research

Co-located collaborative information visualization is a relatively new and still under explored research area. Only a few tools designed specifically to support synchronous collaboration between co-located people using visualizations to explore information have emerged thus far. These are discussed below. However, as noted below, existing visualization tools designed from a single-user perspective have been studied with co-located collaborative tasks [61]. There has been considerable research in the area of scientific visualization in distributed systems (see [38] for an overview). Recently, there has been new primarily web-based research on asynchronous distributed collaborative information visualization systems. This new direction is the focus of section 4.

Co-located Collaborative Visualization: The Responsive Workbench was one of the first visualization systems for developed co-located collaboration around a large horizontal surface [103]. The responsive workbench is a virtual reality environment in which the displayed 3D scene is seen through shuttered glasses and interaction is achieved with a glove which has an attached Polhemus sensor on the back. Agrawala et al. [1] extended this workbench to support two simultaneous users. Several scientific visualization applications were developed for this platform including fluid dynamics and situational awareness applications.

On tabletop displays information visualization interaction techniques have been used to support co-located people in information sharing and exploration tasks. DTLens [35] provides a local non-linear magnification technique enabling multiple lenses for up to four people with for two-handed interaction. Personal Digital Historian uses radial layouts to display photos, video and text documents to supports conversation and story telling for small groups of people [84].

Studying Collaborative Use of Information Visualizations: While research on collaborative data analysis using information visualizations is relatively scarce, collaborative use of existing single user systems has been studied. Mark and Kobsa [61] conducted a user study in which they observed pairs working in co-located and distributed settings with two different visualization systems designed for single users. Their findings suggest that the benefit of collaborative vs. individual problem solving was heavily dependent on the visualization system used and also that, in general, groups were better than individuals working alone at locating errors. From this study, they derive a model for the collaborative problem-solving process. Their model consists of an iterative sequence of five stages: parsing a question, mapping variables to the program, finding the correct visualization, and two validation stages. From studying collaborative work on scientific visualizations in virtual environments using CAVEs, Park et al. [72] report a five-step activity model that was common for the observed collaboration sessions. Their study also noted that participants showed a strong tendency for independent work, if the option was available. Isenberg et al. studied co-located collaborative data analysis scenarios and posit an eight-process framework that relates to previous work on the Sensemaking Cycle [17] and the two studies by Mark and Kobsa [61] and Park et al. [72]. However, a common temporal order of analysis processes as posited by some previous work did not emerge.

3.2 Choosing Hardware to Support Co-located Collaboration

We start with a discussion of hardware because some of the recent interest in co-located collaboration is at least in part due to new hardware innovations.

Display Size: In information visualization, the size of the available display space has always been problematic for the representation of large datasets (e. g., [65]). In a common desktop environment, typically a single user will use all available screen space to display their visualization and, most commonly, this space will not be sufficient. Frequently, visualization software will include interactive features to help the user cope with limited display space. It seems sensible to think that, if we are going to adequately support collaborative or team exploration of visualizations, available display space will be an important issue. In collaborative systems, screen space not only has to be large enough for the required information display, it might also have to be viewed and shared by several users. As the number of people using a shared information display grows, the size of the display and workspace needs to be increased in order to provide a viewing and interaction area that gives adequate access to all group members.

Display Configuration: Several configuration possibilities exist that could increase the amount of available display space, all of which will affect the type of visualization systems possible and the type of collaboration work that would be most readily supported. Many types of configurations are possible; for instance, one could provide team members with interconnected individual displays, as in the ConnecTable system [89], or one could make use of large, interactive, single-display technology, like display walls or interactive tabletop displays (e. g., [90]). An additional possibility is to link wall, table, and personal displays (e. g., [105]), or to consider immersive displays (e. g., [72]). The type of setup most appropriate for an information visualization system will depend on the specific task and group setup. For example, individual interconnected displays allow for private views of at least parts of the data which might be required if data access is restricted. Tabletop displays have been found to encourage group members to work together in more cohesive ways, whereas wall displays are beneficial if information has to be discussed with a larger group of people [76].

Input: In the common desktop setup, input is provided for one person through one keyboard and one mouse. To support collaboration, ideally, each person would have at least one means of input. In addition, it would be helpful if this input was identifiable, making it possible to personalize system responses. If a collaborative system supports multi-user input, the access to a shared visualization and data set has to be coordinated. Also, synchronous interactions on a single representation may require the design and implementation of new types of multi-focus visualizations. Ryall et al. [77] have examined the problem of personalization of parameter changes for widget design, allowing widgets to be dynamically adapted for individuals within a group. Similar ideas could be implemented for personalization of information visualizations during collaborative work.

Resolution: Resolution is an issue both for the output (the display) and for the input. The display resolution has a great influence on the legibility of information visualizations. Large display technology currently often suffers from relatively low display resolution so that visualizations might have to be re-designed so that readability of text, color, and size not affected by display resolution. Also, large interactive displays are often operated using fingers or pens which have a rather low input resolution. Since information visualizations often display large data sets with many relatively small items, the question of how to select these small items using low input resolution techniques becomes an additional challenge that needs special attention [48].

3.3 Creating a Collaborative Environment

The key characteristics of co-located synchronous interactions as described by Olson and Olson [70] will apply to information visualization scenarios designed to support co-located collaboration. These characteristics include: a shared local context in which participants can interact with work objects, rapid feedback, and multiple channel information exchange (voice, gesture, etc.), and visibility of others' actions. These characteristics are further specified in the mechanics of collaboration [73], which describe basic operations of teamwork, or small scale actions and interactions that help people solve a task as a team. These mechanics apply to a variety of group and task settings. This section is discussed under the two major groupings of the mechanics of collaboration—communication and coordination—and under those issues relating to supporting varying collaboration styles.

Communication: Communication is an important part of successful collaborations. People need to be able to trigger conversations, communicate their intentions, indicate a need to share a visualization, and to be generally aware of their team members' actions. Group members need to be informed that some parameter of a shared display might have changed while they were busy working with an information visualization in a different part of the workspace. There has been considerable discourse on the importance of all team members being aware of all other team members' actions. Pinelle et al. [73] make a distinction between explicit and implicit communications. The design of support techniques for both types of communication needs to respect common social and work protocols [70]. For example, the interface should not require a group member to reach into or across another person's workspace in order to acquire or share visualizations or controls.

Explicit Communication: Enabling direct exchange of information through many channels such as voice, gestures, and deictic references facilitates collaborative work in general [73] as well as co-located collaboration [70]. It has been shown that the ability to annotate data and share insights in a written way is an essential part of the discovery process in distributed information visualization settings [46]. This collaborative need for annotation exists in traditional use of pen and paper based information as was observed in a study of teams working on information analysis tasks in a shared setting [68]. However, in digital systems messages of all types, written, voice, etc. might not always be as easily shared and how best to support this will require further research.

Implicit Communication: In co-located non-digital collaboration people are accustomed to gathering implicit information about team members' activities through such things as body language, alouds, and other consequential communications. This is an active research area in distributed collaboration research since the co-located evidence does not naturally become distributed. Co-located collaboration benefits from many of the co-present advantages, however, issues still arise. Some examples include: digital actions are not always readily visible (cursors are hard to see on large screens), menu actions can affect a remote part of the screen, as well as the general problems of change awareness [74]. Thus while implicit communications do support awareness in a co-located setting already to some extent, some system changes made by a collaborator can still remain unnoticed if the collaborative system does not provide appropriate feedthrough (i. e. a reflection of one person's actions on another person's view). In collaborative information visualization, for example, it might be important to consider appropriate awareness for operations that make changes to the underlying dataset.

Imagine a co-located system in which each collaborator works in parallel on a different view using a different file-system representation. If one collaborator discovers an old version of a file and decides to delete it (a value operation [23]), this change might go unnoticed if the other person is looking at a view of the data that does not include the current file or it might be completely surprising to the other person to see a file in their representation disappear. Some research has proposed policies to restrict certain members from making unsuspected global changes to a dataset [75]; however, while earlier research on information visualization discussed the differences between view and value operators (e. g. [23]), most recent research in multiple-view visualization tends to favour view operations (filtering of unwanted data rather than deletion). This seems likely to be most appropriate during collaboration.

It has also been shown the location and orientation of artifacts is used to support implicit communication in non-digital settings providing information on such things as who is working with which artifacts and when one person wants to initiate communication about a particular artifact [53] and that this translates to digital settings [54]. This consideration, providing for artifact mobility and freedom orientation, will probably also be important in supporting information visualization collaboration.

Coordination: In group settings, collaborators have to coordinate their actions with each other. Here, we describe several guidelines for how to support the coordination of activities in collaborative information visualization applications.

Workspace Organization: Typical single-user information visualization systems impose a fixed layout of windows and controls in the workspace. Previous research has shown that, on shared workspaces, collaborators tend to divide their work areas into personal, group, and storage territories [81]. This finding implies that a group interaction and viewing space is needed for collaborative data analysis where the group works on a shared representation of the data or in which they can share tools and representations. Also, the possibility of exploring the data separately from others, in a personal space, is necessary. Flexible workspace organization can offer the benefit of easy sharing, gathering, and passing of representations to other collaborators. By sharing data in the workspace, representations will be viewed by team members with possibly different skill sets and experiences and, therefore, subjected to different interpretations. Also, by being able to move and rotate representations in the workspace, an individual can gain a new view of the data and maybe discover previously overlooked aspects of the data display.

Collaborative information visualization systems should allow for social interaction around data displays [46]. If visualizations can be easily shared, team members with different skill sets can share their opinions about data views, suggest different interpretations, or show different venues for discovery. By offering mechanisms to easily rotate and move objects, comprehension, communication, and coordination can be further supported [54]. Rotation can support comprehension of a visualization by providing alternative perspectives that can ease reading and task completion, coordination by establishing ownership and categorizations, and communication by signaling a request for a closer collaboration [53]. By allowing free repositioning, re-orientation we can also make use of humans' spatial cognition and spatial memory and possibly better support information selection, extraction, and retrieval tasks [68]. Mechanisms for transfer and access to information visualization in the workspace should be designed in a way that they respect common social work protocols [53,81].

Changing Collaboration Styles: Tang et al. [90] describe how collaborators tend to frequently switch between different types of loosely and closely coupled work styles when working over a single, large, spatially-fixed information display (e. g., maps or network graphs). A study by Park et al. [72] in distributed CAVE environments discovered that, if the visualization system supports an individual work style, users preferred to work individually on at least parts of the problem. For information visualization systems, an individual work style can be supported by providing access to several copies of one representation. The availability of unlimited copies of one type of representation of data allows group members to work in parallel. More closely coupled or joint work on a single view of the data can be supported by implementing the possibility of concurrent access and interaction with the parameters of an information visualization. Free arrangements of representations also support changing work styles. Representations can be fluidly dragged into personal work areas for individual or parallel work and into a group space for closer collaboration.

3.4 Designing Information Visualizations for Co-located Collaboration

Many known information visualization guidelines still apply to the design of information visualizations for co-located collaborative use (e. g., [10,92,99]). In this section, we discuss changes and additions to aspects that need to be considered when designing information visualizations for co-located collaborative settings. Thus, much of this discussion simply delineates research questions that may of specific interest when designing information visualizations to support co-located collaboration.

Representation Issues: Spence [87] defines representation as "the manner in which data is encoded," simplifying Marr's [62] definition of representation as a formal system or mapping by which data can be specified. The concept of representation is core to information visualization since changes in representations cause changes in which types of tasks are most readily supported. As in Marr's [62] example, the concept of thirty-four can be represented in many ways. To look at three of them; Arabic numerals, 34, ease tasks related to powers of ten; Roman numerals, XXXIV, simplify addition and subtraction; and a binary representation, 100010, simplifies tasks related to powers of two. Not surprisingly, Zhang and Norman [110] found that providing different representations of the same information to individuals provides different task efficiencies, task complexities, and changes decision-making strategies. Questions arise as to what are the most effective representations during collaboration. Will certain representations be better suited to support small group discussions and decision making? Will multiple representations be more important to support different people's interpretation processes? Will new encodings or representations be needed for collaborative work scenarios? Appropriate representations might have to be chosen and adapted depending on the display type chosen but whether completely new designs are required is not yet clear.

For example, different representations may have to be accessible in an interface because in a collaborative situation, group members might have different preferences or conventions that favour different types of representations. Gutwin and Greenberg [39] have discussed how different representations of the workspace affect group work in a distributed setting. They point out that providing multiple representations can aid the individual but can restrict how the group can communicate about the objects in the workspace. This extends to co-located settings, in which several representations of a dataset can be personalized according to taste or convention, making it harder to relate individual data items in one representation to a specific data item in another. For example, relating one specific node in a treemap [50] to another node in a node-link diagram might require a search to locate the respective node in the other representation. Implementing mechanisms to highlight individual data items across representations might aid individuals when switching between group and parallel data exploration.

Findings suggest that the availability of multiple, interactively accessible representations might be important for information visualization applications since the availability of multiple data representation can change decision making strategies [52]. Also differing representations have an influence on validation processes in information analysis [79], and more easily support people working in parallel on information tasks [72]. While this is probably applicable, empirical evidence directly linking these finding to collaborative information visualization has not yet been gathered.

It is also possible that the actual mappings used in representations may have to be re-thought. For example, spatiality or the use of position/location is commonly an important aspect of representation semantics. However, spatiality as manifested in territoriality is a significant factor for communication and coordination of small group collaboration. It is an open question as to whether there is a trade-off between these two uses of spatiality.

Presentation Issues: Presentation has been defined as 'something set forth for the attention of the mind' [63] and as 'the way in which suitably encoded data is laid out within available display space and time' [87]. From these definition is clear that changing display configurations, as is usually the case to support co-located collaboration, will impact the types of presentations techniques that are possible and/or appropriate. Common presentation techniques include pan & zoom, focus & context, overview & detail, filtering, scrolling, clutter reduction, etc.

A common theme in information visualization is the development of presentation techniques that overcome the problem of limited display space (e. g. [4,20,49]). In collaborative scenarios, information visualizations might have to cover larger areas than in a single user scenario as group members might prefer to work in a socially acceptable distance from each other. The display space might also have to be big enough to display several copies of one representation if team members want to work in parallel.

If groups are working over a shared presentation of data, presentations might have to be adapted to allow collaborators to drill down and explore different parts of the data in parallel. Collaborative information visualizations will likely have to sup-port multiple simultaneous state changes. This poses additional problems of information context. Team members might want to explore different parts of a dataset and place different foci if the dataset is large and parts of the display have to be filtered out. Information presentations might have to be changed to allow for multi-focus exploration that does not interfere with the needs of more than one collaborator. For example, DOI Trees [18] or hyperbolic trees [55] are examples of tree visualizations in which only one focus on the visualization is currently possible. ArcTrees [67] and TreeJuxtaposer [65], for example, allow for multi foci over one tree display but these were not designed to take the information needs of multiple collaborators into account and might still occlude valuable information.

An example for visualization presentation changes based on a collaborative circular tabletop environment has been presented in [94]. The presentation of the circular node-link tree layout was modified to rotate all nodes towards the boundary and a "magnet" was implemented to rotate nodes towards just one team member. Nodes were also changed in size; as leaf nodes were placed closer towards team members, in their personal space [81], they were decreased in size and the nodes towards the center of the table were enlarged to allow for easier shared analysis of the node contents in the group space [81]. A possible extension of this work is to think about placing and re-arranging nodes automatically based on the placement and discovery interests of team members or based on the individual or shared discoveries that have been made.

The presentation of visualizations might also have to take available input devices on a shared large display into account. If fingers or pens are used as an input device, the selection might not be accurate enough to select small information items. A common task in information visualization is to re-arrange data items (e. g. by placing points of interest), to request meta-information [85] (e. g. by selecting an item), or to change display parameters by selecting an item. If the displayed dataset is large, it often covers the full screen and reduces individual items to a few pixels. Previous research has attempted to solve the issue of precise input for multi-touch screens (e. g. [8]) but they might not be applicable if the whole visual display is covered with items that can possibly be selected. Alternatively, information presentations could be changed to allow for easier re-arrangement and selection of items, for example, with lenses [20]. DTLens [35] presents an initial exploration of the use of lenses in co-located collaboration.

The resolution of a large display has an influence on the legibility of data items. It is known that the reading of certain visual variables is dependent on the size and resolution in which they are displayed [99]. Information visualizations also often rely on textual labels to identify data items which may be hard to read on low-resolution displays. The presentation size of individual items and labels may have to be adapted to compensate for display resolution.

View Issues: The term view is common in information visualization literature and view operations (changing what one currently sees) have been defined as distinct from value operations (changing the underlying data) [23], however, this use of the term view also incorporated changes in visual aspects of representation, presentation. Blurring the distinction between view and presentation changes has not been problematic because with a single viewer and a single display these are often concurrent. A change in view can be simply looking at exactly the same presentation and representation of the same data merely from a different angle or it can include changes in all three factors.

In a co-located collaborative setting, of necessity there are as many views of a given presentation as there are people in the group. Also since collaboration practices often include mobility, a given person's view will change as they move in the physical setup. This factor has recently begun to receive attention in the CSCW community. Nacenta et al. [66] have shown that righting (orient-ing a piece of 2D information into the proper perspective, by means of motion tracking, really aids comprehension. Hancock and Carpendale [40] consider the same problem for horizontal displays looking for non-intrusive interactive solutions. Since a study by Wigdor et al.[105] has indicated that angle of viewing affects readability of certain visual variables; this issue will be an important one for collaborative information visualizations. This research on how view-angle distortion affects perception in a single and multi-display environment suggests that certain types of representations may need to be modified in order to be used on a digital tabletop display and that information visualizations should not be compared across multiple display orientations. However, as visual variables were tested in isolation (e. g. length, direction, only) further evaluations have to be conducted to see whether participants will correct for possible distortion if the variables are presented in conjunction with others or whether view correction [66,40] might compensate.

Visualizations that can be read from multiple angles and orientations (e. g. circular tree layouts vs. top-down layouts) might be more appropriate for display on a horizontal surface. However, it is not clear whether participants would try to read oriented visualizations upside down and make wrong conclusions based on these readings or whether they would simply re-orient the visualization to correct the lay-out. Observations of collaborative information visualization scenarios point in the latter direction [68].

Gutwin and Greenberg [39] discuss issues about viewing a representation in relation to distributed scenarios. However, parallels can be drawn to co-located scenarios in which collaborators work with multiple linked copies of the same representation of a dataset. These essentially represent multiple movable “split viewports” [39]. The suggested solutions for distributed settings include radar views, overview+detail solutions, and cursor eye-view. Whether the benefit of these solutions in a co-located setting outweighs the possible distractions they might create, however, would have to be evaluated. Further guidelines for using multiple views in information visualization can be found in [5] and provide a starting point for tailoring multiple views for collaborative visualization.

Issues of view may develop in collaborative information visualization settings if collaborators want to switch from loosely-coupled to closely-coupled workstyles [90] and share discoveries they have made with the other group members. If one collaborator worked with different view of the dataset it might be difficult to locate the information in the other persons’ view. Another important factor to consider when developing a collaborative viewing strategy is the establishment of territories [81] for personal, group and storage purposes that is suggested as beneficial for group coordination (see “Coordination” above),

A study by Yost and North [109] compared the ability of visualizations to display large amounts of data normalized across either a small or a large high-resolution vertical display. Their study showed that the visualizations used were perceptually scalable but that people preferred different visualizations on the large vs. small display, as some were found to be easier to read than others depending on the screen size. How these preferences would change for collaborative work would have to be evaluated.

Interaction Issues: Most interaction issues deal with interaction with representations, presentations and views, thus discussing them here would overlap with points raised under these headings. However, there are some more general interaction issues. When people are co-located, they are in the situation in which people naturally collaborate, the situation in which people have collaborated for centuries. When face-to-face, people naturally know how to collaborate and are so used to picking up subtle cues from each other that they may do this without even being conscious of the precise details of the underlying coordination and communication practices that are in play. As the developers of co-located collaborative information visualizations, our task is to facilitate information access and exploration without interfering with the social protocols that make collaboration effective. However, to do this we have to understand what these social collaboration practices are and specifically if there are any differences when people are collaborating using visual information. Some factors are:

Interactive Response Rates: Information visualization has always had a lot of requirements in that it deals with extremely large and complex data sets and in that it can have considerable graphics requirements for these complex representations. Adding larger screens, more screens, higher pixel counts, multiple simultaneous inputs, and possibly multiple representations will increase computational load adding more requirements to the challenge of maintaining good interactive rates. Thus implementations of collaborative information visualizations will have to be carefully designed for efficiency. While continued hardware advances will mitigate this to some extent, it will be important to address issues in both efficient data processing and fast graphic rendering.

Interaction History: A history task has been defined as a task that involves keeping a history of actions to support undo, replay, and progressive refinement [85]. In a collaborative scenario keeping such a history can have other benefits. If a visualization tracks and reveals which data items have been visited and by whom this information could be valuable for collaborators helping them understand their team members' actions, find unexplored parts of a visualization or to confirm discoveries made by others. A visualized interaction history may support collaboration by promoting mutual understanding of team members involvement in the task [24] and may help keep group members aware of each others actions as people shift from individual to shared views of the data [39]. An exploration history can be useful in such activities as validating work done, in explaining a discovery process to other team members, and in supporting discussions about data explorations.

Information Access: Exactly how to handle information access is an important collaboration issue. The main themes in the research discussion thus far have been motivated by social protocol issues and data centric concerns. While these have not been seen as mutually exclusive they are quite distinct ideas. The social protocol theme has made considerable use of observational studies to better understand exactly what are the social protocols and how do they impact collaboration. These understandings are then used as a basis for software design.

The data centric approach discusses factors such as who has (or does not have) rights to which parts of the data?, who can change the scale, zoom, or rotation settings for a shared view of the data? And how does a data item get passed between team members (hand-off). Restriction has been suggested as a means to stop certain members from making unsuspected global changes to the data that might change other members' view of the same data [75]. Similar issues pertaining to workspace awareness (individual vs. shared views), artefact manipulation (who can make which changes), and view representation have been raised [39]. Is a single shared representation adequate? Should a system allow for multiple representations? Should the exploration on multiple representations of the same dataset be linked or be completely independent?

Fluid Interaction: The fluidity of interactions in a shared workspace influences how much collaborators can focus on their task rather than on the manipulation of interface items [82]. This implies that in a collaborative information analysis scenario, parameter changes to the presentation or representation of a dataset should require manipulation of as few interface widgets (menus, slider, etc.) as possible and little or no changes of input modalities (mouse, keyboard, pen, etc.). A study on collaborative information visualization systems has similarly reported that groups worked more effectively with a system in which the required interactions were easier to understand [61]. This poses a challenge to information visualization tool designers as typically a high number of parameters are required in visualization systems to adapt to the variability in dataset complexity, size, and user tasks.

4 Collaborative Visualization on the Web

Visual analysis is rarely a solitary activity. A business analyst may notice an unexpected trend in a chart of sales figures – but then she's likely to confer with a colleague, who may share the chart with a manager, who later might present it to executives. Such scenarios of collaboration and presentation across both time and space are common in business and scientific visualization. Just as a good visualization takes advantage of the power of the human visual system, it can also exploit our natural social abilities. Accordingly, designers of visualization systems should consider not only the space of visual encodings, but mechanisms for sharing and collaboration. At minimum, systems should enable people to communicate about what they see so they can point out discoveries, share knowledge, and discuss hypotheses.

The social aspects of visualization have taken on new importance with the rise of the Web. While collaboration in small groups remains ubiquitous, it is now also possible for thousands of people to analyze and discuss visualizations together. These scenarios are driven by the fact that users can interact remotely from anywhere on the globe and access the system at different times. Partitioning work across time and space holds the potential for greater scalability of group-oriented analysis. For example, one decision making study found that asynchronous collaboration resulted in higher-quality outcomes – broader discussions, more complete reports, and longer solutions – than face-to-face collaboration [6].

Web-based collaboration around visualizations introduces new challenges for research, as most work on collaborative visualization has been done in the context of synchronous scenarios: users interacting at the same time to analyze scientific results or discuss the state of a battlefield. As described in the previous section, co-located collaboration usually involves shared displays, including large wall-sized screens and table-top devices (e. g., [28,31]). Systems supporting remote collaboration have primarily focused on synchronous interaction [3,14]), such as shared virtual workspaces (e. g., [1,24]) and augmented reality systems that enable multiple users to interact concurrently with visualized data (e. g., [25,9]). In addition, the increasing availability of table-top and large public displays has prompted researchers to experiment with asynchronous, co-located visualization (same place, different time), often in the form of ambient information displays (e. g., [21]).

In this section, we instead focus on the kind of collaboration that is most common over the Web: remote collaboration across time and space. Our goal is to summarize the work done to date and indicate promising research directions. We first review recent web-based systems supporting social data analysis around visualizations, highlighting the collaborative features provided by these systems and how they have been used in practice. We then discuss a number of outstanding challenges for asynchronous collaborative visualization and identify avenues for future research.

4.1 Web-Based Collaborative Visualization Systems

Though web-based collaboration around visualizations is still in its infancy, a handful of commercial and research systems in this area have recently been introduced. Here we discuss contemporary visualization systems that support asynchronous collaborative analysis (shown in Fig. 2), documenting the collaborative features supported by these tools and initial reports of their usage.

DecisionSite Posters: Adding Collaboration to a Single-User Tool: DecisionSite Posters is a feature of the Spotfire product sold by TIBCO, Inc. Users of Spotfire's desktop-based visualization system can capture snapshots of their analyses and publish them on an intranet as "posters." View sharing is supported, as each poster has a unique URL than can be easily distributed. Each poster also sup-ports unthreaded text comments on a side panel. However, posters do not allow annotations, limiting the ability of collaborators to point at specific trends or outliers.

As described in [96], the communication capabilities in DecisionSite Posters have been used in an unexpected way. Instead of engaging in complex conversations by using the comment panel, as envisioned by the system's designers, users have largely used the tool for presenting their findings to colleagues. The ability

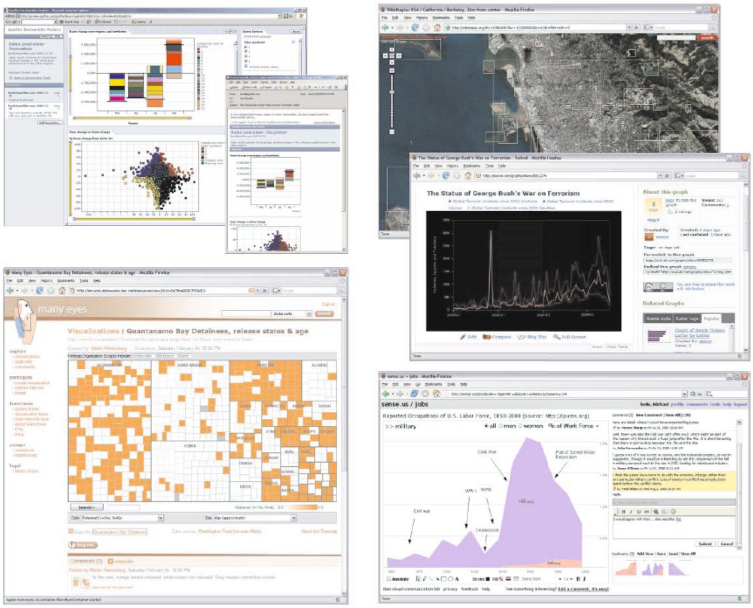


Fig. 2. Asynchronous Collaborative Visualization Systems. Clockwise from top-left: Spotfire DecisionSite Posters, Wikimapia, Swivel, sense.us, and Many Eyes.

to create comments with pointers into the visualization provides an easy way to choreograph a step-by-step presentation.

Swivel: Sharing Data on the Web: Swivel.com is a web site that supports sharing and discussion around data. The service appears to be modeled on sites such as YouTube that support sharing of other media. In keeping with this model, Swivel allows users to upload data sets and talk about them in attached discussion forums. In addition, the site automatically generates graphs by combining columns from uploaded data sets into bar charts, pie charts, and scatter plots. Pointing behavior on the site appears limited.

Although the graphs on Swivel are not interactive, the site provides an example of social data analysis in action, in particular the importance of collaborative publishing and sharing of visualizations. While there do not seem to be many extensive conversations in Swivel's discussion area there has been significant use of Swivel's graphs among bloggers to discuss statistics. In other words, it appears that the ability to publish graphs for use in other contexts is most valuable to Swivel's users.

Wikimapia: Collaborative Geographic Annotation: Wikimapia.org is a web site enabling collective annotation of geographic satellite imagery, and is representative of similar efforts such as Google Earth and mash-ups created with web APIs to mapping services. The site provides a zoomable browser of satellite photos across the globe, along with the ability to select geographic regions for annotation with names and additional data (Fig. 2). View sharing is supported through automatically updating URLs. As the view is panned or zoomed, the current URL updates dynamically to reflect the current zoom level and latitude and longitude values. Pointing is supported through annotations. Users can draw rectangular and polygonal annotations, which scale appropriately as the map is zoomed. To avoid clutter, annotations are filtered as the view is zoomed; the viewer does not see annotations that are too small to be legible or so large they engulf the entire display, improving the scalability of the system.

Wikimapia supports conversation using an embedded discussion technique. Each annotation is a link to editable text. Descriptive text about a geographic region can then be edited by anyone, similar to articles on Wikipedia. Discussion also occurs through voting. When annotations are new, users can vote on whether they agree or disagree with the annotation. Annotations that are voted down are removed from the system. For instance, the small town of Yelapa, Mexico is located on an inlet in a bay near Puerto Vallarta. However, the bay has a number of inlets very close together. As a result, multiple conflicting annotations for Yelapa appeared. Through voting, the incorrect regions were discarded and the correct annotation was preserved.

sense.us: Social Data Analysis of U.S. Census Data: Sense.us is a prototype web application for social visual data analysis [46]. The site provides interactive visualizations of 150 years of United States census data, including stacked timelines, choropleth maps, and population pyramids. With a URL-based bookmarking mechanism, it supports collaboration through doubly-linked discussion, graphical annotations, bookmark trails, and searchable comment listings.

Discussion occurs via a doubly-linked conversation model. Searchable comment listings provide links back into the visualization, while navigating in a visualization automatically causes related comments to be retrieved for the current view. By tying commentary directly to specific view states, comments become less ambiguous, enabling remarks such as “that spike” or “on the left” to be more easily understood. Pointing occurs through freeform graphical annotations and view sharing is facilitated by URLs that automatically update as users navigate the visualizations. Sense.us also allows users to collect view links in a “bookmark trail” of view thumbnails. Users can then drag-and-drop view thumbnails from the trail into a comment text field, thereby adding a hyperlink to the saved view within the comment. In this way, users can provide pointers to related views and create tours through the data.

Studies of the sense.us system revealed interesting patterns of social data analysis. Users would make observations about the data, often coupled with questions and hypotheses about the data. These comments often attracted further discussion. For example, within a visualization of the U.S. labor force over time, a spike and then decline in the number of dentists prompted discussion ranging from the fluoridation of water to stratification within the dentistry profession, with a rise in the number of hygienists corresponding to the decline of dentists. There was also an interesting interaction between data analysis and social activity. Users who tired of exploring visualizations turned their focus to the comment listings. Reading others' comments sparked new questions that led users back into the visualization, stimulating further analysis. The sense.us prototype was initially available on a corporate intranet which provided employees with blogs and a social bookmarking service. Users of sense.us found ways to publish their findings, typically by taking screenshots and then placing them on blogs or the bookmarking service with application bookmarks. These published visualizations drew additional traffic to the site.

Many Eyes: Web-Based Visualization and Publishing: Many-Eyes.com [97] is web-based service that combines public data sharing with interactive visualizations. Like social data analysis sites such as Swivel, site members can upload data sets and comment on them. Unlike Swivel, however, Many Eyes offers a palette of interactive visualization techniques – ranging from bar charts to treemaps to tag clouds – that visitors may apply to any data set. Users may post comments on the visualizations, including bookmarks for particular states.

The pointing and discussion capabilities of Many Eyes are used in a variety of ways. The site contains some lengthy conversations around visualizations, although the great majority of visualizations have no comments. One class of visualizations, however, did lead to lengthy onsite discussions: visualizations that sidestepped sober analysis and were instead playful or comical. One person, for example, initiated a game based on a visualization of Shakespearean poetry in which he used the highlighting mechanisms to pick out alphabetically ordered words to make pseudo-Elizabethan epithets. These games frequently attracted many "players."

The Many Eyes site also can be viewed as a publishing platform, since the visualizations that users create are publicly visible and may be linked to from other web pages. Many bloggers have taken advantage of this, and perhaps as a result the deepest analyses of Many Eyes visualizations have occurred as part of blog entries that reference the site. In one example, a blog at the Sunlight Foundation (a political reform organization) published a Many Eyes tag cloud to analyze messages between the U.S. Congress and the U.S. Department of Defense. The blog entry framed the results as a funny-but-sad surprise: the most common phrases had nothing to do with current pressing issues, but rather requests for congressional travel. In another case, a user created a visualization of the "social network" of the New Testament. Not only was this visualization linked to from more than 100 blog entries, but another user went to the trouble of recording a YouTube video of himself interacting with the visualization and narrating what he learned. These phenomena again underscore the importance of publishing mechanisms for collaborative visualization.

Summary: The previous examples of web-based collaborative visualization present a number of common design decisions, but also some important differences. All systems support view sharing through URL bookmarks and enable discussion through text comments. Furthermore, usage examples from these systems suggest that users derive great value from being able to share and embed visualizations in external media such as blogs. One salient difference between systems is the varied forms of pointing within visualizations: selecting individual items in Many-Eyes, creating polygonal geographic regions in Wikimapia, and drawing freeform graphics in sense.us. Another difference is the way commentary is attached to visualization views. Spotfire Decision Site Posters, Swivel, and Many Eyes all support blog-style unthreaded comments for individual visualizations. In contrast, Wikimapia supports commentary attached to geographic annotations, while sense.us provides threaded comments tied to specific states of the visualization and retrieved dynamically during exploration.

4.2 Research Issues in Web-Based Collaborative Visualization

As described in the previous section, developers of collaborative visualization systems face design decisions of how to support discussion, annotation, and integration with external services. Future research in asynchronous collaborative visualization needs to provide guidance through this design space, as well as develop novel techniques for better facilitating collaborative analysis. In this section, we identify five areas in which we expect additional research to make important contributions to improving the state-of-the-art: the structure and integration of collaborative contributions; engagement and incentives; coordination and awareness; pointing and reference; and presentation, dissemination, and story-telling.

Structuring and Integrating Contributions: A fundamental aspect of successful collaboration is an effective division of labor among participants. This involves both the segmentation of effort into proper units of work and the allocation of individuals to tasks in a manner that best matches their skills and disposition. Primary concerns are how to split work among multiple participants and meaningfully aggregate the results.

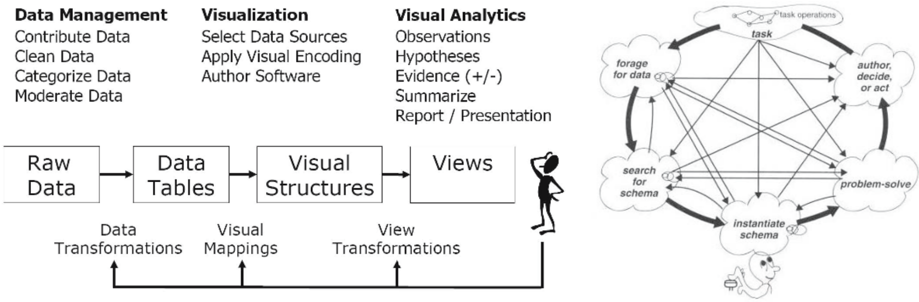
Drawing on examples such as online discussions, open source software, and Wikipedia, Benkler [7] introduces the concepts of modularity, granularity, and cost of integration in the peer production of information goods. Modularity refers to how work is segmented into individual units of contribution, while granularity refers to the scope of these units and how much effort they require. For example, in online scenarios where incentives tend to be small and non-monetary, a small granularity may encourage people to participate in part due to the ease of contributing. The cost of integration refers to the effort required to usefully synthesize contributions into a greater whole. Collaborative work will only be effective if the cost of integration is low enough to warrant the overhead of modularization while enforcing adequate quality control. There are a number of mutually inclusive approaches to handling integration: automation (automatically integrating work through technological means), peer production (casting integration as an additional collaborative task given to trusted participants), social norms (using social pressures to reduce vandalistic behavior), and hierarchical control (exercising explicit moderation).

Collaborative visualization can similarly be viewed as a process of peer production of information goods. Stages in this process include uploading data sets, creating visualizations, and conducting analysis. To support this process, it is important to identify the specific forms of contribution (modules) that users might make and how to integrate these contributions. Existing frameworks for aiding this task include structural models of visualization design and sensemaking processes [17]. As shown in Fig. 3, each of these models suggests tasks that contribute to collaborative analysis, including data cleaning, moderation, visualization specification, sharing observations, positing hypotheses, and marshaling evidence. These concerns are given further treatment in [43].

Once modules have been identified, one can then attempt designs which reduce the cost structure of these tasks. Consider the issue of scale. Most of the examples in the previous section use sequential text comments to conduct analytic discussion. However, it is unclear how well this form of communication will scale to massive audiences. An open research problem is the creation of new forms of managed conversation that have a lower cost of integration, enabling people to understand and contribute to analysis without having to wade through hundreds of individual comments. For example, Wikipedia relies on human editing coupled with a revision management system to integrate and moderate contributions. Alternatively, systems with highly structured input such as NASA ClickWorkers [7] or von Ahn's (2006) "games with a purpose" [98] rely on purely automated techniques. Some middle ground between these approaches should be possible for collaborative analysis, such as argumentation systems that model hypotheses and evidence as first class objects. One example of such a system is CACHE [11], which maintains a matrix of hypotheses and evidence, with collaborators providing numerical measures of the reliability of evidence and the degree to which evidence confirms or disconfirms the hypotheses. These scores can then be averaged to form a group assessment. Other possibilities include augmenting graphical workspaces such as the Analysis Sandbox [107] with collaborative authoring features or automatic merging of representations (c.f., [13]).

Engagement and Incentives: If collaborators are professionals working within a particular context (e. g., financial analysts or research scientists) there may be existing incentives, both financial and professional, for conducting collaborative work. In a public goods scenario, incentives such as social visibility or sense of contribution may be motivating factors. Incorporating incentives into the design process of collaborative visualization systems may increase the level of user contributions, and could even provide additional motivation in situations that already have well established incentive systems.

Benkler posits an incentive structure for collaborative work consisting of monetary, hedonic, and social-psychological incentives [7]. Monetary incentives refer to material compensation such as a salary or cash reward. Hedonic incentives refer to well-being or engagement experienced intrinsically in the work. Social-psychological incentives involve perceived benefits such as increased status or social capital.



(a) Collaborative activity might be introduced at any phase of the information visualization pipeline.

(b) The sensemaking model in [17] can be applied to identify potential mechanisms for collaborative analysis (e. g., [43])

Fig. 3. Models of the Visualization Process.

Observations of social use of visualization have noted that visualization users are attracted to data which they find personally relevant [45,96,100]. For example, in collaborative visual analysis of the occupations of American workers [46]), users often search for their own profession and those of their friends and family, similar to how people search for names in the popular NameVoyager visualization [100]. The hypothesis is that by selecting data sets or designing their presentation such that the data is seen as personally relevant, usage rates will rise due to increased hedonic incentive. For example, geographic visualizations may facilitate navigation to personally relevant locations through typing in zip codes or city names, while a visualization of the United States' budget might communicate how a specific user's taxes were allocated rather than only listing total dollar amounts.

In the case of social-psychological incentives, the visibility of contributions can be manipulated for social effects. Ling et al [56] found that users contributed more if reminded of the uniqueness of their contribution or if given specific challenges, but not under other theoretically-motivated conditions. Cheshire [22] describes a controlled experiment finding that, even in small doses, positive social feedback on a contribution greatly increases contributions. He also found that visibility of high levels of cooperative behavior across the community increases contributions in the short term, but has only moderate impact in the long term. These studies suggest that social-psychological incentives can improve contribution rates, but that the forms of social visibility applied have varying returns. One such incentive for visual analysis is to prominently display new discoveries or successful responses to open questions. Mechanisms for positive feedback, such as voting for interesting comments, might also foster more contributions.

Finally, it is worth considering game play as an additional framework for increasing incentives. In contrast to environments such as spreadsheets, many visualizations already enjoy game-like properties, being highly visual, highly interactive, and often animated. Heer [42] discusses various examples in which playful activity contributes to analysis, applying insights from an existing theory of playful behavior [16] that analyzes the competitive, visceral, and teamwork building aspects of play. For example, scoring mechanisms could be applied to create competitive social-psychological incentives. Game design might also be used to allocate attention, for example, by creating a team-oriented "scavenger hunt" analysis game focused on a particular subject matter. Salen and Zimmerman [78] provide a thorough resource for the further study of game design concepts.

Coordination and Awareness: An important aspect of collaborative action is awareness of others' activities, allowing collaborators to gauge what work has been done and where to allocate effort next [29,19]. Within asynchronous contexts, participants require awareness of the timing and content of past actions. This suggests that designs should include both history and notification mechanisms (e. g. [15]) for following actions performed on a given artifact or by specific individuals or groups. Browseable histories of past action are one viable mechanism, as are subscription and notification technologies such as RSS (Really Simple Syndication).

User activity can also be aggregated and abstracted to provide additional forms of awareness. Social navigation [30] involves the use of activity traces to provide additional navigation options, allowing users to purposefully navigate to past states of high interest or explore less-visited regions (the "anti-social navigation" of Wattenberg & Kriss [100]). For example, navigation cues may be added to links to views with low visitation rates or to action items such as unanswered questions and unassessed hypotheses. One recent study [106] provides evidence that social navigation cues can simultaneously promote revisitation of popular or controversial views while also leading to a higher rate of unique discoveries. Future research is needed to further develop and evaluate other forms of awareness cues for supporting collaborative analysis.

Pointing and Reference: When collaborating around visual media, it is common for one to refer to visible objects, groups, or regions [26,12]. Such references may be general ("north by northwest"), definite (named entities), detailed (described by attributes, such as the "blue ball"), or deictic (pointing to an object and saying "that one"). Hill and Hollan [47] discuss the various roles that deictic pointing gestures can play, often communicating intents more complicated than simply "look here". For example, different hand gestures can communicate angle (oriented flat hand), intervals (thumb and index finger in "C" shape), groupings (lasso'ing a region), and forces (accelerating fist). While other forms of reference are often most easily achieved through speech or written text, deictic reference in particular offers important interface design challenges for collaborative visualization. Nuanced pointing behaviors can improve collaboration by making it easier to establish the object of conversation. Hill and Hollan argue for "generally applicable techniques that realize complex pointing intentions" by engaging "pre-attentive vision in the service of cognitive tasks."

A standard way to point in a visualization is brushing: selecting and highlighting a subset of the data. Naturally, these selections should be sharable as part of the state of the visualization. In addition, a palette of visual effects richer than simple highlighting can let users communicate different intents. For example, following time-varying values of selected items in a scatter plot is easier when the selected items leave trails as they move over time. The selected items and their trails are even more salient if non-selected items are simultaneously de-emphasized. Brushing-based forms of pointing have the advantage that the pointing action is tied directly to the data, allowing the same pointing gesture to be reapplied in different views of the same data. As "data-aware" annotations are machine-readable, they can also be used to export subsets of data and help steer automated data mining [108].

Freeform graphical annotations are a more expressive method of pointing in visualizations. Drawing a circle around a cluster of items or pointing an arrow at a peak in a graph can direct the attention of remote viewers; at the same time, the angle of the arrow or shape of the hand-drawn circle may communicate emotional cues or add emphasis. However, while such drawing and vector graphic annotations allow a high degree of expression, they only apply to a single view in the visualization, without any explicit tie to the underlying data. Freeform annotations can persist over purely visual transformations such as panning and zooming, but they are not data-aware and may become meaningless in the face of data-oriented operations such as filtering or drill-down. A promising research direction is hybrid approaches that combine aspects of both brushing and graphical annotation. The resulting techniques could create graphical annotations that are tied to data points so that they can be reapplied in other views of the data.

Presentation, Dissemination, and Story-Telling: Common forms of information exchange in group sensemaking are reports and presentations. Narrative presentation of an analysis "story" is a natural and often effective way to communicate findings, and has been observed as a primary use of Decision Site Posters. Furthermore, usage of Swivel, sense.us, and Many Eyes leverages external media such as blogs and social bookmarking services as additional communication channels in which to share and discuss findings from visualizations. The challenge to collaborative visualization is to provide mechanisms to aid the creation and distribution of presentations. For example, sense.us [46] allows users to construct and share trails of related views to create tours spanning multiple visualizations and the GeoTime Stories [32] system supports textual story-telling with hyperlinks to visualization states and annotations. However, neither system yet allows these stories to be exported outside the respective applications. In future work, such mechanisms could be improved with support to build presentations semi-automatically using interaction histories, export such presentations into external media, and apply previously discussed pointing techniques. A related issue is to enable follow-up analysis and verification for parts of the analysis story, enabling presentations to serve as a catalyst for additional investigation.

4.3 Summary

In this section, we introduce an emerging use of interactive visualization: collaborative visual analysis across space and time. The Web has opened up new possibilities for large-scale collaboration around visualizations and holds the potential for improved analysis and dissemination of complex data sets. A new class of systems explores these possibilities, enabling web-based data access, exploration, view sharing, and discussion around both static and interactive visualizations. Already, these systems exhibit the promise of web-based collaboration, providing examples of collective data analysis in which group members combine their knowledge to make sense of observed data trends and disseminate their findings.

Still, many research questions remain on how to structure collaboration. For example, how can we move beyond simple textual comments to better scale and integrate diverse contributions? Interested readers may wish to consult [96,46,43] for further discussions on this topic. As described in section 2, another open question is how to design for particular audiences. Different scenarios – including scientific collaboration, business intelligence, and public data consumption – involve different skill sets, scales of collaboration, and standards of quality. Going forward, case studies in these scenarios are crucial to better tailoring visualization tools to such varied audiences. By enabling users to collectively explore data, share views and findings, and debate competing hypotheses, the resulting collaborative visual analysis systems hold the potential to improve the number and quality of insights gained from our ever-increasing collections of data.

5 Conclusion

The adoption of visualization technologies by people from different walks of life has important implications for visualization research and development. Visualization construction tools are lowering barriers to entry, resulting in end-user created visualizations of every kind of data set imaginable. Concurrently, new technologies enabling collaborative use of visualizations in both physical and online settings hold the potential to change the way we explore, analyze, and communicate. In this paper, we have sought to identify these emerging trends and provide preliminary design considerations for advancing the state-of-the-art of visualization and visual analytic tools.

As a parting comment, we note that the release of visualization tools "into the wild" will undoubtedly result in a plethora of unexpected developments. Equipped with new creation and collaboration tools, users will almost certainly re-appropriate these technologies for unexpected purposes. Already, use of systems like Many-Eyes has revealed new genres of data-oriented play and self-expression that complement more traditional analytic activities.

As researchers, it is imperative that we interface with these developments in a productive fashion. It is likely that visualization tools will not only be used in unexpected ways, but in ways we actively dislike. As new audiences are exposed to visualization technologies, "bad" or "chart junk" visualizations will be generated. Furthermore, visualizations will be used to support actions or points of view we may find distasteful, and any communication medium that is sufficiently powerful to inform may also be used to lie or misrepresent. We as a community should not be so concerned with trying to control the medium or prevent people from lying or creating bad visualizations. As audiences get more comfortable communicating with visualizations, we optimistically expect the quality of visualizations and nuance of interpretation to improve.

However, this proscription does not mean that researchers should idly sit on their hands. Rather, there will be an expanded role for visualization experts to play. Issues of data provenance, cleaning, and integrity will force the research community to focus on the visualization pipeline in a more holistic manner. Supporting data at varied levels of structure will become increasingly necessary. New genres of visualization use may require new designs and new systems to support emerging practices, and the design of visual exploration tools that both empower and educate will take on new importance. Consequently, the entrance of visualization technologies into the mainstream offers a new horizon of research opportunities.

References

1. Agrawala, M., Beers, A.C., McDowall, I., Fröhlich, B., Bolas, M., Hanrahan, P.: The Two-User Responsive Workbench: Support for Collaboration Through Individual Views of a Shared Space. In: International Conference on Computer Graphics and Interactive Techniques (Siggraph '97), pp. 327–332. ACM Press, New York (1997), doi:10.1145/258734.258875
2. Ahlberg, C., Wistrand, E.: IVEE: An information visualization & exploration environment. In: Proceedings of the IEEE Symposium on Information Visualization, Atlanta, GA, October 1995, pp. 66–73. IEEE Computer Society Press, Los Alamitos (1995)
3. Anupam, V., Bajaj, C.L., Schikore, D., Shikore, M.: Representations in distributed cognitive tasks. IEEE Computer 27(7), 37–43 (1994)
4. Artero, A.O., Ferreira de Oliveira, M.C., Levkowitz, H.: Uncovering clusters in crowded parallel coordinates visualizations. In: Proceedings of the IEEE Symposium on Information Visualization (InfoVis), pp. 81–88. IEEE Computer Society Press, Los Alamitos (2004)
5. Baldonado, M.Q.W., Woodruff, A., Kuchinsky, A.: Guidelines for using multiple views in information visualization. In: Proceedings of AVI '00, pp. 110–119. ACM Press, New York (2000), doi:10.1145/345513.345271
6. Benbunan-Fich, R., Hiltz, S.R., Turoff, M.: A comparative content analysis of face-to-face vs. asynchronous group decision making. Decision Support Systems 34(4), 457–469 (2003)
7. Benkler, Y.: Coase's penguin, or, linux and the nature of the firm. Yale Law Journal 112(369) (2002)
8. Benko, H., Wilson, A.D., Baudisch, P.: Precise Selection Techniques for Multi-Touch Screens. In: Proceedings of the Conference on Human Factors in Computing Systems (CHI'06), Montréal, Canada, April 22-27, 2006, pp. 1263–1272. ACM Press, New York (2006), doi:10.1145/1124772.1124963
9. Benko, H., Ishak, E.W., Feiner, S.: Collaborative mixed reality visualization of an archaeological excavation. In: IEEE International Symposium on Mixed and Augmented Reality (ISMAR 2004), Arlington, VA, pp. 132–140 (2004)

10. Bertin, J.: Semiology of Graphics: Diagrams Networks Maps (Translation of: Sémiologie graphique). The University of Wisconsin Press, Madison (1983)
11. Billman, D., Convertino, G., Shrager, J., Pirolli, P., Massar, J.P.: Collaborative intelligence analysis with cache and its effects on information gathering and cognitive bias. In: Human Computer Interaction Consortium Workshop (2006)
12. Brennan, S.E.: How conversation is shaped by visual and spoken evidence. In: Trueswell, Tanenhaus (eds.) Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions, pp. 95–129. MIT Press, Cambridge (2005)
13. Brennan, S.E., Mueller, K., Zelinsky, G., Ramakrishnan, I.V., Warren, D.S., Kaufman, A.: Toward a multi-analyst, collaborative framework for visual analytics. In: IEEE Symposium on Visual Analytics Science and Technology (2006)
14. Brodlie, K.W., Duce, D.A., Gallop, J.R., Walton, J.P.R.B., Wood, J.D.: Distributed and collaborative visualization. Computer Graphics Forum 23(2), 223–251 (2004)
15. Brush, A.J., Bargeron, D., Grudin, J., Gupta, A.: Notification for shared annotation of digital documents. In: Proc. ACM Conference on Human Factors in Computing Systems (CHI'02) (2002)
16. Caillois, R.: Man, Play, and Games. Free Press of Glencoe (1961)
17. Card, S., Mackinlay, J.D., Shneiderman, B.: Readings In Information Visualization: Using Vision To Think. Morgan Kauffman Publishers, Inc., San Francisco (1999)
18. Card, S.K., Nation, D.: Degree-of-Interest Trees: A component of an attention-reactive user interface. In: Proceedings of the Working Conference on Advanced Visual Interfaces, May 2002, pp. 231–245. ACM Press, New York (2002), http://www2.parc.com/istl/projects/uir/pubs/items/UIR-2002-11-Card-AVI-DOITree.pdf
19. Caroll, J., Rosson, M.B., Convertino, G., Ganoe, C.H.: Awareness and teamwork in computer-supported collaborations. Interacting with Computers 18(1), 21–46 (2005)
20. Carpendale, M.S.T., Montagnese, C.: A framework for unifying presentation space. In: Schilit, B. (ed.) Proceedings of ACM Symposium on User Interface Software and Technology (UIST), pp. 61–70. ACM Press, New York (2001), http://pages.cpsc.ucalgary.ca/~sheelagh/personal/pubs/2001/carpendaleuist01.pdf, doi:10.1145/502348.502358
21. Carter, S., Mankoff, J., Goddi, P.: Building connections among loosely coupled groups: Hebb's rule at work. Journal of Computer-Supported Cooperative Work 13(3), 305–327 (2004)
22. Cheshire, C.: Selective incentives and generalized information exchange. Social Psychology Quarterly 70(1) (2007)
23. Chi, E.H.-H., Riedl, J.T.: An Operator Interaction Framework for Visualization Systems. In: Wills, G., Dill, J. (eds.) Proceedings of the IEEE Symposium on Information Visualization (InfoVis '98), pp. 63–70. IEEE Computer Society Press, Los Alamitos (1998)
24. Chuah, M.C., Roth, S.F.: Visualizing Common Ground. In: Proc. of the Conf. on Information Visualization (IV), pp. 365–372. IEEE Computer Society Press, Los Alamitos (2003)
25. Chui, Y.-P., Heng, P.-A.: Enhancing view consistency in collaborative medical visu-alization systems using predictive-based attitude estimation. In: First IEEE International Workshop on Medical Imaging and Augmented Reality (MIAR'01), Hong Kong, China (2001)

26. Clark, H.H.: Pointing and placing. In: Kita, S. (ed.) Pointing. Where language, culture, and cognition meet, pp. 243–268. Lawrence Erlbaum, Mahwah (2003)
27. Cluxton, D., Eick, S.G., Yun, J.: Hypothesis visualization. In: Proceedings of the IEEE Symposium on Information Visualization (Posters Compendium), Austin, TX, October 2004, pp. 9–10. IEEE Computer Society Press, Los Alamitos (2004)
28. Dietz, P.H., Leigh, D.L.: Diamondtouch: A multi-user touch technology. In: Proc. ACM Symposium on User Interface Software and Technology, pp. 219–226 (2001)
29. Dourish, P., Belotti, V.: Awareness and coordination in shared workspaces. In: Proc. ACM Conference on Computer-Supported Cooperative Work, Toronto, Ontario, pp. 107–114 (1992)
30. Dourish, P., Chalmers, M.: Running out of space: Models of information navigation. In: Proc. Human Computer Interaction (HCI'94) (1994)
31. Dynamics, G.: Command post of the future. Website (accessed November 2007)
32. Eccles, R., Kapler, T., Harper, R., Wright, W.: Stories in geotime. In: Proc. IEEE Symposium on Visual Analytics Science and Technology (2007)
33. Ellson, J., Gansner, E.R., Koutsofios, E., North, S.C., Woodhull, G.: Graphviz and Dynagraph – static and dynamic graph drawing tools. Online Documentation (accessed November 2007)
34. Fekete, J.-D.: The Infovis Toolkit. In: Ward, M., Munzner, T. (eds.) Proceedings of the IEEE Symposium on Information Visualization (InfoVis), pp. 167–174. IEEE Computer Society Press, Los Alamitos (2004)
35. Forlines, C., Shen, C.: DTLens: Multi-user Tabletop Spatial Data Exploration. In: Proc. of User Interface Software and Technology (UIST), pp. 119–122. ACM Press, New York (2005), doi:10.1145/1095034.1095055
36. Gahegan, M., Wachowicz, M., Harrower, M., Rhyne, T.-M.: The integration of geographic visualization with knowledge discovery in databases and geocomputation. Cartography and Geographic Information Society 28(1), 29–44 (2001)
37. Gershon, N.: What storytelling can do for information visualization. Communications of the ACM 44(8), 31–37 (2001)
38. Grimstead, I.J., Walker, D.W., Avis, N.J.: Collaborative visualization: A review and taxonomy. In: Proceedings of the Symposium on Distributed Simulation and Real-Time Applications, pp. 61–69. IEEE Computer Society Press, Los Alamitos (2005)
39. Gutwin, C., Greenberg, S.: Design for individuals, design for groups: Tradeoffs between power and workspace awareness. In: Proceedings of Computer Supported Cooperative Work (CSCW), pp. 207–216. ACM Press, New York (1998), doi:10.1145/289444.289495
40. Hancock, M., Carpendale, S.: Supporting multiple off-axis viewpoints at a tabletop display. In: Proceedings of Tabletop, pp. 171–178. IEEE Computer Society Press, Los Alamitos (2007)
41. Heer, J.: The flare visualization toolkit. Website (accessed November 2007)
42. Heer, J.: Socializing visualization. In: Proc. CHI 2006 Workshop on Social Visualization (2006)
43. Heer, J., Agrawala, M.: Design Considerations for Collaborative Visual Analytics. In: IEEE Symposium on Visual Analytics Science and Technology (VAST), pp. 171–178. IEEE Computer Society Press, Los Alamitos (2007), `http://vis.berkeley.edu/papers/design_collab_vis/2007-DesignCollabVis-VAST.pdf`
44. Heer, J., Card, S.K., Landay, J.A.: prefuse: A toolkit for interactive information visualization. In: Proceedings of the Conference on Human Factors in Computing Systems (CHI), pp. 421–430. ACM Press, New York (2005), doi:10.1145/1054972.1055031

45. Heer, J., boyd, d.: Vizster: Visualizing Online Social Networks. In: Proceedings of the IEEE Symposium on Information Visualization (InfoVis '05), pp. 33–40. IEEE Computer Society Press, Los Alamitos (2005)
46. Heer, J., Viégas, F.B., Wattenberg, M.: Voyagers and voyeurs: Supporting asynchronous collaborative information visualization. In: Proceedings of the Conference on Human Factors in Computing Systems (CHI), pp. 1029–1038. ACM Press, New York (2007)
47. Hill, W.C., Hollan, J.D.: Deixis and the future of visualization excellence. In: Proc. of IEEE Visualization, pp. 314–319 (1991)
48. Isenberg, T., Neumann, P., Carpendale, S., Nix, S., Greenberg, S.: Interactive annotations on large, high-resolution information displays. In: Conference Compendium of IEEE VIS, InfoVis, and VAST, pp. 124–125. IEEE Computer Society Press, Los Alamitos (2006), `http://cpsc.ucalgary.ca/~isenberg/papers/Isenberg_2006_IAL.pdf`
49. Jerding, D.F., Stasko, J.T.: The Information Mural: A Technique for Displaying and Navigating Large Information Spaces. In: Proceedings of the IEEE Symposium on Information Visualization (InfoVis), pp. 43–50. IEEE Computer Society Press, Los Alamitos (1995)
50. Johnson, B., Shneiderman, B.: Tree-maps: A space-filling approach to the visualization of hierarchical information structures. In: Proceedings of IEEE Visualization, pp. 284–291. IEEE Computer Society Press, Los Alamitos (1991)
51. Kerren, A., Stasko, J.T., Fekete, J.-D., North, C.J. (eds.): Information Visualization. LNCS, vol. 4950. Springer, Heidelberg (2008)
52. Kleinmutz, D.N., Schkade, D.A.: Information Displays and Decision Processes. Psychological Science 4(4), 221–227 (1993)
53. Kruger, R., Carpendale, S., Scott, S.D., Greenberg, S.: Roles of orientation in tabletop collaboration: Comprehension, coordination and communication. Journal of Computer Supported Collaborative Work 13(5–6), 501–537 (2004)
54. Kruger, R., Carpendale, S., Scott, S.D., Tang, A.: Fluid integration of rotation and translation. In: Proceedings of Human Factors in Computing Systems (CHI), pp. 601–610. ACM Press, New York (2005), doi:10.1145/1054972.1055055
55. Lamping, J., Rao, R., Pirolli, P.: A focus + context technique based on hyperbolic geometry for visualizing large hierarchies. In: Proceedings of the Conference of Human Factors in Computing Systems,CHI, pp. 401–408. ACM Press, New York (1995), doi:10.1145/223904.223956
56. Ling, K., Beenen, G., Lundford, P., Wang, X., Chang, K., Li, X., Cosley, D., Frankowski, D., Terveen, L., Rashid, A.M., Resnick, P., Kraut, R.: Using social psychology to motivate contributions to online communities. Journal of Computer-Mediated Communication 10(4) (2005)
57. Liu, Y., Gahegan, M., Macgill, J.: Increasing geocomputational interoperability: Towards a standard geocomputation API. In: Proceedings of GeoComputation, Ann Arbor, MI (2005)
58. Livny, M., Ramakrishnan, R., Beyer, K., Chen, G., Donjerkovic, D., Lawande, S., Myllymaki, J., Wenger, K.: DEVise: Integrated querying and visualization of large datasets. In: Proceedings of the International Conference on Management of Data (SIGMOD), Tucson, AZ, pp. 301–312. ACM Press, New York (1997)
59. Mackinlay, J.D.: Automating the design of graphical presentations of relational information. ACM Transactions on Graphics 5(2), 110–141 (1986)
60. Mackinlay, J.D., Hanrahan, P., Stolte, C.: Show me: Automatic presentation for visual analysis. IEEE Transactions on Visualization and Computer Graphics 13(6), 1137–1144 (2007)

61. Mark, G., Kobsa, A.: The effects of collaboration and system transparency on CIVE usage: An empirical study and model. Presence 14(1), 60–80 (2005)
62. Marr, D.: Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. W.H. Freeman, New York (1982)
63. Merriam-Webster: Webster's english dictionary. Website (accessed November 2007), `http://www.cs.chalmers.se/~hallgren/wget.cgi?presentation`
64. Munzner, T.: Guest Editor's Introduction: Information Visualization. Computer Graphics and Applications 22(1), 20–21 (2002)
65. Munzner, T., Guimbretiére, F., Tasiran, S., Zhang, L., Zhou, Y.: TreeJuxtaposer: Scalable Tree Comparison Using Focus+Context with Guaranteed Visibility. ACM Transactions on Graphics 22(3), 453–462 (2003)
66. Nacenta, M.A., Sakurai, S., Yamaguchi, T., Miki, Y., Itoh, Y., Kitamura, Y., Subramanian, S., Gutwin, C.: E-conic: a perspective-aware interface for multi-display environments. In: Proceedings of ACM Symposium on User Interface Software and Technology (UIST'01), pp. 279–288. ACM Press, New York (2007)
67. Neumann, P., Schlechtweg, S., Carpendale, M.S.T.: ArcTrees: Visualizing Relations in Hierarchical Data. In: Proceedings of Eurographics / IEEE VGTC Symposium on Visualization (EuroVis 2005, June 1–3, 2005, Leeds, England, UK), Aire-la-Ville. Eurographics Workshop Series, pp. 53–60. Eurographics (2005)
68. Neumann, P., Tang, A., Carpendale, S.: A Framework for Visual Information Analysis. Technical Report 2007-87123, University of Calgary, Calgary, AB, Canada (July 2007)
69. North, C., Shneiderman, B.: Snap-together visualization: A user interface for coordinating visualizations via relational schemata. In: Proceedings of the Working Conference on Advanced Visual Interfaces (AVI), May 2000, pp. 128–135. ACM Press, New York (2000)
70. Olson, G.M., Olson, J.S.: Distance Matters. Human-Computer Interaction 15(2 & 3), 139–178 (2000)
71. Olston, C., Woodruff, A., Aiken, A., Chu, M., Ercegovac, V., Lin, M., Spalding, M., Stonebraker, M.: Datasplash. In: Proceedings of the International Conference on Management of Data (SIGMOD), Seattle, WA, June 1998, pp. 550–552. ACM Press, New York (1998)
72. Park, K.S., Kapoor, A., Leigh, J.: Lessons learned from employing multiple perspectives in a collaborative virtual environment for visualizing scientific data. In: Proceedings of Collaborative Virtual Environments (CVE), pp. 73–82. ACM Press, New York (2000), doi:10.1145/351006.351015
73. Pinelle, D., Gutwin, C., Greenberg, S.: Task analysis for groupware usability evaluation: Modeling shared-workspace tasks with the mechanics of collaboration. ACM Transaction of Human Computer Interaction 10(4), 281–311 (2003)
74. Rensink, R.A.: chapter Change Blindness. In: McGraw-Hill Yearbook of Science and Technology, pp. 44–46. McGraw-Hill, New York (2005)
75. Meredith, Ryall, K., Shen, C., Forlines, C., Morris, F.V.R.: Beyond "social protocols": Multi-user coordination policies for co-located groupware. In: Proceedings of Computer-Supported Cooperative Work (CSCW), pp. 262–265. ACM Press, New York (2004), doi:10.1145/1031607.1031648
76. Rogers, Y., Lindley, S.: Collaborating around vertical and horizontal large interactive displays: Which way is best? Interacting with Computers 16(6), 1133–1152 (2004)
77. Meredith, Everitt, K., Ryall, F.V.K., Esenther, A., Forlines, C., Shen, C., Shipman, S., Morris, R.: Identity-differentiating widgets for multiuser interactive surfaces. IEEE Computer Graphics and Applications 26(5), 56–64 (2006)

78. Salen, K., Zimmerman, E.: Rules of Play: Fundamentals of Game Design. MIT Press, Cambridge (2003)
79. Saraiya, P., North, C., Duca, K.: An Insight-Based Methodology for Evaluating Bioinformatics Visualizations. IEEE Transactions on Visualization and Computer Graphics 11(4), 443–456 (2005)
80. Scott, S.D., Carpendale, M.S.T., Habelski, S.: Storage bins: Mobile storage for collaborative tabletop displays. IEEE Computer Graphics and Applications 25(4), 58–65 (2005), http://doi.ieeecomputersociety.org/10.1109/MCG.2005.86
81. Scott, S.D., Carpendale, M.S.T., Inkpen, K.M.: Territoriality in collaborative tabletop workspaces. In: Proceedings of Computer-Supported Cooperative Work (CSCW), pp. 294–303. ACM Press, New York (2004), http://innovis.cpsc.ucalgary.ca/pubs/2004/Territoriality.CSCW/scott_cscw2004.pdf, doi:10.1145/1031607.1031655
82. Scott, S.D., Grant, K.D., Mandryk, R.L.: System guidelines for co-located collaborative work on a tabletop display. In: Proceedings of the European Conference on Computer-Supported Cooperative Work (ECSCW), pp. 159–178. Kluwer Academic Publishers, Dordrecht (2003), http://www.ecscw.uni-siegen.de/2003/009Scott_ecscw03.pdf
83. Seo, J., Shneiderman, B.: Knowledge discovery in high dimensional data: Case studies and a user survey for an information visualization tool. IEEE Transactions on Visualization and Computer Graphics 12(3), 311–322 (2006)
84. Shen, C., Lesh, N., Vernier, F.: Personal digital historian: Story sharing around the table. ACM Interactions 10(2), 15–22 (2003)
85. Shneiderman, B.: The eyes have it: A task by data type taxonomy for information visualizations. In: Proceedings of the IEEE Symposium on Visual Languages, pp. 336–343. IEEE Computer Society Press, Los Alamitos (1996)
86. Shum, S.B., Li, V.U.G., Domingue, J., Motta, E.: Visualizing internetworked argumentation. In: Kirschner, P.A., Shum, S.J.B., Carr, C.S. (eds.) Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making, December 2002, pp. 185–204. Springer, Heidelberg (2002)
87. Spence, R.: Information Visualization, 2nd edn. Pearson Education Limited, Harlow (2007)
88. Takatsuka, M., Gahegan, M.: GeoVISTA Studio: A codeless visual programming environment for geoscientific data analysis and visualization. Computational Geoscience 28(10), 1131–1144 (2002)
89. Tandler, P., Prante, T., Müller-Tomfelde, C., Streitz, B., Steinmetz, R.: ConnecTables: Dynamic Coupling of Displays for the Flexible Creation of Shared Workspaces. In: Proceedings of User Interface Software and Technology (UIST), pp. 11–20. ACM Press, New York (2001)
90. Tang, A., Tory, M., Po, B., Neumann, P., Carpendale, S.: Collaborative coupling over tabletop displays. In: Proceedings of Human Factors in Computing Systems (CHI), pp. 1181–1290. ACM Press, New York (2006), doi:10.1145/1124772.1124950
91. Thomas, J.J., Cook, K.A.: Illuminating the Path: The Research and Development Agenda for Visual Analytics. National Visualization and Analytics Center (2005), http://nvac.pnl.gov/agenda.stm
92. Tufte, E.R.: The Visual Display of Quantitative Information. Graphic Press, Cheshire (2001)
93. van Wijk, J.J.: The value of visualization. In: Proceedings of IEEE Visualization (VIS), pp. 79–86. IEEE Computer Society Press, Los Alamitos (2005), http://www.win.tue.nl/~vanwijk/vov.pdf

94. Vernier, F., Lesh, N., Shen, C.: Visualization techniques for circular tabletop interfaces. In: Proceedings of Advanced Visual Interfaces (AVI), pp. 257–263. ACM Press, New York (2002)
95. Viégas, F.B., boyd, d., Nguyen, D.H., Potter, J., Donath, J.: Digital artifacts for remembering and storytelling: PostHistory and social network fragments. In: Proceedings of the Hawaii International Conference on System Sciences (HICCSS), pp. 105–111 (2004)
96. Viégas, F.B., Wattenberg, M.: Communication-minded visualization: A call to action. IBM Systems Journal 45(4), 801–812 (2006), doi:10.1147/sj.454.0801
97. Viégas, F.B., Wattenberg, M., van Ham, F., Kriss, J., McKeon, M.: Many Eyes: A site for visualization at internet scale. IEEE Transactions on Visualization and Computer Graphics (Proceedings Visualization / Information Visualization 2007) 12(5), 1121–1128 (2007), `http://www.research.ibm.com/visual/papers/viegasinfovis07.pdf`
98. von Ahn, L.: Games with a purpose. Computer 39(6), 92–94 (2006)
99. Ware, C.: Information Visualization – Perception for Design, 2nd edn. Morgan Kaufmann Series in Interactive Technologies. Morgan Kaufmann Publishers, San Francisco (2004)
100. Wattenberg, M., Kriss, J.: Designing for Social Data Analysis. IEEE Transactions on Visualization and Computer Graphics 12(4), 549–557 (2006)
101. Weaver, C., Fyfe, D., Robinson, A., Holdsworth, D.W., Peuquet, D.J., MacEachren, A.M.: Visual analysis of historic hotel visitation patterns. In: Proceedings of the Symposium on Visual Analytics Science and Technology (VAST), Baltimore, MD, October 31–November 2 2006, pp. 35–42. IEEE Computer Society Press, Los Alamitos (2006)
102. Weaver, C.E.: Improvise: A User Interface for Interactive Construction of Highly-Coordinated Visualizations. Phd thesis, University of Wisconsin–Madison (June 2006)
103. Wesche, G., Wind, J., Göbe, M., Rosenblum, L., Durbin, J., Doyle, R., Tate, D., King, R., Fröhlich, B., Fischer, M., Agrawala, M., Beers, A., Hanrahan, P., Bryson, S.: Application of the Responsive Workbench. IEEE Computer Graphics and Applications 17(4), 10–15 (1997), doi:10.1109/38.595260
104. Weskamp, M.: newsmap. Website (accessed November 2007), `http://marumushi.com/apps/newsmap/index.cfm`
105. Wigdor, D., Shen, C., Forlines, C., Balakrishnan, R.: Perception of elementary graphical elements in tabletop and multi-surface environments. In: Proceedings of Human Factors in Computing Systems (CHI), pp. 473–482. ACM Press, New York (2007)
106. Willett, W., Heer, J., Agrawala, M.: Scented widgets: Improving navigation cues with embedded visualizations. IEEE Transactions on Visualization and Computer Graphics 13(6), 1129–1136 (2007)
107. Wright, W., Schroh, D., Proulx, P., Skaburskis, A., Cort, B.: Advances in nSpace – the sandbox for analysis. In: International Conference on Intelligence Analysis, McLean, VA (May 2005)
108. Yang, D., Rundensteiner, E.A., Ward, M.O.: Analysis guided visual exploration to multivariate data. In: Proc. IEEE Visual Analytics Science and Technology (2007)
109. Yost, B., North, C.: The perceptual scalability of visualization. IEEE Transactions on Visualization and Computer Graphics 12(5), 837–844 (2005)
110. Zhang, J., Norman, D.A.: Representations in distributed cognitive tasks. Cognitive Science 18(1), 87–122 (1994)

111. Zuk, T., Schlesier, L., Neumann, P., Hancock, M.S., Carpendale, M.S.T.: Heuristics for Information Visualization Evaluation. In: Proceedings of the Workshop Beyond Time and Errors (BELIV), held in conjunction with AVI, pp. 55–60. ACM Press, New York (2006), doi:10.1145/1168149.1168162