

Information Visualization for the People

by

Michael Danziger

B.A. Physics
Boston University, 2001

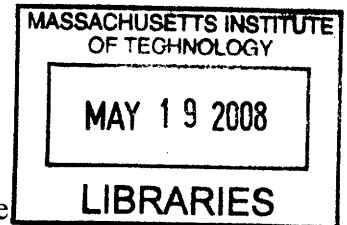
SUBMITTED TO THE DEPARTMENT OF COMPARATIVE MEDIA STUDIES IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN COMPARATIVE MEDIA STUDIES
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2008

© 2008 Michael Danziger. All rights reserved.

The author hereby grants to MIT permission to reproduce
and to distribute publicly paper and electronic
copies of this thesis document in whole or in part
in any medium now known or hereafter created.



ARCHIVES

Author: _____
Program in Comparative Media Studies
May 9, 2008

Certified by: _____
Nick Montfort
Assistant Professor of Digital Media
Thesis Supervisor

Accepted by: _____
Henry Jenkins
Peter de Florez Professor of Humanities
Professor of Comparative Media Studies and Literature
Co-Director, Comparative Media Studies

Information Visualization for the People

by

Michael Danziger

Submitted to the Program in Comparative Media Studies
on May 9, 2008 in Partial Fulfillment of the
Requirements for the Degree of Master of Science in
Comparative Media Studies

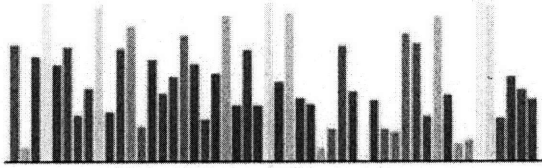
ABSTRACT

The design of information visualization, defined as the interactive, graphical presentation of data, is on the verge of a significant paradigm shift brought on by the continued maturation of the Information Age. Its traditional role as a scientific tool deployed by rigorous data analysts is in the process of expanding to include more mainstream uses and users, reflecting fundamental changes to the role of information and data in our increasingly digital society. However, visualization design theory remains rooted in earlier conceptions of its use, largely ignoring the needs of this new, non-expert audience. Accordingly, this thesis attempts to re-contextualize information visualization as a public-facing practice, and explores ways in which its design can shift from being described as “by experts, for experts” to a new characterization as “for the people.”

Thesis Supervisor: Nick Montfort
Title: Assistant Professor of Digital Media

ACKNOWLEDGEMENTS

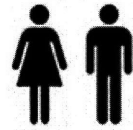
Many thanks to my thesis committee, Nick Montfort, Fernanda Viégas, and Martin Wattenberg, for helping me focus my thinking and getting me closer to the thesis I wanted to write.



information visualization



for



the people

INTRODUCTION	11
CONTEXTUALIZING INFORMATION VISUALIZATION FOR THE PEOPLE	17
THE STANDARD RATIONALE	17
HISTORICAL PERSPECTIVE	18
CUSTOMERS	20
MASS APPEAL	23
“INFOVIS AT THE EDGES”	24
DATA CULTURE: REVISITING THE STANDARD RATIONALE	27
THE DESIGN OF INFORMATION VISUALIZATION FOR THE PEOPLE	31
STANDARDIZATION: INFOVIS AS GRAPHICAL USER INTERFACE	33
INFOVIS DESIGN IN THE CONTEXT OF OUR DIGITAL EXPERIENCE	37
THEORETICAL AND METHODOLOGICAL DESIGN FRAMEWORKS FOR THE PEOPLE	41
<i>Semantics of Infovis Design</i>	43
<i>Affective Infovis Design</i>	53
<i>Social Infovis Design</i>	64
<i>Narrative Infovis Design</i>	69
CONCLUSION	73
LIST OF FIGURES	81
WORKS CITED	85

INTRODUCTION

The design of information visualization, defined as the interactive, graphical presentation of data, is on the verge of a significant paradigm shift brought on by the continued maturation of the Information Age. The ubiquity and centrality of digitization and networking technologies have radically increased our bandwidth for information exchange and consumption across all levels of society, resulting in the need for similarly scaled—and broadly accessible—methods of representation. While this type of information processing has historically been limited to scientific and industrial domains, we are now entering an era where these technologies are bringing quantitative data analysis to our living rooms, our breakfast tables, or anywhere else we can get access to the internet. Information visualization (or “infovis”), taking advantage of modern computer graphics technology, has established itself as an effective means of attending to large scale analysis in the former specialized, utilitarian context. However, despite the supposed intuitiveness of visualization, infovis design has struggled in its attempts to address the more mainstream, non-expert audience suggested by the latter one. The following analysis explores this transitional crisis in an attempt to help facilitate the shift from an information visualization design paradigm characterized as “by experts for experts” to one described as *information visualization for the people*.

Why information visualization for the people? If we accept that infovis is a valuable means of communicating information, why not? This is not a difficult claim to justify, as the visualization of information has a long history that predates the digital age. Canonical examples of effective quantitative data display range from William Playfair’s economic charts, to John Snow’s use of information graphics to resolve a cholera epidemic, to a culture of infographic use in the sciences and business. However, with a few notable exceptions, it has had less impact as a public-facing practice. While some types of visualization, such as cartography, have achieved social and cultural ubiquity, and basic charts and graphs have maintained a regular presence in the media and our daily lives, their value and use as information channels for the general public has not been clearly established.

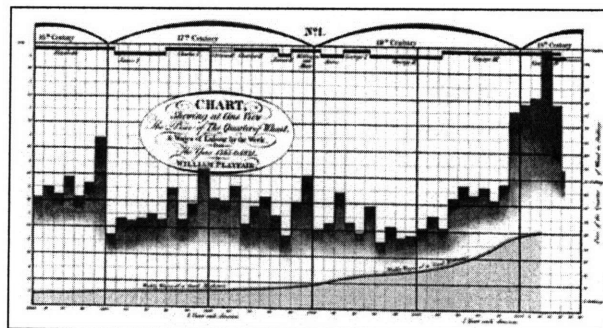


Figure 1. William Playfair’s 1786 line/bar chart depicting the price of wheat and wages in England.



Figure 2. Ross Perot’s use of infographics in his 1992 television ads.

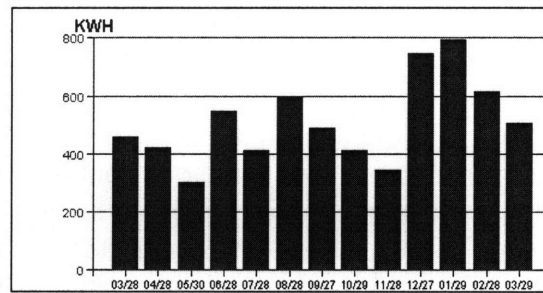


Figure 3. Histogram on my electricity bill from NSTAR.

This is arguably the result of a popular perception of statistical graphics that characterizes them as dull or personally irrelevant. The former concern, certainly well justified, is clearly a result of a design philosophy arising from utilitarian definitions of analysis and function that do not apply to a wider, non-expert audience. This philosophy will need to evolve as visualization becomes a greater part of an internet-enabled era. The latter concern is dissolving in the face of the Information Age, providing the space for evolution to occur.

The Information Age has (re)defined information visualization as a consequence of new technological affordances and a corresponding change in the way we relate to information. Computers with sophisticated graphics hardware are the norm, allowing for the processing and rendering of data on a greatly increased scale, and in real time. Interactivity is now integral to the display of data, providing a wealth of new techniques for exploring its content. Simultaneously, this new relationship with information is breaking down previous models of data consumption that conceptualize raw data as something only trained analysts and other experts deal with: First, the notion of a single, trusted information authority, such as the morning paper or the evening news, is giving way to a vast milieu of sources providing information across the internet. Gaining insight from this web of data becomes an act of aggregation and filtering, where we extract meaning from large-scale features as well as the individual details. Second, digitization is simply making (much) more of this information freely available – more than can be vetted by standard channels. Finally, new types of data are appearing that represent the artifacts of our social, cultural, economic, and scientific processes as we become more digital, which are also in need of representation. In other words, information itself is not just for experts anymore.

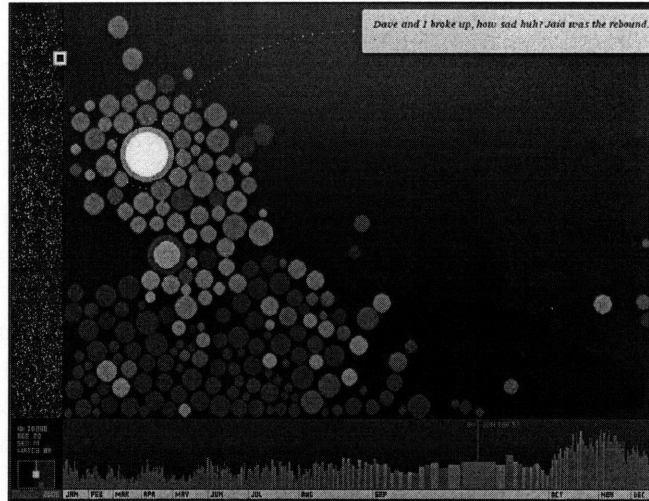


Figure 5. Screenshot from *The Dumpster*, an interactive “visualization of romantic breakups for 2005” that tracks blog posts related to romantic relationships (Golan Levin, et al.).

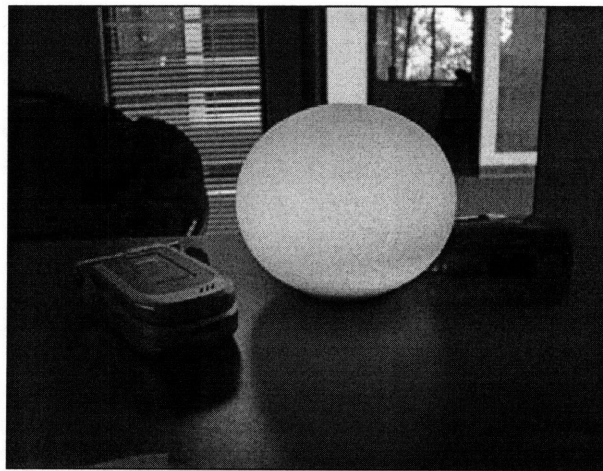


Figure 6. Photograph of the *Ambient Orb*, “a glass lamp that uses color to show weather forecasts, trends in the market, or the traffic on your homeward commute” (Ambient Devices).

The first part of this analysis, *Contextualizing Information Visualization for the People*, further motivates the definition of information visualization for the people by contextualizing the practice of infovis design both historically and in terms of the various communities engaged in it. It describes some of the theoretical and methodological concerns of these groups, from the scientific to the artistic, and positions their work within the broader spectrum of information visualization. Analysis of these concerns suggests the degree to which each group addresses a specific audience, which in turn helps define a new public audience that has so far remained underrepresented as a consumer of information visualization.

The second part, *The Design of Information Visualization for the People*, explores specific design paradigms relevant to public-facing infovis. Beginning with a comparative look at information visualization as graphical user interface that considers how public infovis design might be informed by the design of the ubiquitous desktop metaphor, this section is then devoted to identifying four design vectors that are relevant to the design of infovis systems for non-expert users. They are presented as broad,

overlapping approaches to design that revolve around a user-centered philosophy that characterizes information visualization for the people:

1. ***Semantic design*** addresses visualization as a visual language and sign system.
2. ***Aesthetic and affective design*** addresses the how the emotional impact of visualization design can impact its usability.
3. ***Social design*** addresses the intersection of infovis with social media.
4. ***Narrative design*** addresses the need for infovis systems to tell a coherent story.

Ultimately, while this document was written in an academic context, it is conceived as addressing a broad audience of interested parties, the range of which is represented by the groups to which it refers: academics, scientists, artists, designers, users, or any combination thereof. It has two goals in mind: First, to present a view of the current state of information visualization that both contextualizes its fragmented design space and motivates the need for a paradigm shift. Second, to suggest, based on a synthesis of existing research and the successes and failures of examples in the world, a framework for thinking about popular infovis design that better attends to the needs of non-expert users. Towards this end, it focuses quite specifically on recent developments in information visualization, particularly those that have been most public-facing. Accordingly, much of this activity has emerged in non-traditional infovis contexts—through the work of artists, designers, businesses, journalists, and the like—and represents both the diffusion of infovis research into the public sphere as well as vernacular design that stands entirely outside the scope formal academic infovis study. Therefore, it perhaps represents a people’s history of information visualization.

CONTEXTUALIZING INFORMATION VISUALIZATION FOR THE PEOPLE

The Standard Rationale

Any cursory survey of the discourse within and around the field (including this document) uncovers a deeply pervasive and optimistic rhetoric about the need for, and value of, information visualization in an increasingly information-overloaded digital age. The introductory paragraphs or statements of almost any infovis-related writing or presentation consistently feature what infovis researcher Jarke van Wijk identifies as “the standard rationale of our field,” (van Wijk, 2005) in which visualization is deployed against this information onslaught to help us efficiently make sense of and gain insight from the digital deluge. As far as rhetoric goes, it is a noble statement of purpose. Though its continual reiteration certainly serves the obvious need to justify the time, manpower, and research funds invested in the work that follows the introductions of these texts, it also serves a function that is at least somewhat novel in academic work, particularly in the “scientific” domains, which is to associate focused research goals with broader social value and implications outside of the lab. For the most part, it is from this moment of association that the fundamental argument of this thesis arises, which poses questions that ultimately revolve around the impact of information visualization in the “real” world. Accordingly, in the spirit of defining terms, it is important to unpack this rhetoric slightly in order to frame the discussion that follows.

Though instances of this “standard rationale” can appear with varying degrees of specificity, there is a generalized ambiguity about the actors and operations it describes. The “us” that gains insight and makes sense through visualization is underdetermined. So are the acts of “gaining insight” and “making sense” themselves. Certainly particular researchers and writers have particular ideas about how these terms are defined, and some even explicitly reference a mass audience (“InfoVis for the Masses,” 2007), but as it turns out, despite the socially progressive language of the “standard rationale,” much of current and past information visualization research has been based on implicit assumptions concerning who uses information visualization, and what it is used for. This section attempts to enumerate some of these assumptions, as well as motivate the particular assumptions and meanings that this thesis envisions.

As suggested above, when trying to articulate the “standard rationale” assumptions or the meaning of “information visualization for the masses,” we are really getting at two questions: who are “the masses” and what does information visualization “for” them look like? The first is largely a question of demographics; the second explicitly speaks to the scope of the research in the field, but also implicitly alludes to a conception of information visualization as a “mass medium” that “the masses” are engaged with.

Historical Perspective

To understand the changing landscape of information visualization users, it's worthwhile to briefly position the historical development and use of information visualization against that of infographics more generally. Although these two fields overlap, the generally accepted distinction between them is that infographics typically represent static visual displays of information (charts, graphs, tables, diagrams, etc.), and hence are not technologically dependent in their design or presentation. In contrast, information visualization typically refers to the use of interactive computer-driven systems to visually display information, focusing on dynamic representations that support more robust exploration of larger data sets. As such, the history of information visualization is obviously much shorter than that of infographics, being tied to technologies that have only seen major development in the last twenty years.

This technological distinction is important in that it frames a corresponding distinction in usage and presumed "audience" that characterizes the field today. While information visualization clearly has a historical antecedent in info-graphic design, the two have arguably diverged in terms of their uses. Looking at the canonical examples in the historical development of infographics, there is an established origin in scientific study: John Snow's use of visualization to determine the cause of a cholera outbreak in London in 1854, Mendeleev's creation of the periodic table of the elements in 1869, and many other examples that could be extended back to the scientific illustrations of Leonardo da Vinci, Galileo Galilei, or even Euclid. However, there is a parallel history of info-graphic development that was perhaps more public-facing: William Playfair, widely considered to have invented the bar chart in 1786, was a political economist interested in making economic data easier to comprehend, both for government officials and, presumably, a wider public audience. Similarly, Charles Joseph Minard, who created what Tufte has declared "the best statistical graphic ever drawn" (Tufte, 1983, p. 40) in 1861, was a civil engineer in France, developing graphics that "endeavored to popularize" the "great questions of public works" (Chevallier, 1871). Otto Neurath, an Austrian sociologist, developed the pictographic language *Isotype* with the goal of improving "cultural communication" and "democratizing cultural life" through visual representations of information whose meanings could be understood regardless of a viewer's cultural, social, or educational background (Neurath, 1937).

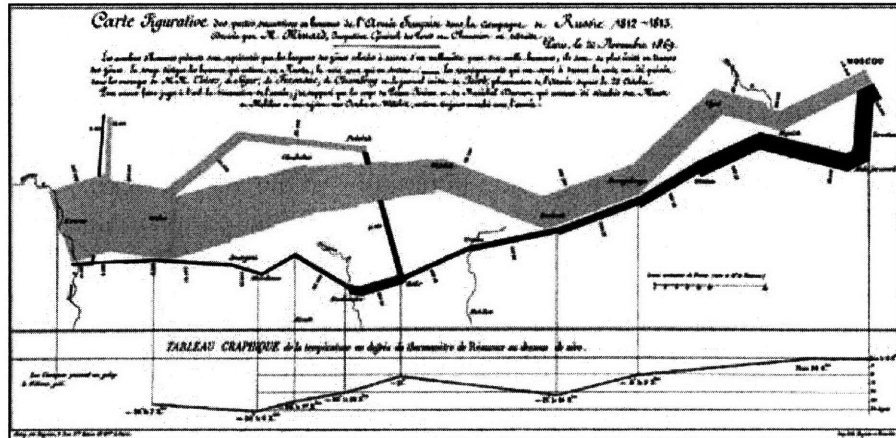


Figure 7. Charles Minard's infographic describing Napoleon's march to Russia.

However, despite this variety in info-graphic usage, the field of information visualization arose largely as a scientific tool. In response to the data analysis needs of the scientific community, and most likely empowered by the availability of computing technology to that community (long before similarly powerful computers became available on the consumer market) and the corresponding developments in computer science, the first forays in to information visualization (or, strictly speaking, “scientific visualization”) focused on supporting scientific research (Wattenberg, 2005). It is perhaps significant that, in addition to its focus as a tool for scientists, it was the technological demands of infovis that necessarily kept it out of the public eye (and consciousness); in the 80’s and early 90’s personal computers were comparatively underpowered, and the internet was barely in its infancy. It’s therefore not surprising that the design of infovis tools did not reflect any particular need to cater to non-expert users. Therefore, the argument could be made that despite a rather drastic change in the techno-social landscape since then, the current discourse within the field is still burdened by this inherited legacy.

Jumping ahead in the development of the field, the target audience for traditional visualization systems and tools has continued to be a largely scientific one. However, as a result of increased interest in “information” visualization techniques (as opposed to “scientific” visualization) that aim to address data of an arbitrary form, and undoubtedly coupled with the need to commoditize visualization research, it has expanded into other fields that are dependent on data analysis. Financial, medical, and military institutions involved in large-scale data analysis became the first wave of customers for the work being done in the field of information visualization. The outgrowth of infovis to this new community of users served to reinforce the prevailing notions of what infovis does, and who it serves, in two ways. First, it continued to support the definition of infovis as a utility to facilitate rigorous analytic work. Second, it characterized infovis users, to the researchers within the field, as not only “analysts,” but also as “customers.” Both of these notions have had a significant influence on the current state of information visualization and its users.

Customers

The bulk of the academic research being done in information visualization today focuses on the development of novel visualization techniques and the refinement of existing ones. While the majority of this research appears to remain “in the lab” or deployed under very specific research conditions, some of it has been spun off into commercial products and companies with varying degrees of success. Two of the most well known—having been initiated by two forefathers of infovis—are Spotfire and Tableau, both of which function as all-in-one visual data analysis packages. These are high-end software packages, featuring powerful visualization and analysis tools, and are aimed at the sorts of industrial users (or, more accurately, the organizations for which they work) described above. Indicative of how these companies perceive their users, these products are priced as enterprise-level software, putting them well beyond the financial reach of most individuals (ironically, Spotfire’s tagline is “Analytics for Everyone”) (TIBCO Spotfire, 2008). While lower-priced infovis software certainly exists (at some level, Microsoft Excel can be considered an infovis tool), packages such as Spotfire and Tableau undoubtedly represent the definitive “face” of information visualization to the customers they are seeking, or are at least characteristic of “serious” infovis software on the market. However, academic researchers in need of visualization tools and the market demographic that can afford enterprise-level software are about as “mass” as their audience gets.

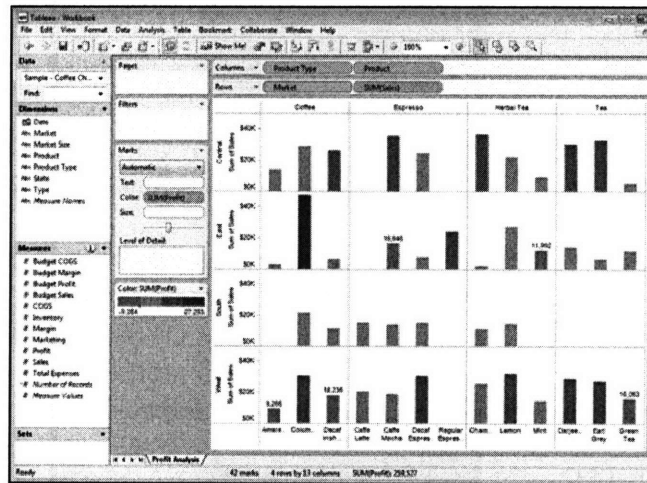


Figure 8. Screenshot from Tableau (Tableau Software).

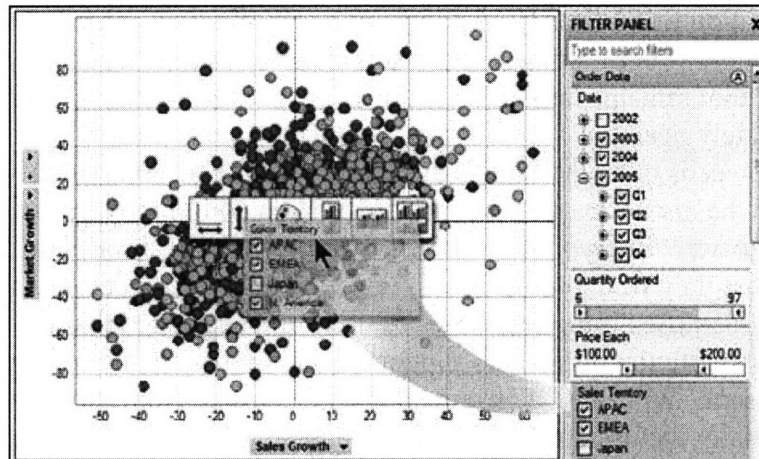


Figure 9. Screenshot from Spotfire (TIBCO).

Interestingly, there is recognition within the field that this is an issue. Some infovis researchers are beginning to become more introspective, expressing concern over the applications of current infovis research. As when any research field matures, the “low-hanging fruit” has been grabbed, resulting in more specialization and incremental advancements. In what turns out to be a problematic counter-force to the “standard rationale,” some fear infovis research has become too insulated from the information needs of outside users, and that new techniques are being developed on a trajectory characterized by an increasing complexity that takes them further away from what most of their users find tractable. Van Wijk cites a “growing gap between the research community and its prospective users,” and points out that “many [...] new methods are not used in real-world situations” (2005). Shneiderman repeatedly encourages researchers to connect with end users so that new techniques can make it to market (2007). Bill Lorenson presents a more ominous view in his paper titled “On the Death of Visualization”:

[...] If we don’t adapt and change, the field (and [the Visualization Conference]) will disappear. Our customers have disappeared. In some cases, they don’t need us because they can purchase robust, well-engineered instances of our techniques. In other cases, we have not been responsive to their requirements, mainly because we don’t understand their problems. Certainly, there are still some customers still present at the conference, namely the National Labs. But even there, it’s the computer scientists, not the scientists that attend Visualization.

The evident problem, which will lead us toward the definition of “information visualization for the people” suggested in this thesis, is that despite concerns over a lack of research applications, the scope of the target audience conceived of by the infovis community remains the same; they are apparently unwilling or unable to expand their notion of “prospective users” beyond the “customers” they currently serve.

This isn’t for any lack of interest in expanding the reach of their research. As suggested at the beginning of this section, the field often explicitly encourages itself to address the masses. At the 2007 InfoVis Conference, infovis consultant Stephen Few gave the

capstone presentation titled “InfoVis as Seen by the World Out There.” In it, Few delineated a disconnect between the work being done by infovis researchers and “the world out there” that remains largely unaware of what information visualization can do (despite “desperately needing” it). Pointing out that the vast majority of infovis research is focused on the needs of a small minority of users, he asked “Who makes up this world out there, those who are not part of the infovis research community but stand to gain from what we have to offer? Everyone who needs to make sense of information and to communicate what they find to others in an effort to make better decisions” (2007). Accordingly, he identified three categories of users and suggested that the infovis community redistribute their efforts to support them proportionally to the number of users the groups represent. Working backwards, Few’s groups go from “sophisticated data analysts,” essentially the users referred to here thus far – highly trained analysts doing robust analysis on large data sets – to “informal information consumers and presenters.” He characterized the latter as business workers that are required to interact with information, often visually, in a regular but limited way.

Few’s distinction attempts to broaden the potential audience for visualization research but simultaneously reinforces a long-held assumption about visualization audiences. He envisions them all as “information workers” of some sort within particular domains – people whose livelihood revolves around the explicit interaction with particular data at some level, who are required to make sense of that data in order to do their jobs, and who ostensibly have some domain expertise. For information visualization to be more widely applicable, this is a problematic assumption. It not only excludes the vast majority of the population who are not information workers, but also calls in to question the defining tenets of information visualization. The standard rationale correctly suggests that we are entering in to an age of information inundation. But again, “we” are underdetermined. Where the field of information visualization has stumbled in its progressive rhetoric is in not being progressive enough.

This thesis agrees with the necessity of visualization as a means to attend to our increasing information needs, but disagrees with the scope of those needs. Of course data analysts will always need more effective tools, but the significance of the Information Age is in the degree to which it will impact “users” at all levels of society. It’s not just information workers that need to navigate this flood, it’s just about everyone. More importantly, the nature of that navigation will vary wildly; some will need to perform “mission critical” analysis on which to base decisions, while some will just want to make sense of it. For some it will be a means to a paycheck, while for others it will simply be an opportunity for informal learning or entertainment. In principle, information visualization can help mediate all these types of interactions. However, as will be examined later, the current prevailing assumption concerning infovis users begets certain visual design theories and methodologies that may not be appropriate for wider audiences. These wider audiences are the real “masses,” and reaching them with infovis will require some broader consideration.

Mass Appeal

To better understand these “masses” as potential infovis users, we can return to information visualization’s historical antecedent, the infographic. As suggested earlier, infographics have some historical precedent as being public-facing visual information displays that address audiences of the scale described here. Following from the likes of Playfair, Minard, and Neurath, infographics have been deployed across a variety of contexts to convey quantitative information to non-expert audiences. Perhaps the most “mass”-reaching example is the newspaper graphic.

While the history of newspaper infographics is arguably as long as the history of newspapers themselves, they have seen a dramatic increase in popularity since the late 70’s and early 80’s and now appear as regular features in most major newspapers. This popularity can be traced in part to improvements in color printing techniques and the rise of desktop graphic design and publishing software (Adobe Illustrator, etc.), but also to an increased need to present quantitative information clearly and quickly to the public (or, more dubiously, a need to increase the visual appeal of the paper itself). Despite differences in design methodology, scope and usage, as well as a history of criticism of those methodologies (see Tufte), these infographics represent a type of visual communication deployed expressly to address a non-expert audience. Given that newspapers are about as “mass media” as you can get, they suggest themselves as a model for infovis design that addresses “the masses.” It is also important to point out that despite their earlier characterization as an antecedent, the practice of newspaper infographic design has continued to develop in parallel to that of information visualization. These parallel and largely independent development tracks present an interesting contrast in design and reception that becomes particularly relevant as the technologies and guiding principles (i.e. the standard rationales) of each practice begin to converge. The digitization of newspapers, and their movement online, brings the design of their infographics explicitly in to the space of infovis, which makes it all the more possible for the two fields to benefit from each other’s strengths.

Without going in to too much detail here concerning the particulars of their design it is helpful to briefly qualify their usage against that of traditional infovis in order to better understand their relevance, but also their limits, as a model for infovis design. The primary difference between the two, despite the technological basis for the proliferation of newspaper infographics described above, is in their dependency on digital technologies at the level of both design and presentation. Ignoring the latest developments in online newspaper infographics for a moment, they represent a static, non-interactive display of information. This is in obvious distinction to infovis, which is almost always interactive or animated in some way. Another significant difference is in scope; newspaper infographics are typically more focused in their approach to quantitative data. For instance, the dimensionality of data presented in a newspaper graphic is usually much lower than that of typical infovis, which often prides itself on encoding a large number of variables in a single display. Similarly, the number of detailed data points emphasized in newspaper graphics tends to be relatively small (this is distinct from the size of the data sets being represented, which can of course be arbitrarily large; it essentially refers to the

amount of space that is given to “highlighting” individual data points). On the other hand, the technological affordances of infovis systems allow for “details-on-demand,” where users can drill down from an overview of the data to the details of specific data points (by clicking on various parts of the display, for example). The final difference, which perhaps encompasses those described so far, is one of usage goals: In information visualization, visualization systems are sometimes characterized by whether they are *exploratory* or *presentational*. Exploratory visualizations are used to discover unknown trends in the data, while presentational ones are used to convey its known characteristics. While this distinction is arguably becoming less meaningful as infovis matures (careful interface design considerations allow for a single visualization tool to support both types of usage, as seen in IBM’s Many Eyes system; see *Social Infovis Design* in the next chapter), it does seem relevant with respect to newspaper graphics, if only by virtue of their static presentation.

Despite these differences, however, the user (or in this case, reader) centered design paradigms employed in the production of newspaper infographics are worth analyzing. While there hasn’t been as much explicit study of visual representation in newspaper graphics as there has in infovis, their vernacular methodologies, undoubtedly influenced by the methodologies of journalism in general, have achieved a measure of success in reaching non-expert audiences. As will be discussed later, an emphasis on story-telling in infographic design, the use of widely recognizable signs, symbols and icons, the integration of explanatory text within visuals, and even the addition of chartjunk and sub-optimal data-to-ink ratios (both references to elements of Tufte’s guidelines) help to engage the non-expert reader. Some of these approaches have been largely ignored in information visualization design; some actually contradict its best practices. Nevertheless, any attempt to design towards a mass audience should take them seriously.

Similarly, though not treated separately here, we might look to the design principles employed in the development of infographics, animations, and visualizations that are more expressly educational. Visual representations of information in educational software, museum displays, and other informal learning contexts also serve a broad audience and attempt to convey insight with no assumptions of prior knowledge on behalf of their users.

“Infovis at the Edges”

Another context in which information visualization addresses a broader audience is through what Zachary Pousman, John Stasko, and Michael Mateas identify as the “edge cases” of infovis (2007). These edge cases represent applications that share some defining characteristics of traditional information visualization, but are generally not considered part of the academic discourse in the field. Regardless, they represent an important development toward a more broadly defined domain of infovis.

The characterization of infovis as an outgrowth of computer science established here (referred to largely as “traditional” infovis) reflects its presence as an academic research field. Most visualization research is published within the domains of the ACM and/or IEEE, which, despite the suggestion that it is a truly interdisciplinary field, again reinforces the notion that it is first and foremost a scientific discipline. However, the increasing availability of visualization development tools (or more general purpose technologies suitable for creating interactive visual displays, such as Java, Flash, Director, etc.), as well as the increasing availability of data to visualize, has resulted in a proliferation visualization work by—and for—people outside of the academic infovis scene. Designers, new media artists, “mashers,” and data-philes of various backgrounds have brought their own sensibilities to bear on the practice of visualizing information.

The types of visualization that this group has produced have been categorized in various ways by Pousman (et al.) and others, but ultimately tend to be classified along a spectrum that places information visualization (in the traditional sense) at one end, and “information art” on the other. Where “information visualization” refers to systems designed to support rigorous visual data analysis, “information art” refers to artworks that are in some way derived from, or based on data, though do not necessarily maintain an explicit or well-defined mapping from the data to its visual representation; as “impressionistic” approaches to infovis they typically don’t support particularly rigorous analysis of the data they represent. Instead, they serve a variety of alternative functions, from the purely aesthetic, such as Alex Dragulescu’s *Malwarez* (visualizations of computer viruses, <http://www.sq.ro/malwarez.php>), to the rhetorical, such as Josh On’s *They Rule* (a visualization of relationships between board members of Fortune 500 companies, <http://www.theyrule.net>). The less populated space in between these extremes is made up of various other classifications, including “ambient visualization,” “information aesthetics,” and “casual infovis,” which often share many similarities in scope and focus to traditional infovis systems. However, the particular details of these distinctions are unimportant to this argument except insofar as they all represent flavors of information visualization that have received little attention within the academic infovis community.



Figure 10. *Malwarez* visualization of the *MyDoom* computer virus (Alex Dragulescu).

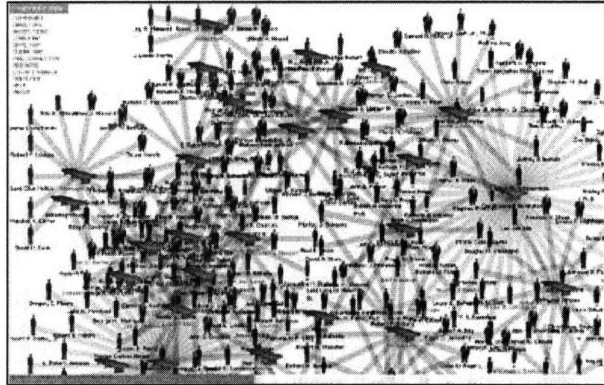


Figure 11. Screenshot from *They Rule* (Josh On).

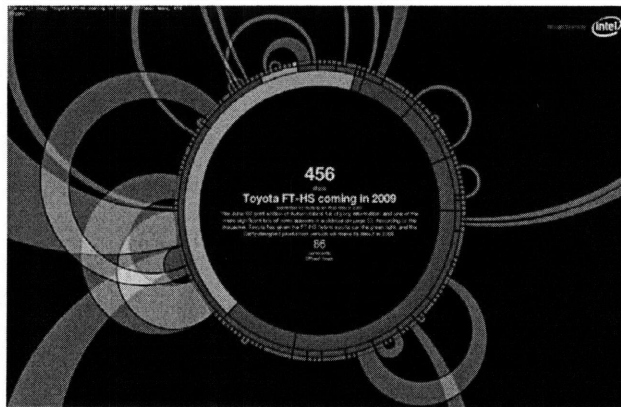


Figure 12. Screenshot from *Digg Arc* (Stamen Design).

Nevertheless, they do see more exposure to a general audience. Though certainly not widespread, they are often deployed on the internet or in high-traffic physical spaces (galleries, museums, etc.), sometimes in association with popular websites, and most significantly, almost always free to use. Stamen Design, for example, has produced a series of aesthetic visualizations for popular websites Digg and Twitter, and YouTube has recently begun visualizing some of its content. Taking advantage of new Web 2.0-related affordances, many individual designers have produced attention-grabbing infovis based on publicly available data from sites like Google or Amazon. Through internet word-of-mouth, the novelty of these visualizations regularly generates a fair amount of buzz and interest on the web. In doing so, they have established another discontinuity between the vernacular infovis design practices that often appeal to a wide audience (it's no stretch to suggest that such visualizations would not receive as much attention if they were built in *Excel*) and the more formal ones characteristic of traditional infovis. This has generally muddied the public consensus on what information visualization is all about. Is it an aesthetic artifact to be admired as such? Is it a means of doing data analysis or gaining analytic insight? Is it something in between? This public exposure has shown how infovis design can meaningfully move beyond traditional types of data representation, but has also raised concerns, such as those voiced by Few, that the traditionally-defined analytic usefulness of infovis is being misrepresented and devalued in the public eye.

Regarding this point, the position taken in this discussion is that in order for information visualization to become a “mass medium,” its value as a useful information processing

aid must be emphasized. More artistic explorations of the medium are of course valuable, but visualization implementations characterized as “beautiful but useless” will be just as ineffective at reaching a broad audience as overly technical, expert-level ones. However, this perspective does not discount the valuable and engaging design methodologies often deployed in these “edge cases.” In fact, it argues that these design principles will become necessary to the production of information visualization that is accessible and engaging to a broad audience. Furthermore, balancing the useful/beautiful dichotomy is by no means a zero sum game; this discussion suggests that it is possible to have both through the application of principles synthesized from the entire range of visualization design theory. Finally, and perhaps most importantly, as Pousman (et al.) suggests, it is essential that the expanded uses of infovis be taken seriously as applications of information visualization. The “information explosion” referenced by infovis’ standard rationale will include “important” statistical data that requires serious analysis, but it will also include more mundane information streams (the output of social networks, consumer information, personal data of various types, etc.) that general users will want to digest. Information visualization has proved its value with regard to the former; it should now be purposed towards the wider information needs emerging in the Information Age.

Data Culture: Revisiting the Standard Rationale

To further motivate the need to consider “information visualization for the people,” it makes sense to briefly return to the rhetoric of the standard rationale and attempt to qualify it in such a way that it better reflects the Information Age as a moment of social and cultural change. The picture it paints in its original formulation is one of impending information inundation, but it is perhaps rooted in a perception of this inundation as it was understood in the early days of information visualization development. Granted, this wasn’t long ago –15 or 20 years ago at most– but the information landscape has changed drastically since then. Without resorting to too many Information Age clichés, few would have predicted, even that recently, the extent to which our relationship to information would change at all levels of society. The internet has certainly fulfilled its destiny as an “information super-highway,” but developments in the last few years in particular speak emphatically to a need for more efficient, scalable means of attending to the threat of information overload for all users. In other words, today’s information needs go far beyond just scientific or business-related; information is everywhere.

With the advent of the technologies and practices loosely described as “Web 2.0,” we are beginning to witness the commoditization of information (or data) in a very literal sense. Organizations that used to simply operate web sites or services are now in the business of gathering, storing, and sharing information; Google, of course, continues to index the contents of the web, but other groups, such as the various social networks (including YouTube, Flickr, and the like), the media networks, the blogosphere, the e-commerce sector, and governmental organizations (among many, many more) now represent nodes around which information is constantly accreting. Simultaneously, many of these

organizations are beginning to open their databases directly to the public, offering programmatic access to information that used to be available solely through their websites, if at all. Along with the development of software tools that allow for easy access to this data (Flash, Java), this presents a whole new way to navigate the super-highway. One example of this is the recent popularity of the “mashup” phenomenon, where intrepid users develop novel interfaces to these databases, often combining different data sets from different providers to produce new perspectives on the information. Such mashups have addressed a diverse range of information, from urban crime statistics to popular music purchases on Amazon. There is evidence of increased interest in explicitly statistical data as well: Sites like Swivel and Many Eyes encourage users to upload their own data sets, and provide tools with which to analyze them. Recently, Yahoo! And Google have released “charting APIs” that allow users to dynamically generate charts and graphs that can be embedded in a web page, using whatever data they like.

Perhaps more telling of a popular willingness to embrace data culture is the playful appropriation of information visualization as a means to express the “statistics” of popular culture. A variety of blog sites such as *GraphJam* (<http://www.graphjam.com>) aggregate creative renditions of song lyrics and other pop culture references in infographic form:

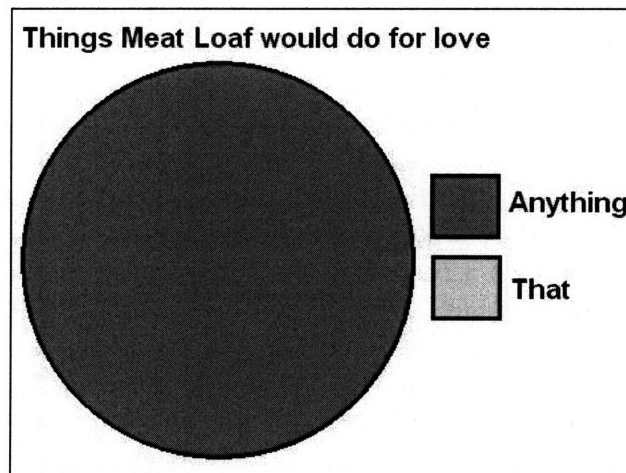


Figure 13. *Things Meat Loaf would do for love* (original designer unknown).

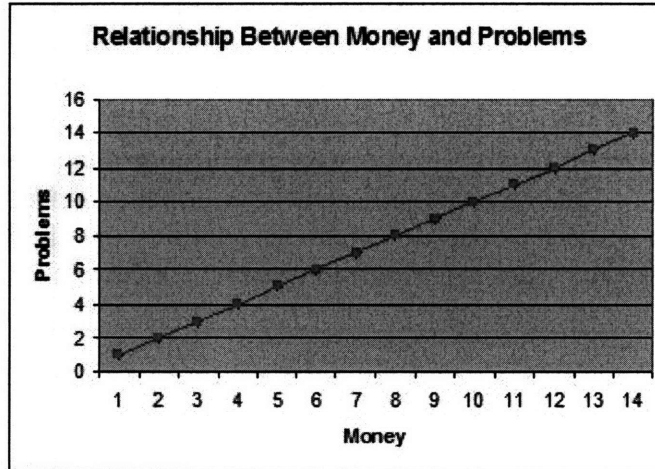


Figure 14. *Relationship Between Money and Problems* (original designer unknown).

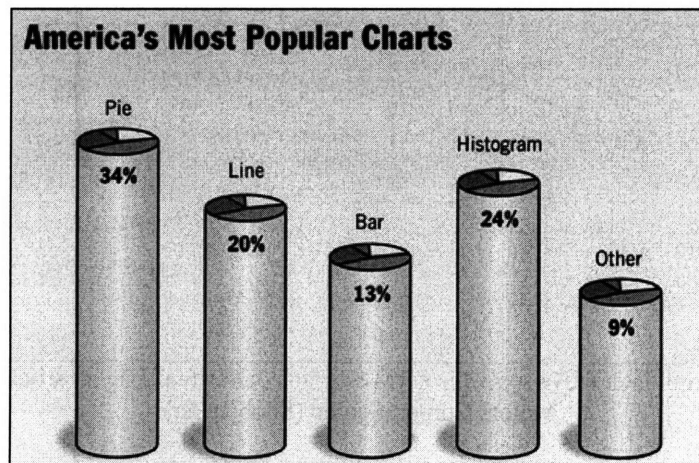


Figure 15. *America's Most Popular Charts* (The Onion).

Thinking optimistically, what is most important about this new “data culture” is its social and cultural impact. While an information deluge has the potential to be paralyzing, it also promises to change the way we think about the world. The increased availability of information of any kind not only makes us “smarter,” (if only in the sense of being able to Google an unfamiliar topic and learn something about it in five minutes), but also promotes a relationship with ideas or facts or opinions that is grounded in context. The trivial example of this is the way in which we can now use the aggregated ratings and reviews from thousands of users to help select what products to buy, rather than relying solely on advertisement or “expert” opinion. A more socially and culturally significant example can be seen in Ethan Zuckerman’s *Global Attention Profiles*, (2003) where he has analyzed, via automated web searches, the degree to which each country in the world receives media attention, presenting the results in an easily understandable (and visual!) form, and giving context to our media consumption. Bold new approaches to news aggregation, as exemplified by sites like *Silobreaker* (<http://www.silobreaker.com>), take this idea to the next level, explicitly contextualizing news items within a rich semantic network, using visualization as means to explore contextual connections. Another example on a smaller scale involves Switzerland’s *Smartvote* system, where visualization

techniques are used to present voters with a “map” of potential candidates to vote for, based on the degree to which their political views match those of the available candidates. There are many more such examples, but the point is that the increasing availability of such data can give us a much broader view of the world, leaving us less reliant on the narrower information streams we are accustomed to (expert opinion, broadcast media, etc.). The broader significance of information visualization in this context is the way in which it can help mediate this relationship with data; the availability of information is irrelevant without a means to interpret it effectively.

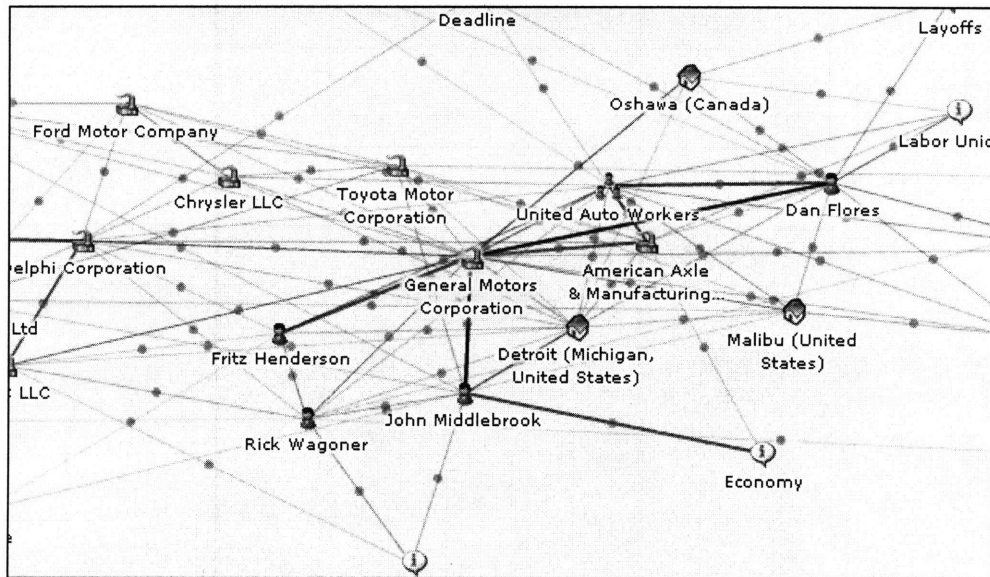


Figure 16. Detail from interactive semantic network graph contextualizing an article about the General Motors Corporation on (Silobreaker).

The notion of using infovis to mediate our interactions with information also implicitly recognizes infovis as a medium. This is an important understanding, as it then allows for all of the subtleties of use that characterize other media, like text, film, or the internet. This includes an acceptance of visualizations as rhetorical objects, and the additional power that affords them, or simply as venues for expression. In other words, if we can become as free and comfortable in our use of visualization as we are in writing, its effectiveness is bound to increase. Media theorist Lisa Gitelman, channeling Marshall McLuhan, points out that “as critics have long noted, the success of all media depends at some level on inattention or ‘blindness’ to the media technologies themselves (and all of their supporting protocols) in favor of attention to the phenomena that they represent [...]” (2007). This is the state information visualization should strive to achieve. It therefore seems inadvisable to enforce particular types of use on it; rather, it should be treated a platform for experimentation.

THE DESIGN OF INFORMATION VISUALIZATION FOR THE PEOPLE

The practice of information visualization, whether in the traditional sense or otherwise, draws on a wide variety of disciplines in its construction of a working—but often conflicting—set of theories and methodologies that define its best practices. While this is undoubtedly its greatest strength as a communicative medium, it is also potentially its greatest weakness with respect to achieving widespread popular adoption and use, particularly when these theories and methods are not invoked consistently. In other words, the choices designers make when producing visualizations beget visual languages, and the form of these languages, just like written ones, are largely defined by the motivations behind the design choices. This is true at both the “structural” level (used, in this case, to refer to the multitude of visual data encodings—bar graphs, scatter plots, treemaps, etc.) as well as the “presentation” or “interface” level (used here to mean the “look and feel” of a visualization). The end result is a wide range of visual information display languages, each derived from some mixture of theoretical best practices and ostensibly targeted towards a particular audience. A survey of the visualization examples aggregated at *visualcomplexity* (<http://www.visualcomplexity.com>), for instance, bears this out. This site focuses on examples of network graphs, and yet there is a huge variety in their visual presentation. No two look exactly alike, and some look nothing alike, despite being based on similar data structures.

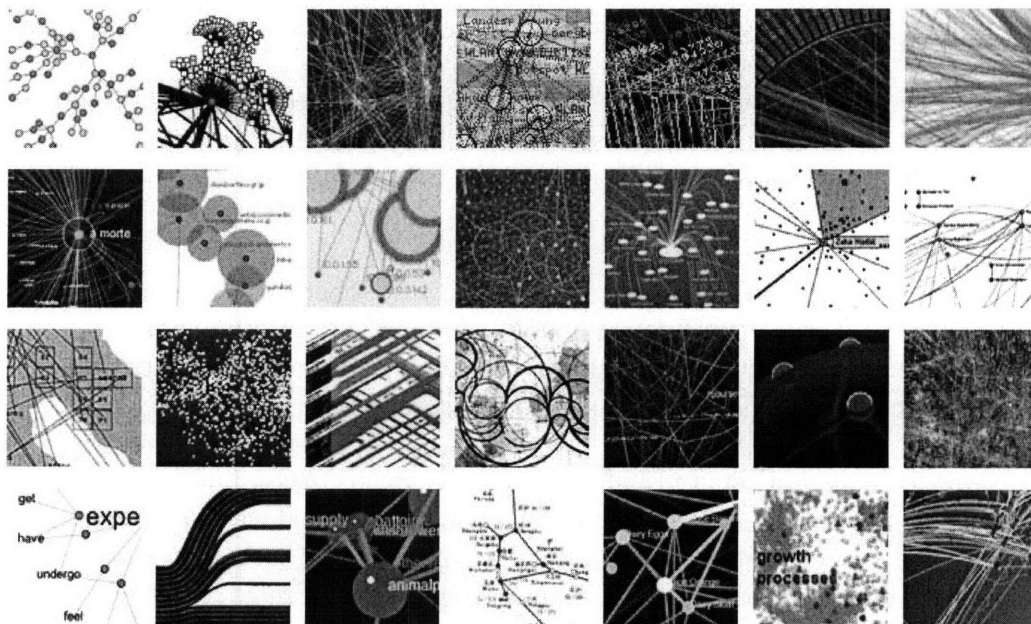


Figure 17. Details from network visualizations listed at *visualcomplexity.com*.

However, while experimentation and diversity in infovis design are generally features to be applauded and encouraged, they can become problematic when trying to establish infovis as a usable communicative medium for a general audience. Continuing the

analogy to written language, communicating to this general audience requires a consistent language with familiar structure and syntax. Fundamentally, learning to read is impossible if this structure is constantly in flux, or if the definitions of its words are arbitrarily subject to change. At a higher level, writing is only easily comprehensible to a reader when it is written in a style appropriate to their understanding of the language. So, just as overly academic, simplistic, technical, or poetic prose might confound or disinterest the general reader, the same principle applies to the visual languages and encodings of information visualization. Appropriately, Manuel Lima of *visualcomplexity* invokes a quotation from Matt Woolman's 2002 book, *Digital Information Graphics*, on his site:

Functional visualizations are more than innovative statistical analyses and computational algorithms. They must make sense to the user and require a visual language system that uses colour, shape, line, hierarchy and composition to communicate clearly and appropriately, much like the alphabetic and character-based languages used worldwide between humans.

Of course, the meaning of this statement depends on the definition of the term “functional,” which will be addressed shortly, but it references an understanding of visual language that, while long studied in more traditional applications of the visual (the fine arts and graphic design, for instance), has not had as much consideration within the relatively young field of information visualization despite being extremely relevant, particularly when considering its potential mainstream usefulness. This is presumably another reflection of its origins as a scientific practice with a narrow disciplinary focus.

It is interesting to note, however, that there are historical precedents to the discussion of visual language in information displays. A particularly compelling example comes from Otto Neurath, the Austrian sociologist responsible for inventing the *Isotype* pictographic language in 1937:

The usual visual methods—even the most careful charts and the most elaborate exhibits—are frequently confusing rather than enlightening, because their elements are unfamiliar. It is almost as though people had to learn a new language for each new communication.

Neurath was interested in using visual language as a means to communicate across (verbal) language barriers, as well as cultural, economic, and educational ones. He argued the need for “consistent visualization,” suggesting that such a language could “humanize and democratize the world of knowledge of and of intellectual activity.” (Neurath, 1937). *Isotype* itself became the precursor to the now universally familiar iconographic language used to identify locations of services (such as in airports), serve as warning signs, and, most recognizably, tell us where to look for restrooms.

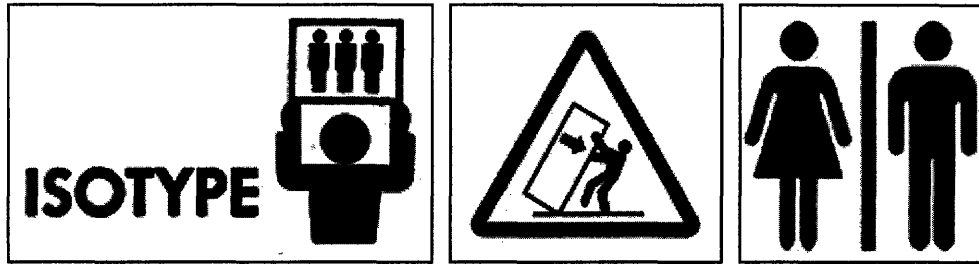


Figure 18. *Isotype* and its derivatives.

Neurath's reflections on *Isotype*, though over 70 years old, serve as a conceptual model for how infovis could position itself in today's far more information-intensive society. Like his vision for *Isotype*, information visualization for the people should be accessible to anyone, constituting a literacy as basic, or more so, than the ability to read and write. However, the degree to which the visual standardization he advocated, particularly at the "presentation" level, is the appropriate way to achieve this is still debatable. The following section takes a closer look at this question.

Standardization: Infovis as Graphical User Interface

As exemplified by *Isotype* and many other design philosophies across a wide spectrum of domains, one of the most straightforward ways to address the issue of usability is through standardization. In the case of interface design, standardized displays attempt to invoke a similar layout and look-and-feel across similar applications so as to promote familiarity and transferability of user expertise. On the surface, to the degree that information visualization can be standardized in any broad sense, this seems like a compelling way to foster literacy for its design. However, a look at the impact of standardization on other software interfaces complicates this notion. Does standardization serve the goal of popularizing information visualization in the long term, or will it ultimately hinder the development of the field in this direction?

There are obvious precedents for constraining design to address non-expert users, particularly with regard to software. We don't have to look any further than the operating systems of our computers to see this principle in action. Both the look-and-feel and underlying functionality of the Windows and Macintosh operating systems have remained largely unchanged since their inceptions, at least as far as the average user is concerned. Arguably, providing a consistent "desktop" experience across multiple applications is one of the fundamental purposes of a GUI-based OS. A consistent interaction paradigm (the mouse-based windowing system), a host of familiar application-agnostic functions (open, save, close, cut, copy, and paste, etc.), and a standard layout and visual style (the "look and feel" of buttons and other interface widgets) serve to mitigate confusion when performing the multitude of different tasks supported by the operating system. When new applications are introduced to the user, this standardization provides a framework from which to understand them; if we're not

familiar with their particular functionality, we at least have an entry point, even if it is as simple as knowing how to access its help files, or where the “preferences” page is located.



Figure 19. Evolution of the Apple Mac OS and Microsoft Windows graphical user interfaces.

In other words, the 20+ year persistence of the desktop metaphor in the two major graphical operating systems and its corresponding visual design certainly speak to the relevancy of familiarity and the power of a well-defined interaction model. Historically, this design provided a consistent, over-arching framework that defined personal computing to a public that was just beginning to embrace the notion of computer interaction as a part of their daily lives, and successfully (by most considerations) ushered us in to the computing age. Therefore, ignoring the obvious limitations of a comparison—the desktop GUI represents a specific type of interface to a specific type of information (files and applications on a computer), while information visualization refers broadly to the practice of producing visual representations of information in general—can the design of user-friendly operating systems inform the design of user-friendly infovis systems? Assuming that the variety of visual encodings employed in infovis implementations could be mapped to a standardized design format analogous to the look-and-feel of applications running on a specific operating system, would this be a productive direction for infovis design to take? Given that an operating system GUI is an early example of information visualization itself, it makes sense that the design of public-facing infovis could follow from its most ubiquitous implementation.

There are certainly compelling reasons to think so, the most obvious of which being the aforementioned familiarity it would facilitate. However, there are some distinct differences between the contexts in which operating system GUIs and more generalized infovis interfaces have developed that potentially complicate the issue. It is also important to first be aware of the primary limitation of the comparison: as will be discussed in the next chapter, one of the fundamental issues facing modern infovis design involves the degree to which the semantics of different fundamental visual constructions are unfamiliar to most users. This was earlier referred to as the “structural” layer of design, which distinguishes one type of visual construction from another (i.e. bar graph, scatterplot, network graph, treemap, etc.), and arguably represents the core of what traditional information visualization is about. As such, this is probably an element of infovis design that standardization cannot directly influence; for instance, no amount of standardization can transform a confusing network diagram into a well understood histogram – these are fundamentally different structures. However, standardization of information visualization techniques might dictate which structure should be used for a specific type of data. It also certainly has influence over the “presentation” layer, which refers to the way in which these fundamental structures are packaged for the user. By this logic, we might read the desktop metaphor as a standardized choice of iconography and operations used to represent a computer’s files, applications, and processes (largely characterized by the so-called WIMP paradigm: Windows, Icons, Menus, and Pointers).

It’s also important to recognize the differences in usage patterns between the OS GUI and more typical information visualization. The former underlies all the operations we execute on our local computers, thus representing a significant investment of time by its users, and, for the purposes of this discussion, is offered in essentially two flavors (Microsoft or Apple). The latter is usually encountered in more incidental, short, and discrete episodes, and could be supplied by a variety of vendors or other content providers. This has a significant impact on the design principles underlying their production. However, practically speaking, it also renders the discussion somewhat hypothetical, since the reality of standardized infovis design across a multitude of providers is admittedly a far-fetched one.

In any case, the experience of non-expert users acclimating to the desktop metaphor remains an interesting model for popular infovis design. The visual desktop metaphor entered the public consciousness at the same time as the personal computer itself; in effect, it represented the only type of computer interaction most users had ever known. Everything about the computing experience was novel, including the designers’ understanding of how humans and computers interact, which uniquely positioned this particular paradigm to become the de facto standard for public-facing computing. Whether motivated by a genuine desire to foster computer literacy, technical limitations, the inherent quality of the desktop metaphor, or by the need to expand market reach (offering a similar product to your competitor reduces potential barriers dissuading their users to switch to your product), this approach successfully navigated the early phase of personal computer adoption by avoiding the confusion resulting from a market saturated with a variety of different interface formats. Initially, this seems like the sort of course

that would be appropriate for the goal of encouraging public acclimation to visualization, but it ignores over 20 years of technological development and corresponding advances in our understanding of user-centered design. Despite having not yet achieved mainstream appreciation, the infovis design space today is arguably much more developed than that of GUI design 20 years ago. Technological advancements alone allow for the exploration of exotic visual constructions that weren't manageable on the comparatively meager hardware of the 1980s (the massive increase in display resolutions and the development of 3D graphics capabilities are two representative ones), and software interface design theory has evolved accordingly. The result is a much richer palette of design options to draw from; ignoring these options feels like a step backwards in the theory of interface design.

It also points to one of the more problematic issues of standardization that is particularly evident in GUI design: the tendency for early design constraints to perpetuate even when the conditions of constraint no longer apply – the desire for familiarity and “backwards compatibility” becomes a security blanket that inhibits further design innovation. In the case of our operating systems, the familiar desktop metaphor on which all of modern computing was built has become so central to our conception of the user experience that it is now exceedingly difficult to supplant it, even by newer versions of the same operating system: each version of Microsoft Windows released after Windows 95, despite being released in to an increasingly sophisticated technological landscape, offers the option to revert to a “Classic” interface that reproduces the familiar look of Windows 95, presumably to cater to users uncomfortable with even the slight updates to the standard desktop that these new versions provide. Interfaces that attempt a more significant break with the traditional metaphor are relegated to entirely different platforms: the interfaces designed for the iPhone and OLPC, for example, are only considered viable precisely because they are not running on traditional personal computers (one is a handheld device and the other is a laptop targeted towards users who have never used a computer before). Meanwhile, as the images in Figure 19 indicate, the standard interfaces remain tied down to “legacy” paradigms – though it is interesting to note how the design of their latest versions takes on an aesthetic dimension once their functionality is normalized. If the general lack of traction that newer alternatives to the desktop metaphor have achieved is any indicator, this is likely not a design cul-de-sac that infovis wants to find itself trapped in. If anything, modern infovis is already battling against a public perception of the field as characterized by charts and graphs of the type generated by popular standardized software such as Microsoft Excel.

Standardization may also inhibit the internalization of our interactions with information, both in the abstract McLuhan-esque sense, and more literally. McLuhan's argument, applied to software interfaces, is that the formality of a standard interface reinforces the presence of the “medium” through which information is conveyed, which at some level prevents us from experiencing it as pure content. More literally, as has been shown in the case of the desktop metaphor, standardized design runs the risk of reinforcing the standard rather than the information model it represents. A quintessential example of this would be the user who is only able to navigate the system through recognition of its symbols themselves, rather than the underlying operations or objects they signify (i.e. the

non-expert desktop user who understands internet access simply as a lowercase blue “e,” and is lost in situations when that symbol isn’t present). Despite all the concern over the visual design of information visualization, its value to the world is ultimately tied to the information it represents, rather than its own formal properties. In other words, successful infovis should be about experiencing the nuances of the information rather than those of the visualization. Towards this end, is it more valuable to pursue varied visualization design that speaks to us in terms we understand, rather than requiring our adherence to a more limited or strict design formalism? Theoretically, yes—this is one of the philosophies behind the practice of “human centered design” (HCD) (Norman, 2005). The design methodology advocated here, borrowing from HCD, is one that attempts to cast the technological operations that computers present to us as cultural, social, or psychological ones that we, as humans, are better equipped to understand; interacting with an information visualization system (or any software, for that matter) should invoke those “human” qualities to make the computer interaction more transparent. Because those qualities are constantly evolving and so context-dependent, it seems difficult to encapsulate them in any kind of standard design format.

Another practical effect that is not captured well in low level quantitative theory, but understood at a high level by the concept of the “digital native,” (Prensky, 2005) is the degree to which interface literacy is becoming increasingly innate to generations of computer users who grew up in a digital environment. Originally characterized by the now quaint example of kids being better at programming a VCR than their parents, it is now observed as a generalized expertise or fluency with computer technology that often transcends poor design; as Prensky puts it (in another echo of McLuhan), these technologies are “extensions of their brains” by virtue of their social and cultural closeness to them. Where older users experience “user interface fatigue,” the digital natives abstract away small differences in functionality and approach them simply as generalized interfaces between them and the content they are seeking. Recently, this notion has been invoked to explain, in part, the success of services such as *YouTube*, which attract millions of users daily despite exhibiting an interface design that rates quite poorly by standard usability heuristics (Silva and Dix, 2007) (though, importantly, the case of *YouTube* also exemplifies the power of affective engagement to overcome design flaws). While we could light-heartedly read this as a sign that all design questions are ultimately moot, it is more powerfully another reflection of the design model argued for above, where design is deployed to humanize our experiences with technology. We should try to identify design vectors that position and emphasize infovis as part of our intellectual, social, and cultural experience, rather than a cryptic, computational artifact that requires an explicit set of skills to decode.

Infovis Design in the Context of our Digital Experience

Any attempt to invoke a broad, theoretical infovis design philosophy must first linger on the last point of the previous section: the realities of software use in the world today. The vast majority of us users, being creatures of habit, will have an immediate reaction to

unfamiliar user experiences regardless of their quality. This is particularly true for information visualization on two counts: First, infovis systems are typically delivered in a context that doesn't necessarily enforce familiar interface frameworks. In the case of public-facing visualization, this typically means internet-delivered "applets" running inside a web browser via Java, Flash, or similar plug-in. As such, their design is not implicitly subject to the sorts of underlying look-and-feel constraints (or defaults) of the operating system. This is, of course, one of the valuable aspects of these delivery technologies, as they allow for the development of OS-independent software, but in the case of visual design presents the possibility of unfamiliar layout and structure (the look-and-feel) and corresponding interface discontinuity with respect to the rest of the user's computing environment. Accordingly, this delivery method emphasizes the "anything goes" nature of infovis design at a technical level and the design issues it raises. It also reminds us that this freedom places the burden of usability squarely on the designer –and with great power comes great responsibility (*Spiderman*, 2002).

Second, information visualization itself, beyond simple graphs and charts, represents a world of visual representation that is fundamentally unfamiliar, regardless of look-and-feel; reading a visualization is entirely different than reading email or a website, or operating a media player. It is arguably rare that anyone encounters an artifact (digital or otherwise) in a public-facing context that is utterly alien to them, and yet much of infovis design today presents this way. "What am I looking at?" is a common response to visualizations presented on the web without qualification. Overcoming this basic confusion is the first, fundamental step in their design.

Speaking to this design goal in general, Donald Norman, author of the popular design texts *The Design of Everyday Things* and *Emotional Design*, conceptualizes the relationship between designer, designed artifact, and user as a triad involving the designer's model of how the designed object should function, the user's model of how it should function, and the "system image" projected by the artifact. The system image represents the way in which the design itself suggests how it should be used (through appearance, functionality, etc.), and given that the designer has no direct communication with the user, Norman argues that good design is that which projects a system image that most effectively communicates the designer's model to the user (2004, p. 75).

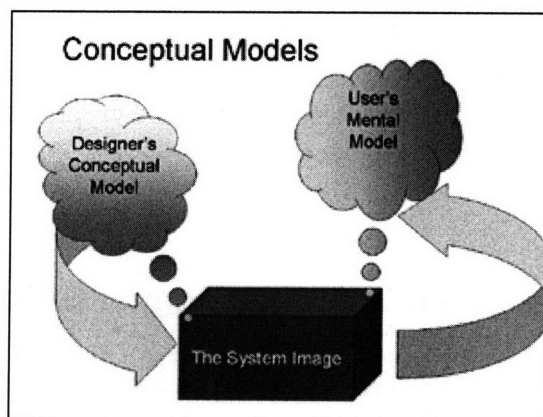


Figure 20. Donald Norman's conceptual models for design.

This model certainly makes sense in the context of industrial design for which it was developed, but in applying it to popular infovis design we have to contend with the possibility that a visualization's system image is too far removed from a user's normal experience of software interaction or information display to render a functional mental model. Given the range of design methodologies alluded to earlier (some of which present intentionally opaque system images) and the relative obscurity of public examples of functional infovis beyond simple charts and graphs, the formulation of a functional mental model cannot be taken for granted. It is entirely possible, for example, for the presentation of an infovis system to be abstract enough that the user fails to recognize it as an instance of information visualization (hence, "what am I looking at?"). At this point, the rest of the design would be utterly undermined.

This last issue is addressed by Lars Erik Holmquist when he suggests a three-stage model for the comprehension of ambient information visualization (2004). "Ambient visualization" is a label given to a particular style of visualization that typically presents information as part of the environment (physical or virtual), and is largely characterized by abstracted displays that focus on simplicity of presentation over detailed representation of data. They are often designed to operate at the periphery of awareness, rather than as tools with which a user actively interacts. As such, awareness that such a display is in fact presenting information is the first of Holmquist's three steps towards comprehension. The second two—understanding *what* kind of information is being visualized, and *how* it is being visualized—can follow only after the first has been established. Holmquist rightly points out that most existing user studies of visualization systems start at the third step, implicitly assuming the first two, and suggests that it is the nature of ambient displays that necessitates the first two steps.

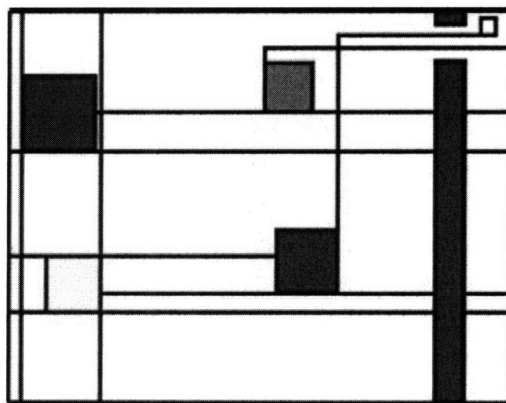


Figure 21. Mondrian-inspired ambient visualization of bus departure times (Skog, et al.).

However, what is true of ambient infovis is arguably true of all instances of information visualization, particularly public-facing examples intended for popular use. This, in fact, gets at one of the fundamental differences between "information visualization design for the people" and more traditional infovis design paradigms, which is an emphasis on the context in which visualization tools are used. A user's encounter with information visualization in the wild is entirely different than the evaluations often conducted in a usability lab: in the lab, the user is not only expecting to interact with a visualization system, but the interaction is artificially orchestrated. Specific questions are asked that

are rarely (if ever) unanswerable from the information provided by the system, and reflect an interaction model that does not correspond to the way we encounter information displays in our normal lives. We don't approach a visualization casually encountered on the New York Times website, for instance, with a list of questions to answer and an implicit imperative to evaluate the display. The assumption that information visualization systems exist solely as software tools used to address a specific task is becoming increasingly quaint, and is longer applicable in our current information and media rich environment. This further emphasizes the importance of Norman's system image in the design of these visualizations. For a passing, non-expert user to make sense of an information visualization display, that display must first be recognizable as such, but also convey its functionality in an intuitive way that fits the context in which it is encountered. There are no lists of questions to guide visualization use in the wild (and if there are, they probably aren't the most natural way to promote exploration), and their designs need to reflect and account for this reality.

To be concrete, the current discussion of infovis design principles for the people assumes a goal of communicating information to a general audience accurately, but also intuitively and in an engaging way. In many ways, this is a direct response to the dominant design axis that has emerged in recent years, as Kosara explicitly suggests (2007), that characterizes information visualization as either analytically or aesthetically driven. Exploration of "the space between," or recognizing that these distinctions are actually somewhat artificial, is important not only on a theoretical or philosophical level, but also because the competing extremes of this axis simply do not serve *the people*. The properties of "intuitiveness" and "engagingness" are often variably devalued by their underlying design methodologies. Simply though generally put, analytic or utilitarian infovis design places little value on engagingness, since it perceives infovis as nothing but a tool—an infovis system need not be any more engaging than a screwdriver. On the other end, aesthetic infovis design often disregards intuitiveness; aesthetic infovis doesn't need to be intuitive because it's "art." Again, these approaches are geared towards very specific populations of users (roughly, data analysts and art enthusiasts), neither of which represent "the masses" in any sense. Arguably, the oft-suggested failure of information visualization to realize its mass-potential is a direct result of a design failure. Infovis designers have either not realized or not cared how their designs are received by average users in the real world, or have simply not designed with that audience in mind. Information visualization design for the people is suggested as an approach that specifically targets this group of users in an effort to promote visualization as a widely usable platform for information exchange. To be clear, however, criticisms of "extreme" design paradigms should be read not as criticisms of those principles in general—there are certainly valid reasons and circumstances for adopting those designs—but simply as a criticism of their applicability and relevance to this design strategy.

For the purposes of popular design, the usual issues of user expertise remain. The infovis system should be neither overly complex, nor overly cryptic, nor so simple that it is perceived as trivial. If the design aims to seriously convey information, it must avoid setting a high barrier to entry, as it doesn't take much to discourage someone from using the system if they are not obligated to do so. However, reflecting the discussion above,

its design needs to be compelling within the context of today's techno-cultural information environment. While most discussions of infovis design are based on the assumption that users have already been convinced, by whatever means, to use the system, infovis design for the people must recognize that it will be competing for user attention with a wide array of other media. Assuming they are being deployed on the web, they will exist along side images, advertisements, links, text, and all the other types of content available on the web. Studies have shown how fickle users can be with their ability to selectively ignore graphical elements of a web page – witness the increasing irrelevance of the banner ad and popular aversion towards Flash-based websites (Benway & Lane, 1998. Rossi, 2006) – so information visualization that aims to engage such an audience needs to take this in to consideration. As these examples suggest, an overly flashy display runs the risk of being ignored (as does an overly “Flash-y” one). Similarly, one can imagine a bland, simple design being ignored as uninteresting background noise or dull statistics. The question again becomes one of discovering the sweet spot where usefulness and readability are successfully married to a compelling, fun presentation.

Theoretical and Methodological Design Frameworks for the People

Given the preceding discussion, how might we identify a design framework for actually producing information visualization for the people? The traditional scientific infovis community has laid the foundation for most infovis design, though it has generally conceived of infovis design as algorithm design, focused on refining specific graphing techniques to improve efficiency or scalability. Accordingly, their approach in developing visualization systems has largely been dictated by the data with which they are working. Though they conduct usability evaluations, these tests are designed to measure specific isolated metrics, such as the time it takes to complete a specific task or the accuracy with which the task is completed. Analysis of the user's experience with a visualization system has often been limited to psycho-physical response (i.e. the ways in which the human eye perceives a visual stimulus and the resulting cognitive operations of the brain). More publicly recognized visualization figures, such as Edward Tufte and Stephen Few, have promoted a design philosophy emphasizing simplicity and clarity of presentation in their criticisms of common design mistakes, though their focus has been more on standard graphical displays (charts and graphs) than on more robust interactive infovis systems. Artists and graphic designers producing infovis on the aesthetic side of the spectrum have deployed sophisticated and effective vernacular methodologies in their designs, though these methods are rarely documented. Finally, there are many design strategies not specific to visualization that have succeeded in capturing the attention of mass audiences in the digital age.

Piecemeal analysis of contributing design factors can be problematic, simply because the complexities of implementation often make it difficult to draw distinct connections between the theoretical motivation of a design choice and its actual significance within a

finished system; when deployed within a complex system, the function of that design choice may change. The design space of information visualization is also sufficiently broad and flexible that many apparently definitive design guidelines can be overturned in certain contexts. This is particularly true for infovis that must appeal to a wide range of users, where specific types of usage and user knowledge cannot be guaranteed. Therefore, rather than attempting to categorize visualization theory and methods in to specific taxonomies, it seems appropriate to take a slightly more holistic approach to the characterization of public-facing design theory.

The following discussion will address four dimensions (or “vectors”) of infovis design that have been reflected in successful implementations of information visualization for the people. These are by no means necessarily mutually exclusive categories (in fact, there is clear overlap between some of them), nor are they necessarily tied to specific design techniques, but rather represent a set of lenses through which popular infovis design can be conceptualized or framed:

1. *Semantic design* refers to the issues surrounding the design of information visualization as a visual language, including the degree to which we are taught to “read” it as such.
2. *Aesthetic / Affective design* refers to the application of “artistic” or affective design principles to information visualization as a means to evoke emotional engagement with its content.
3. *Social design* refers to the incorporation of functionality that facilitates collaborative analysis and the emergence of collective intelligence while also humanizing the process of interacting with data. This is considered distinct from the social view of narrative, which is treated separately.
4. *Narrative design* refers to the conceptualization of information visualization as storytelling, focusing on the ways in which its design can reflect elements of this more fundamental mode of information exchange.

As is perhaps obvious, the first two vectors broadly represent the current dominant approaches to infovis design, as they map to long-standing design dichotomies that are not specific to information visualization, and ostensibly speak closely to the nature of infovis as a type of visual communication. The second two, though having seen less focus in the infovis space so far, represent philosophies that speak to the fundamental nature of human users. As such, they are particularly relevant to a discussion of design that attends to a wide audience.

This being said, the goal of this analysis is not to pit these perspectives against one another or suggest that there are particular “correct” ways of applying them, but rather simply to present them as general frameworks for thinking about compelling information visualization design that resonates with a mass audience. Reflecting the prior discussion of interface design, the hope is that this broader more holistic view helps to advance a

design paradigm that focuses less the specific technical details or narrow views of infovis design, and more on its potential role as a mainstream method of communicating information. It also aims to suggest that information visualization design for the people is about more than “just the data,” and should be guided by a philosophy that attends to the experience of its users, even if doing so occasionally conflicts with existing notions of what constitutes “good” visualization design. The more layers of salience that can be expressed to the user, the better.

Semantics of Infovis Design

“The Semiology of Graphics”

There are (at least) two significant ways in which semiotics, and specifically semantics, intersect with the design of information visualization. The first is in a fairly theoretical sense that nonetheless bears directly on the way information graphics are commonly interpreted today, as it has greatly influenced the way we have traditionally been taught to read simple data graphics such as charts and graphs. As such, it is worth highlighting, if only to understand one of the underlying bases of popular data graphic literacy. The second is more of a practical application of semantics to visualization as a means to enhance readability and understandability; it involves relating visual elements to semantic ones within the data being presented, such as in the use of icons to signify data context.

First, the theoretical: Perhaps one of the earliest attempts at codifying the “language” of visual information displays came by way of semiotic analysis that attempted to map various types of data to types of visual representation. Given that any systematic act of transcribing tabular data to a readable visual form necessitates some contemplation of how that data should be signified, this was certainly implicit in the initial development of data graphics (i.e. the first charts, graphs, etc.) usually attributed to William Playfair in the late eighteenth century. It was, of course, also relevant to the work of later practitioners, such as that of Otto Neurath mentioned earlier. However, it is arguably the more rigorous and aggregative approach published by French cartographer Jacques Bertin in 1967 (and again in 1973 and 1983) that has had the largest theoretical impact on the way data graphics and information visualization are “written” and “read” today. His book, *The Semiology of Graphics*, is often cited as a foundational work in information visualization design.

Bertin was interested in categorizing and organizing the ways in which tabular data could be most effectively displayed visually on a page, and the various functions served by each type of representative element. The fundamental framework around which he developed this formal taxonomy was that of the “ $x y z$ construction,” which essentially referred to the way in which the dimensionality of data could be mapped to the dimensionality of the image plane: “ X and y are the orthogonal dimensions of the table; z is the variation in light energy at each signifying point of the table” (Bertin, 1983). It was the ability to signify relationships between components of tabular data as spatial relationships between marks on an image plane, and the ability to permute the $x y z$

construct (by re-ordering the axes) to explore the “topology” of those relationships, that was considered the basis for the visual representation of information. In other words, this identification formalizes the familiar notion of plotting points in a rectangular coordinate system that we are taught in grade school – for many of us the extent of our education in visualizing data. On top of the $x y z$ construct, Bertin described the different ways in which the marks appearing on the image plane (points, lines, areas, and volumes) could, besides by their positions and sizes, encode information. He called these “retinal variables,” which included color, texture, orientation, and shape. Together with the $x y z$ construction, this characterized the rich set of visual possibilities afforded by the information graphic.

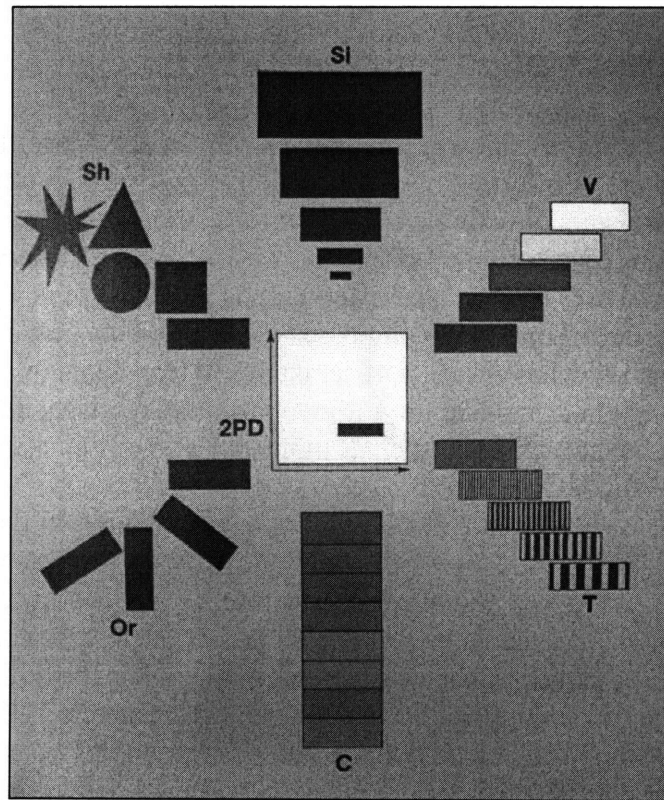


Figure 22. The visual variables as described by Bertin: Size (Si), value (V), texture (T), color (C), orientation (Or), shape (Sh), and the 2D planar projection (2PD) (Bertin, p. 43).

As the modern field of information visualization has evolved, Bertin’s work has been invoked as a model for formalizing the new graphic constructions being developed computationally. In 1996, Stuart K. Card and Jock Mackinlay, who later went on to compile one of the few early “bibles” of information visualization (1999), authored a paper titled “The Structure of the Information Visualization Design Space” in which they expanded on Bertin’s taxonomy (adding “connection” and “closure” to the list of retinal variables) and applied it to the newly developed visualization techniques of the day. In doing so they began to touch on some of the ways in which modern computer-aided visual representations of data were starting to diverge from the established semantic models subtended by Bertin’s $x y z$ construction. New visual representations of data were becoming increasingly unfamiliar, as they relied on more ambiguous semantic mappings that lacked the raw intuitiveness of simple infographics. And this was just within the

strict, “traditional” realm of infovis-as-scientific-research-tool; as will be shown, the new design methodologies employed by the now wider range of infovis practitioners further stress this traditional semantic structure. This becomes relevant to popular infovis design when we consider that most people’s education in reading information visualization is based on Bertin’s formalism – an education that is becoming increasingly outdated.

In their paper, Card and Mackinlay identified a number of semantic mapping ambiguities in the area of “information landscapes and spaces” and large-scale treemaps. Citing a 3D “information landscape” of the New York Stock Exchange, they noted that due to the nature of its presentation (attempting to map the physical NYSE trading room in to an information space that the user “moves through”), certain elements of data within the space are ambiguously represented. In this case, stock volumes were displayed as rectangular blocks whose height represented the stock volume, but whose length, width, and location (within certain constraints) did not. This is perhaps an issue that only a visualization expert would recognize, but it does reflect the confusion anyone might encounter in trying to interpret 3D worlds as information visualizations.

Their analysis of *network graphs* and *treemaps* is more relevant, as these have since become some of the more widely used “advanced” infovis techniques. The network graph in particular is becoming increasingly common and public-facing today, undoubtedly as a result of increasing popular interest in social networks. In principle, a network graph is simple; it is a technique for representing how nodes in a network are connected. However, as seen earlier in the example of *visualcomplexity.com*, network graphs can take on a wide variety of appearances, and thus are potentially difficult to interpret. The most significant source of confusion revolves around the meaningfulness of absolute position so strongly emphasized in the types of graphics Bertin described. As Card and Mackinlay pointed out, the spatial positioning of nodes in network graph is not necessarily semantically meaningful. Their positions are determined by a *layout algorithm* that may place them on the screen randomly, or try to group related nodes together, or attempt an aesthetically pleasing arrangement, or impose some other structure on them. The primary function of the graph is simply to show connections, which is usually done by drawing lines between nodes, so the spatial arrangement of nodes is largely up to the designer. However, users expecting the semantic significance of the $x y z$ construction may find themselves confused by the often arbitrary node positions, looking for order where there may be none. What is learned as being a fundamental component of quantitative data display is rendered less meaningful.

Similarly, Card and Mackinlay note the spatial ambiguity in the treemap construction. A treemap is a “space-filling” representation of hierarchical data (such the folder structure on a hard drive) where branches in the hierarchy are represented as successively embedded rectangles whose areas are proportional to the “size” of the node they represent (in the case of a folder structure, the file size of a given folder). They note that, because of the nature of the treemap layout algorithm, nodes can be arranged either horizontally or vertically depending on their hierarchical “depth” in the tree. Like the network graph, this leads to inconsistent spatial mappings, but also wide variations in size representations: “What the user should be able to take from the image is essentially

Retinal: Size coding, but the same Size can have many different visual manifestations, each with a different aspect ratio. Thus the space-filling property of the visualization comes at a cost.” In other words, the layout algorithm cuts up the space in such a way that identical sizes (i.e. areas) can appear with different aspect ratios making comparisons difficult. Newer approaches to the treemap, such as the Voronoi treemap illustrated in Figure 4, attempt to solve some of these problems with varying degrees of success. The treemap also exhibits similar spatial ambiguities as the network graph, where the absolute positions of its nodes are largely arbitrary.

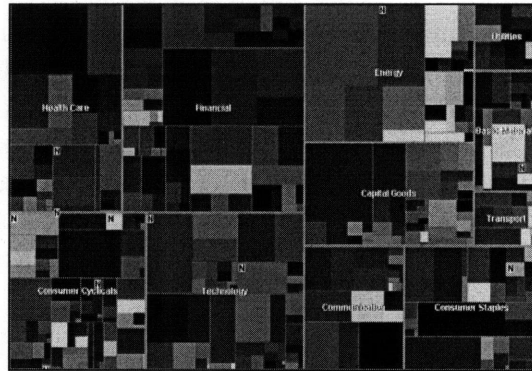


Figure 23. Example of a treemap. While the relative sizes of the rectangles clearly indicate differences in magnitude of some underlying property, the meaning behind their relative positions is less clear.

Moving to more modern implementations of infovis, the issue of semantic mappings becomes increasingly pronounced. The introduction of animation in to information visualization, for instance, introduces new ambiguities. Although Bertin had computer-aided visualization in mind when he developed his *Semiology*, motion did not make it in to his taxonomy. However, modern examples of infovis often use motion and animation in a variety of ways that are not always semantically meaningful. Aside from situations in which motion is used to signify change (in the data) over time, many visualizations feature elements that flash, pulsate, oscillate, jitter, dance, or otherwise animate with no particular significance. Following on the discussion above, one frequent example is in the motion of nodes in a network graph: in Stamen Design’s *Swarm* visualization of the activity on *Digg.com*, not only are the node positions ambiguous, but they move around one another in ways that are not obviously defined. This type of use is potentially problematic because of the high priority motion has to our perceptual systems. Because it draws our attention so well, we are immediately compelled to understand what the motion means, particularly when we are expecting some corresponding significance within the data. If there is no data-based significance, and it is simply being used aesthetically, it runs the risk of distracting users from the informational content of the visualization.

That being said, motion can be semantically powerful when used to animate transitions between visualization states or highlight relationships between elements of the visualization. A common example of this is the use of animated zooming and panning to transition between views on the data. The animation helps to guide the eye by suggesting how the two views are connected. This can help to lessen the jarring disorientation that sometimes occurs when switching between two entirely different states (or views)

instantaneously. A semantically meaningful implementation of this technique can be seen in the zoom feature of the treemap visualizations featured on the social visualization site *Many Eyes* (<http://www.many-eyes.com>). By selecting a node and right-clicking, you can switch the focus to that node, at which point the view will smoothly zoom and/or pan in real-time to bring the selected node to the center of the screen and maximize its size; the animation signifies the operation taking place. In contrast to this, Stamen Design's *Twitter Blocks* visualization of content on *Twitter* (<http://www.twitter.com>) implements both semantic and non-semantic transitions. *Twitter Blocks* visualizes the relationship between *Twitter* users as a "block" of neighbors, represented as a series of connected tiles, each corresponding to a particular user. When first clicking on a tile, the view smoothly transitions from an isometric perspective to a zoomed-in overhead view of the clicked tile in a similar fashion to the *Many Eyes* example. The animation relates the current view to the impending one. However, if you then choose to "load the neighbors" of the *Twitter* user represented by that tile, the entire "block" structure smoothly "dissolves" and is reconstructed with the selected user as its focus. This latter transition is not semantic, in that it does not connect the initial view to the final view in a semantically meaningful way. For instance, the visual representation of the selected tile does not become the "root" tile in the new view. However, that isn't to say that such a transition isn't compelling or appropriate in other ways (which will be discussed later).

When moving outside of traditional information representations, semantic mappings can become even more abstract. As suggested earlier (and to be expanded on later), the recent appropriation of information visualization by artists, designers, and analysts of various types has introduced new visual methods that can stray quite far from the $x y z$ construction and its descendants. There are often compelling reasons for this, aesthetic or otherwise, but their mappings run the risk of being confused or misinterpreted. The Ambient Orb is often used as a pathological example of this; the Orb, a physical device, relies purely on changing colored light to visualize its information stream (stock market conditions, etc.). Aside from giving no explicit indication of what data is being visualized, its semantic mapping is unclear, or unintuitive, without knowledge of how the system works. In some cases, visualizations deploy sophisticated semantically-charged visual metaphors: InfoCanvas presents data as elements of "abstract scenes or 'virtual paintings' in which the images in the scenes change subtly to convey the state of an information resource" (Stasko, et al., 2008). In one example, a beach scene is used as the substrate on to which the data is mapped; the data is suggested by the appearance and states of various beach-related objects, such as seagulls, beach-balls, boats (although the data itself doesn't necessarily have any semantic connection to seagulls, beach-balls, or boats). Google's *iGoogle* service (<http://www.google.com/ig>) takes a similar approach in its implementation of "themes" that reflect the passage of time and weather conditions. Some are semantic mappings that follow the $x y z$ construction (time zone maps, satellite images of the Earth as it rotates), some are in the style of InfoCanvas (cartoon scenes of a bus stop or a Japanese teahouse), and some are utterly abstract (John Maeda's "Simplicity is Complex" theme). VisitorVille (<http://www.visitorville.com>) maps website analytics to a SimCity-style game environment, where web pages are houses, search engines are buses that "drop off" visitors their destinations, and the visitors themselves are little characters walking around the "neighborhood." Finally, visualizations like Jonathan

Harris' *We Feel Fine* (<http://www.wefeelfine.org>) exhibit elements of both abstract and concrete representation of data. While the default display maps its data (in this case, blog postings) in an entirely abstract and chaotic manner (perhaps symbolizing the chaos and turmoil of the “feelings” represented in the piece), a mode is available that presents it as a more “quantitative” $x y z$ mapping that allows for more intuitive analysis.

Ultimately, the question of whether to adhere to the formal semantic mappings characterized by Bertin's $x y z$ construction or to forgo them in favor of other design considerations will depend on specific applications. It is, of course, entirely possible that particular information visualizations cannot be designed to respect these mappings. The purpose of this discussion is not necessarily to suggest that it is preferable to abide by them, but simply to motivate their significance to the way non-expert users attend to information displays. Bertin's construction is reflective of the way in which we learn to read quantitative data graphics and are therefore reflective of our basic intuition about visualization in general. It is important to remember that our literacy for infovis, at this point in time, largely derives from our lessons in grade school (“the independent variable goes on the X axis, and the dependent variable goes on the Y axis”), and is further reinforced by the most traditional venues of data presentation (newspaper graphics, Microsoft Excel, etc.). Significantly, as suggested by Matt Ericson of *The New York Times* during his keynote presentation at *InfoVis 2007*, even this level of literacy cannot be guaranteed: in designing visualizations for their newspaper, they discovered that many readers have trouble interpreting scatterplots in which the X axis does not represent time. While this is an anecdotal observation, it certainly reflects on our common experience with data graphics, both as students and beyond, where the vast majority of graphics to which we are exposed represent the change in some variable (income, gas prices, deaths, stock volumes, etc.) over time. However, information visualization, regardless of type or agenda, aims to go far beyond the simple plotting of “change over time.” For it to achieve the desired effect on mass, non-expert audiences, we must always be conscious of how that audience will approach such displays in their attempts to make sense of them.

Many types of infovis are successful as a result of this type of consideration. Despite recently becoming an object of ridicule amongst infovis aficionados, the pie chart is a fundamentally understandable construction because of the immediate intuitiveness that its name suggests. A more modern example, the tag (or word) cloud has rapidly proliferated across the internet for similar reasons. In this case, it succeeds because of a simple mapping (quantity to font size) and because it signifies its data with the data itself—the number of occurrences of a particular word or tag is signified in the cloud by the font size of the word.

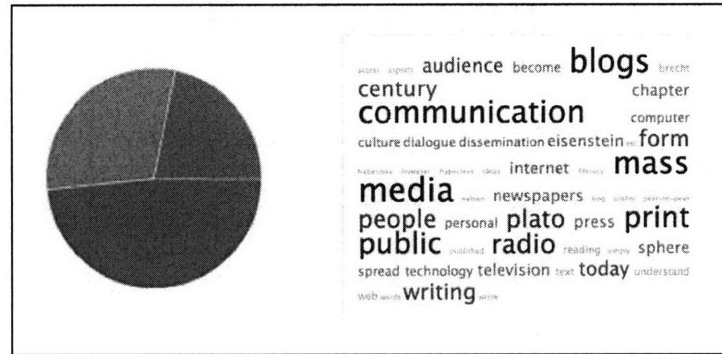


Figure 24. Pie charts and tag clouds – simple but intuitive.

Of course, these are relatively simple examples that present simple mappings, but as much as is possible, the design of more complex visual semantics should begin from the same simple principles of intuitiveness. As suggested by the analogy to written language, the semantic mappings of a visualization are fundamental to its readability. Without a clear relationship between the components of the image, the corresponding relationships in the data cannot be clearly expressed.

Icons, Chartjunk, and Information

The second application of semantics to information visualization is more specific and practical to popular infovis design, and is a direct response to the inherent abstractness of all forms of visualization. Influenced again by the work of Otto Neurath, it is the technique of embedding icons, symbols, or small images in to a visualization to help signify the nature of the data being visualized. This could be any sort of “meta” information about the data, or about specific trends within it. For instance, appropriately chosen icons might make a visualization of food sales more quickly recognizable; the icons could be matched to the semantic content of the data (i.e. coffee mugs representing coffee, a cow representing beef, etc.) such that a viewer can immediately grasp what the visualization is about. The suggestion here is that many forms of visualization are sufficiently complex (or, conceivably, dull) as to lack an entry point for the non-expert or casual viewer. Confronted by the need to decode abstract visuals, the viewer may lose sight of—or potentially interest in—what is actually being represented. Rather than relying on additional abstractions to convey this relationship, such as colors or shapes defined in a legend, icons and symbols provide a more direct representational way to express this information and maintain the connection.

Given the broader social and cultural precedent for the use of icons and symbols (and obviously photographic images or illustrations) to convey information, particularly now with their relevance to common graphical user interfaces fundamental to today’s computer use, it makes sense to incorporate them in to information visualization systems aimed at wide audiences. However, in addition to the evidence emphasizing the efficiency and memorability of icons that underpins their use in software interfaces (Ferreira, et. Al, 2006), and the observation from a cognitive and perceptual perspective

that “object displays will be most effective when the components of the objects have a natural or metaphorical relationship to the data being represented” (Ware, 2004), they also work well both as symbolic and rhetorical devices. Newspaper visualizations, for example, often make use of icons, symbols, and images to enhance the semantic meaning of the graphics themselves. In some cases, such as in *31 Days in Iraq*, the use of iconic imagery serves as a literal reminder of what the data represents, but also makes a more deeply rhetorical statement in its representation of human deaths with human forms rather than, for example, abstract colored dots. As another example, geographic maps have long been used as the basis for data visualization, given their immediately recognizability, allowing associated data to be quickly correlated to places in the real world.

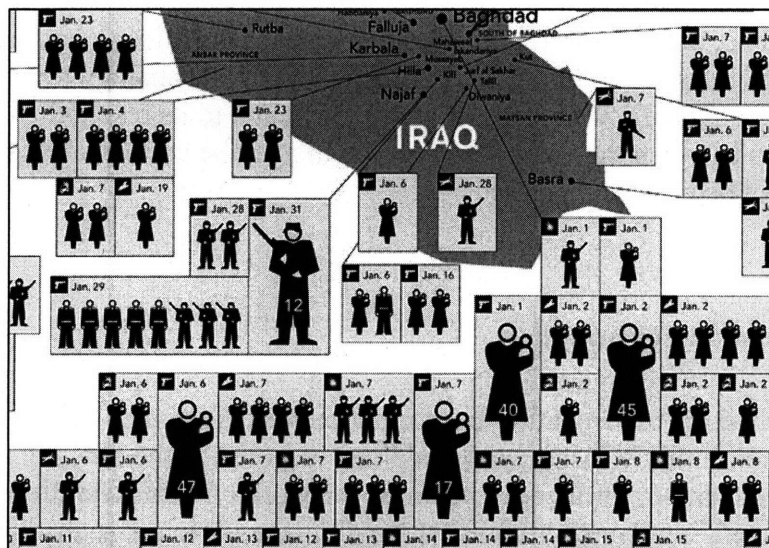


Figure 25. Detail from *31 Days in Iraq*, illustrating Iraqi deaths as human icons (Adriana Lins and Alicia Cheng).

As another example, compare Jeffrey Warren’s *ARMSFLOW* visualization (<http://www.armsflow.org>), which visualizes the trade of weapons between different countries, to Jonathan Harris’ infographic for the International Networks Archive, *Choose Your Weapon*, covering a similar topic. Though this isn’t a completely fair comparison – the two examples don’t address exactly the same data, and *ARMSFLOW* is an interactive, exploratory visualization while *Choose Your Weapon* is a static, presentational one with a more overtly rhetorical angle – their presentational strategies are significantly different. Ignoring their titles, the abundant use of icons related to weaponry (bombs, bullets, military vehicles) in *Choose Your Weapon* makes it immediately obvious what the graphic is about. *ARMSFLOW*, on the other hand, uses a much more abstract and minimal presentation. While there is nothing wrong with this in principle (and there are certainly many valid criticisms that could be leveled at the data presentation in *Choose Your Weapon*, including its high “lie factor” – the term Tufte uses to describe when the size of data representations scale disproportionately to the data itself, as with the areas of the bullets in the graph that are only meant to indicate their data values by their height), its presentation enforces a level of semantic abstraction between the visual and the data being represented. Looking at the image alone, nothing

about it suggests that we are looking at the flow of weapons in particular across the globe. In other words, it doesn't reinforce its meaning in an explicit and obvious way.

While it is arguable that this type of abstraction might encourage user exploration in some cases, there is a fine line between visualizations that are mysterious and intriguing, and ones that are mysterious and incomprehensible. Unfortunately, much of current infovis falls in to the latter category with respect to general audiences, either due to excessive abstraction, excessive complexity, or simply a lack of explanatory documentation. This type of semantic enhancement aims to mitigate these sorts of issues.

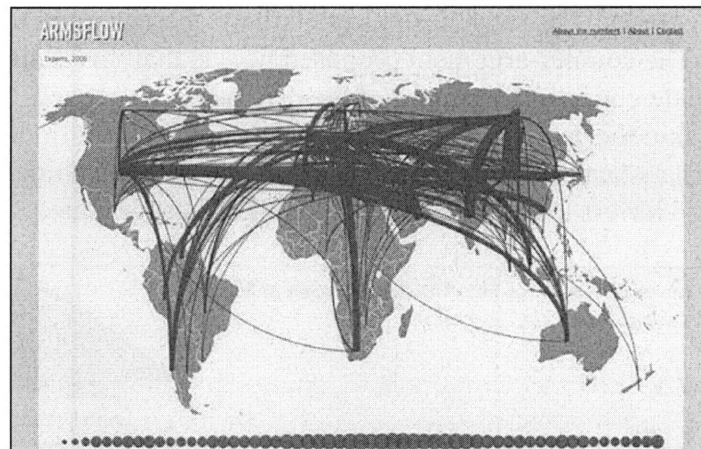


Figure 26. *ARMSFLOW*, semantic content is abstracted in the visual (Jeffrey Warren).

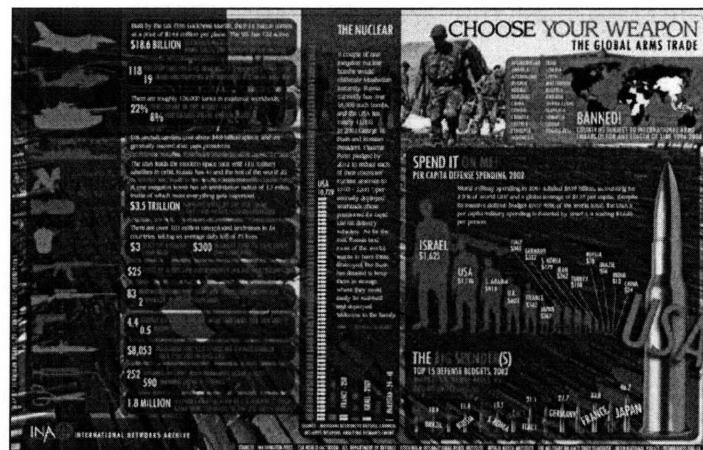


Figure 27. *Choose Your Weapon*, semantic content is emphasized in the visual (Jonathan Harris).

This being said, it also often potentially challenges two long-standing conventions of infovis design. The first revolves around Edward Tufte's oft-cited principles of "chartjunk," "data-to-ink ratios," and "graphical ducks." In his series of well-known books on the design of infographics, which refer primarily to printed graphics, Tufte has argued that the key to clear and informative data graphics is the removal of any distracting elements (or "decorations") that do not explicitly represent the data being displayed, or are "content-free" (1983, p. 177). While he does advocate the use of image

“Semantic enhancement” is also difficult to reconcile with the infovis convention that advocates content-agnostic design. This is a convention that implicitly arises out of the very nature of visualization: most common graphing methods exist because they are useful “standardized” ways of representing information. Bar charts, for example, exist because they afford a way to compare quantities of *anything* in a straightforward manner. Thus, there is a tendency, particularly in the design of visualization tools that allow you to author your own visualizations (from the high-end, such as *Tableau*, to the more common *Microsoft Excel*) to abide by a “neutral” display style so as to accommodate any particular data sets you might want to visualize. This is, of course, partially a technical limitation – *Excel* can’t yet determine what the numbers in your spread sheet *really mean*, let alone produce icons or picture that signify that meaning – but it also serves to provide a transferrable understanding of data graphics. Kosara notes that “[these] visualization techniques are also often general, and can be applied to many different data sets. This is considered a strength, because the user can gain experience with the method and apply that to different data, rather than having to start from scratch again.” Catherine Plaisant, though referring mostly visualization in a research context, articulates the opposing view:

When utility has been demonstrated in one setting, potential adopters often still need to jump across application domains and envision how the tool could be used. What could we do to encourage this leap of faith? Paradoxically our quest for generality can become an impairment. Researchers are encouraged to – and rewarded for – designing techniques that are generic in nature, can be used with a wide variety of data, and used in many application domains. However, potential adopters might be turned off if they perceive that the tool they are evaluating has not been designed specifically for their particular needs [...]

Plaisant goes on to suggest that visualization designers invest time in tailoring their tools to the specific needs of various domains (2004), which is essentially what “semantic enhancement,” as envisioned here, should aim to do. If domain-specific design is attractive for research applications, it is certainly appropriate for public-facing visualization, where the extraction of meaning from a visual display must occur over a much shorter period of time. However, these are not mutually-exclusive approaches to design. Reinforcement of the data’s semantic content can be gratuitous or subtle; as always, a successful implementation depends on arriving at an appropriate balance.

Affective Infovis Design

Culture Clash

While originally deployed within the confines of the laboratory, the practice of infovis design has more recently been appropriated by artists and graphic designers who produce work that appears in both in galleries and across the web. In many ways, this work in fact characterizes the public face of information visualization, reaching a significantly larger public audience than its more scientific applications while simultaneously representing vernacular design theories and methodologies that remain largely outside the

scope of academic infovis research. Andrew Vande Moere characterizes this new movement as follows (Burkhard, et al., 2007):

In recent years, a stream of mainly young and self-motivated people is experimenting to visualize fashionable realworld datasets in ‘artistic’ ways. Independent from institutional or commercial pressure, their visualizations demonstrate the initiative, enthusiasm, interest and skill to tackle complex issues that were previously reserved for the expert visualization researcher or developer. Captured in the relatively limited fame of online weblogs, galleries or museum exhibitions, their contributions seem to be ignored by the academic world. This emergent movement is driven by a combination of contemporary social phenomena, including the current seductiveness of the ‘information society’, novel software tools specialized in the creation of visual artefacts that are specifically developed for designers, the cross-fertilization of computer science insights in design schools (and vice versa), and the online creation and sharing of interesting datasets.

While the origins of this division as it relates to the practice of infovis are questionable, the division itself is clearly focused around the age-old Modernist dichotomy of *form versus function*, or more recently, *beauty versus usefulness*. There is seemingly an assumption amongst infovis designers that these dichotomies are not reconcilable; if a visualization design is beautiful, it cannot be useful, and vice versa. However, the challenge to this assumption, and the fact that the dichotomy exists in the first place, is significant in that it suggests an important paradigm shift in the way information visualization is being conceptualized. Kosara argues that this comes from the realization that infovis is not, as traditionally assumed, a hard science. Instead, by virtue of its explicit communicative nature, it shares properties with other approaches to communication:

The most basic description of visualization is that we produce images from data. In other words, we depict data, and we do that to communicate information. This is not at all unlike what artists, illustrators and designers do (and have done for a long time). Like artists—and unlike traditional scientists—we build artifacts. Unlike artists, we design fairly general tools rather than specific pieces for communicating just one idea for one purpose; and we also need to produce things that are useful.

As he goes on to point out, the Modernist conception of infovis as a scientific instrument should give way to a postmodern one that considers infovis as an extension of a broader visual culture, and its design should reflect this (Jankun-Kelly, et al., 2006). Accordingly, before continuing an exploration of aesthetics in information visualization, it is worth emphasizing that the differences in infovis theory embodied in this dichotomy are, in fact, mostly sociological or cultural, rather than hinging on any inherent principles of design. As Fishwick suggests, they are a result of a “culture clash” between scientific and artistic communities (2008).

It is, of course, somewhat counter-intuitive to imagine a theory of visualization design that advocates the intentional production of “unattractive” images, and no such theory actually exists. Instead, we can understand the historical development of the traditional infovis design philosophy primarily as arising from an incubation period spent within the confines of scientific research environments. The initial forays in to computer-aided infovis were a response to a need to process large amounts of scientific data. In that context, visualization systems were being developed by the researchers themselves to address specific needs. Aesthetic concerns were likely either not considered (research scientists are typically not designers), not feasible (given technological limitations at the time), or not budgeted (why spend time and energy on the aesthetic component of a research tool when there was actual research to be done?). Tools simply serve functions, and you don’t necessarily need to be happy about using them as long as they give you the results you want.

The second major influence on this design paradigm arguably came, once again, from the philosophy emphasized most strongly by the work of Edward Tufte. Tufte’s four books, particularly the three written during the period that information visualization was first coalescing as a field (late 80’s – 90’s), have had a profound impact on “the visual display of quantitative information,” whether in the context of infovis or print-based infographics (prompting the New York Times to declare him “the Da Vinci of data,” for instance (Shapely, 1998)). Many of his insights are undoubtedly profound, most significantly his exposé of the ways in which data graphics can distort and misrepresent data (the “lie factor”). However, his position on the aesthetics of data graphics follows entirely from the belief that “form follows function” and “ornamentation is a crime.” Though he agrees that artistry is a necessary part of the production of effective visualizations, his conception of the term does not seem to include the important affective role it can serve. Instead, he worries about a growing “art bureaucracy” conspiring to replace “content” with “style.” (1983, p. 80).

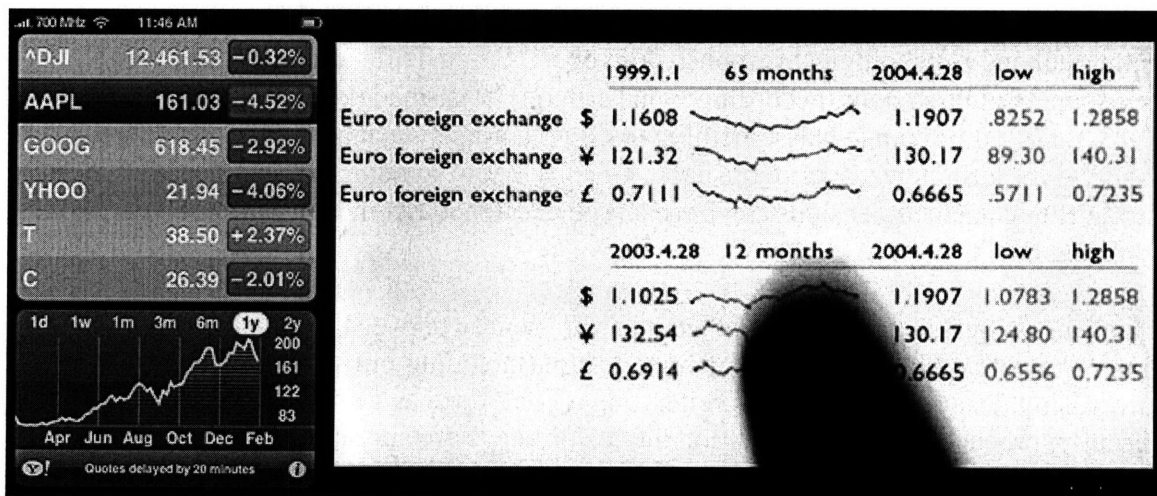


Figure 29. Tufte’s response (right) to Apple’s iPhone stock market graphics (left).

However, in practice the dichotomy that exists today is mostly likely a result of a fairly distinct separation of the contexts in which visualization is produced. Designers in the

“scientific” camp, designing mainly for research purposes, favor methodologies that maximize a set of usability criteria (efficiency and accuracy of the tasks undertaken by a user working with the system, for example), and are often simply unaware or unconvinced of the benefits of good aesthetic design. Similarly, “artistic” designers, targeting art enthusiasts, strive to implement maximally affective designs, often without regard for the value of the methodologies deployed by the former group. Neither is typically designing for a mass audience, so once again, the sweet spot is actually somewhere in between. The purpose of this sociological detour is to reiterate that any division between the aesthetic and the pragmatic is largely social and cultural, having more to do with ideological differences than irreconcilable design rules. While a simple model of the relationship between the aesthetic and pragmatic approaches to infovis may suggest that the pragmatic requirement to accurately represent the data constrains aesthetic creativity (and vice versa) (Lau and Vande Moere, 2007), the potential gains from aesthetic design can often outweigh any perceived loss in analytical accuracy. In other words, good aesthetic design is more than just another pretty face.

Why should aesthetic concerns drive design to the same degree as standard questions of usability and function, particularly with respect to infovis? As reflected in critiques such as Tufte’s, there is often a negative connotation associated with aesthetic design that implicates it as extraneous fluff whose designers are peddling style over substance. It is often regarded with the same skepticism as the practices of the marketing and advertising worlds, where aesthetics are seen as a means to distract us from the deficiencies of a product by appealing to our basic attraction to “beautiful things,” even if those things are completely irrelevant to its function (think beer commercials). However, while a concern over a lack of substance in design is certainly legitimate, particularly with respect to the design of ostensibly functional artifacts like information visualization systems, these criticisms ultimately represent a short-sightedness over the scope of design and its effects. Returning to the pathological case, it may be intellectually distasteful to see images of scantily clad models used to sell low-quality domestic beer, but there is something undeniably effective, if not compelling, about its application of aesthetic design to the advertisement at hand. Design like this is deployed for a reason: the emotional and psychological response it invokes. When Tufte asks why major newspapers publish poor (according to his criteria) artist-made data graphics (Tufte, 1983, p. 79), it of course has something to do with artists’ lack of training in statistics, but it is also because newspaper publishers evidently know something about design that Tufte doesn’t: that aesthetic designs can have a type of impact that more sterile, utilitarian ones typically don’t.

This tension gets at one of the fundamental differences between the conceptual frameworks that have governed software design (including infovis) and those that inform advertising design. At the risk of reiterating a point once again, software design has, until recently, proceeded from the position that its products are simply tools that accomplish tasks according to objectively measurable usage criteria. The “user” is positioned as a black box that exists only to invoke the functions of the software. The field of advertising, on the other hand, has been more inclined to study the psychology of the user, even if only one-dimensionally, in an attempt to understand what makes them tick.

The emerging field of *affective design* is, in many ways, the application of the latter methodology to the world of software development. Popularized by Donald Norman, affective design refers to design that attempts to elicit a pleasurable emotional response from its users. In practice, such designs attempt to be fun, playful, beautiful, engaging, or otherwise alluring, and stand in contrast to more Spartan, purely utilitarian design paradigms that emphasize the functionality of a designed artifact over the quality of the experience one has while using it. Affective design recognizes users as emotional beings whose emotions influence their functional interactions with their tools. Based on experimental evidence suggesting that a positive emotional state facilitates more effective interactions (Isen, 1983), this model argues that the consideration of aesthetics can inform design practices. In the context of information visualization, because of its visual nature, these affective principles are most often conveyed in the visual aesthetic of the image – visual constructions are chosen to maximize “beauty” in the presentation of data – but are also present, when applicable, in the interaction design of a system. Accordingly, regardless of sociological considerations, what is significant about the new infovis design culture described above, whether cast as artistic practice or otherwise, is its focus on the exploration the affective in a context usually focused on more quantifiable conceptions of usability. But, as work by Norman, Tractinsky, and others has suggested: “the beautiful is usable.” (Tractinsky, et al., 2000).

It is difficult to establish definitive practical rules for aesthetic or affective design. Beauty is not easily quantifiable. However, the following two sections will present two frameworks that aim to provide guidelines for understanding the role of these principles. The first, based on the work of Vande Moere and his students, attempts to disambiguate the notion of aesthetics in infovis design. The second, influenced by the worlds of advertising and affective design, takes a more holistic approach to the role of aesthetics and affect by conceptualizing them as implicit elements within a broader design paradigm whose goal is the “seduction” of the user.

Information Aesthetics

In 2007, Vande Moere and his students Andrea Lau and Nick Cawthon published a series of papers that are among the few attempts, academic or otherwise, to analyze the use of aesthetic design in information visualization from a practical perspective. Inspired by Manovich’s notion of “info-aesthetics,” which refers to the analysis of the “aesthetics and computer-based cultural forms specific to information society” (2001), and in support of the intuitive observation that aesthetics are integral to visual displays of information, their work attempts to construct a taxonomy of infovis within an aesthetic continuum, formalizing the aesthetic/utilitarian dichotomy, as well as describes a number of different ways in which infovis implementations can embody aesthetic principles. Finally, they motivate the need to understand the specific impact of aesthetics on usability while simultaneously arguing, as was hinted at in the previous chapter, the inadequacy of standard usability metrics with regard to capturing the impact of the sorts of affective

interactions that are facilitated by aesthetic design. Vande Moere argues that “these software tools should not only be designed by interaction engineers and software experts, but that designers are required to create novel forms of *information art*, that appeal to user engagement and aesthetics and provoke long and enjoyable usages” (2005).

Described as “information aesthetics,” their model is characterized by two factors that determine how aesthetics influence visualization design: *data focus* and *mapping technique*. Mapping technique refers to the ways in which data is visually encoded, and defines a spectrum that puts “direct” mappings on one end, and “interpretive” mappings on the other. Similar to Kosara’s pragmatic-aesthetic continuum, “direct” mapping represents the traditional infovis design paradigm that emphasizes efficiency and readability of the data, while “interpretive” mappings represent the “artistic” extreme that typically devalues readability in favor of other aesthetic considerations. Data focus refers to the degree to which the visualization supports rigorous data analysis versus more reflective interpretation of the data.

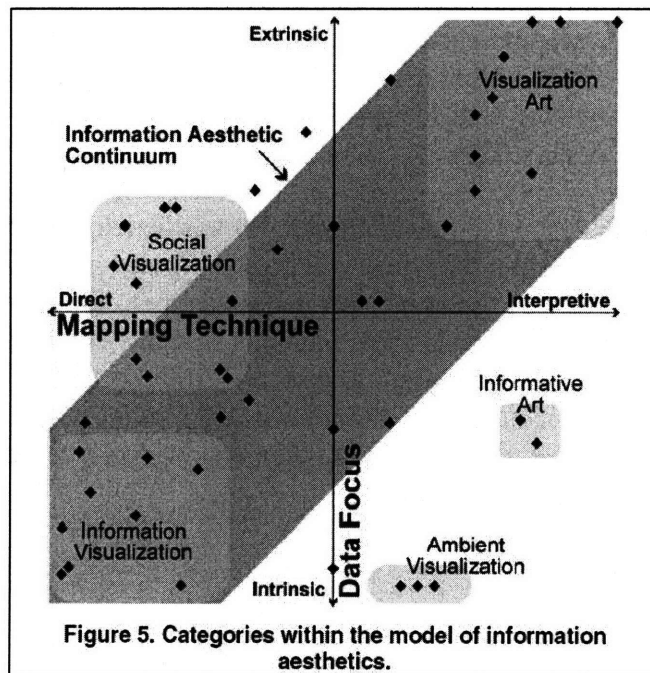


Figure 30. A model of information aesthetics (Lau and Vande Moere, 2007).

These axes define a design space that reasonably encapsulates the range of aesthetic implementations in infovis design. Technical, analytic implementations, such as the TouchGraph Google browser (<http://www.touchgraph.com>), inhabit the direct/intrinsic “corner,” while examples of “information art,” such as Cory Arcangel’s “Data Diaries” (<http://turbulence.org/Works/arcangel/>), occupy the interpretive/extrinsic realm. However, in mapping these and forty-five other existing examples of information visualization to a graph of this design space, the significance of this taxonomy immediately emerges. The vast majority of the examples in this set exist either in the lower-left direct/intrinsic quadrant, or in the upper-right interpretive/extrinsic quadrant. Of course, there is a clear correlation between the mapped variables, but it this suggests that the extremes are in fact the rule. In other words, infovis design tends to be either

extremely technical, or extremely “artistic.” The space between remains largely unexplored. As Robert Kosara puts it, “the question remains, though, if an aesthetically pleasing visualization has to appear like a work of art, or can be easily recognizable at the same time” (2007).

The next question this study provokes is: where is the sweet spot within this space for effective, public-friendly information visualization? While continued exploration of designs that incorporate elements from each of the “edges” of the spectrum seems like an intuitive course, it is still unclear how aesthetics ultimately impact the reception of visualization systems. What proportions of intrinsic vs. extrinsic data focus and direct vs. interpretive mapping technique represents the most compelling experience when actually sitting in front of a visualization system? How are these proportions implemented in practice? What psychological or emotional responses are they actually invoking? The answers to these questions are likely as varied as the range of personal preferences of infovis users, making it all the more difficult to design for a mass audience. So far, the user studies that attempt to make sense of these questions remain largely inconclusive in determining anything other than that users are typically drawn to pretty pictures (Lau and Vande Moere, 2007). This may suggest that the influence of aesthetics is too subtle to quantify, particularly with standard evaluation techniques. It may also be impossible to isolate aesthetic considerations, in practice, from other elements of infovis design; in other words, the final product may always be more than the sum of its parts – particularly given that a great deal of its appeal is undoubtedly related to personal relevance of the data it represents, independent of the “interface” to that data. Or, it could be that establishing a visualization as “pretty,” with no further qualification, is sufficient to capture the useful effects of aesthetically pleasing design.

The model of “information aesthetics” described above is undoubtedly useful academically, as it establishes aesthetics as an active component of infovis design and describes its influence in fairly objective (if generalized) terms. However, to address the subtlety of its impact, it is perhaps useful to move to a more holistic view aesthetic design. Such an approach might begin with more ethnographically inspired evaluation methods; rather than trying to isolate and dissect the role of aesthetic design elements, we should follow the current trends in the fields of HCI and turn our attention towards the unified user experience. Ultimately, the application of aesthetic design principles will be effective (and affective) only as long as such design fits within the larger infovis experience, which includes many factors both within the design of the system and in the context in which it is encountered.

A holistic approach may also provide more valuable, practical design guidelines that address a wide audience. Clearly, the use of aesthetics is not defined with enough specificity to suggest many particular rules (although there are exceptions; the use of certain fonts, for example, can evoke particular reactions), so we may be better served by taking a step backwards and addressing what value aesthetics provides in the big picture. Returning again to the principles that have long guided the mass-targeted fields of advertising, industrial design, and, more recently, affective design, the next section explores a model of aesthetics in infovis that sees as its goal the seduction of the user.

Seduction Design

In their 1999 paper, “Understanding the Seductive Experience,” Julie Khaslavsky and Nathan Sherdoff discuss the ways in which principles of seduction can be applied to software design. They note that, unlike other more mature fields, software design, with its traditional emphasis on features and functionality, has lagged behind in its understanding of how to entice users in to the sorts of relationships with software products that “transcend issues of price and performance.” The notion that the appearance of a product could have a significant impact on its market success was, at that point, only starting to be recognized, despite the obvious seductive qualities long emphasized in the design of cars, buildings, and entertainment media (take, for example, Ferrari’s long history of aesthetic automotive design inspired by the female form). Though the situation has clearly improved somewhat in the intervening ten years (the authors point to Apple’s iMac as a success story, but the article predates the massive success of the iPod and iPhone—which went on to become cultural icons—and the resulting attention they drew to issues of aesthetic design), there is still an apparent resistance towards thinking about software aesthetically or seductively. Information visualization is no exception.

In characterizing suggested steps to encourage seduction in software design, Khaslavsky and Sherdoff present two brief case studies. The first examines a non-software design: the famous Philippe Starck juicer featured on the cover of Donald Norman’s book, *Emotional Design*. The second, coincidentally or not, looks at an example of information visualization. It is one of the first well-known examples of popular, public-facing infovis on the web: Plumb Design’s (now ThinkMap, Inc.) *Visual Thesaurus*. The *Visual Thesaurus* is a thesaurus presented with a network-graph interface, where selected words are represented as nodes connected by lines to their synonyms, allowing you to browse through and explore word relationships by expanding and collapsing parts of the graph.

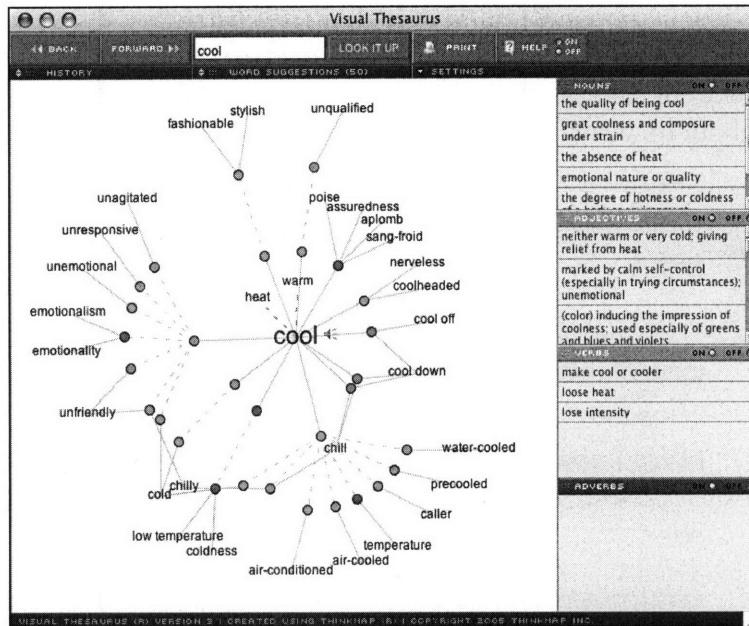


Figure 31. Screenshot from *Visual Thesaurus*, a visual interface to an online thesaurus (Thinkmap, Inc.).

The first four of their steps to good seduction design represent, at a broad level, goals that are coincident with those of aesthetic or affective design:

1. Entice by diverting attention
2. Deliver surprising novelty
3. Go beyond obvious needs and expectations
4. Create an instinctive emotional response

The case study implies that information visualization is by its nature attention-grabbing and novel as compared to non-visual representations of data, which is true, but these factors can still distinguish compelling visualizations from less interesting ones. Despite Tufte's insistence, the vast majority of data graphics and visualizations out in the world are fairly bland by most imaginable measures, and in many cases are more likely to be ignored than to entice attention. Truly compelling implementations that aim to draw in casual users must overcome this tendency.

Though it is impossible to objectively define the successful application of seductive aesthetic design to visualization, it is perhaps useful to contrast implementations that address similar data sets in order to see how attending to these principles might yield compelling experiences. As an example, compare *Google Analytics* to the previously mentioned *VisitorVille*. Ignoring issues of accessibility (*Analytics* is free and web-based, *VisitorVille* is not), both of these visualization systems provide a wide array of website analytics to the user, tracking the usual metrics of hit counts, referrals, visitor loyalty, and the like. *Google Analytics* presents the results in a series of elegant but straight-forward charts and histograms with a functional but minimalist aesthetic. As with the *Visual Thesaurus* example, neither *Google Analytics* nor *VisitorVille* can really divert attention, since you must explicitly launch both in order to experience them (as opposed to incidentally encountering them within the context of a web site). However, it is difficult

to argue that the visual component of *Google Analytics* delivers surprising novelty, goes beyond obvious needs or expectations, or creates an instinctive emotional response. It is certainly functional, but aside from a compulsion to track the statistics of your web sites, there is nothing about its visual presentation that makes you want to interact with it.

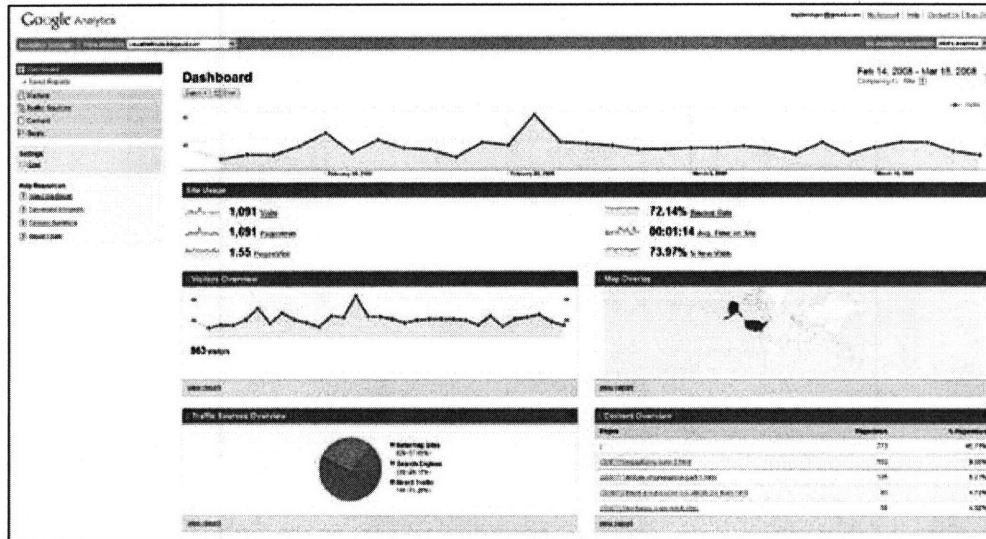


Figure 32. Web analytics in *Google Analytics* (Google).



Figure 33. Web analytics in *VisitorVille* (World Market Watch).

In contrast, *VisitorVille*, as described in the previous section, presents an entirely different visual metaphor in its display of website analytics. Instead of charts and graphs, the analytic metrics are mapped to a virtual “information landscape” similar in style to the popular video game *SimCity*. Web pages are represented as buildings in a neighborhood, hits are represented as little avatars entering the buildings, search engine referrals are represented by buses dropping off the avatars at their destinations. While more standard charts and graphs are available, the primary display represents an utterly novel way to interpret web analytics. It goes beyond expectations in its use of an

intuitive metaphor that emphasizes the personal and social nature of web traffic. It evokes an emotional response in the same way that *SimCity* or *The Sims* might: first by being “cute,” second by being fun, and third establishing an emotional attachment to the data; this isn’t just your web site, it’s your “city,” and visitors aren’t just anonymous numbers, they are real people that you are trying to engage with. This attachment, along with the game-like interface, might compel someone to use the system who would otherwise avoid more statistical displays.

That being said, the relative effectiveness of these systems is ultimately dependent on a variety of additional factors that speak to the limitations of the comparison (including issues of cost and the role of branding on perceived value). The most significant is the fact that web analytics reflect a type of data that is typically studied over a long period of time. As such, analytics tools are often used by a more invested, arguably expert audience, which complicates the use of seductive design somewhat. Implicit in the seductive philosophy is a focus on designing towards the initial attraction a user feels towards the designed artifact; the degree to which seduction plays a long-term role in the user experience is more subtle. In many cases, once a user develops familiarity with a system, they prefer a more streamlined interface which may forgo the four seductive steps in order to keep things simple. Furthermore, for expert (or semi-expert) systems like this, the more direct *Google Analytics* style interface may in fact be read as more seductive because of this. In any case, the comparison here is used merely to motivate the way seductive, aesthetic, and affective visual design can be mobilized to produce more compelling visualizations of information. In keeping with the argument that addresses the use of visualization by non-experts in the ways described earlier, a user’s “first impression” is extremely important, and can be served through this seductive approach. However, it is also important to reiterate the fact that seduction should not be a consideration that isn’t weighed against usability more generally. A well-executed seductive visualization should remain seductive throughout its relationship with the user.

Accordingly, it is important to note that seductive design need not present a radical aesthetic departure from more traditional design paradigms. As suggested earlier, while infovis lends itself to the exploration of the affective in visual aesthetics, simple choices in interaction design can be powerfully affective as well. Stamen Design’s *Stack* visualization of *Digg.com* (<http://labs.digg.com/stack>) is fundamentally a simple histogram measuring the number of “Diggs” (votes) current stories have received, but the designer’s choice to refresh the underlying data in real-time (as opposed to refreshing every few minutes, or requiring a manual refresh by the user) provides an opportunity to draw users in further. As votes are received by the site, they are represented as small blocks falling down from the top of the screen and landing on the appropriate histogram bin, adding to its height. This simple mechanic imparts a game-like spin to the aggregation of data: watching the blocks fall and accumulate, you can’t help but pit the different stories against each other in a race for votes (i.e. “Which will reach the top of the screen first??”). Even the simple emotional engagement this feature affords makes the system more compelling – and in this case, it’s hard to argue that the affective design choice interferes with the accurate representation of the data in any way.

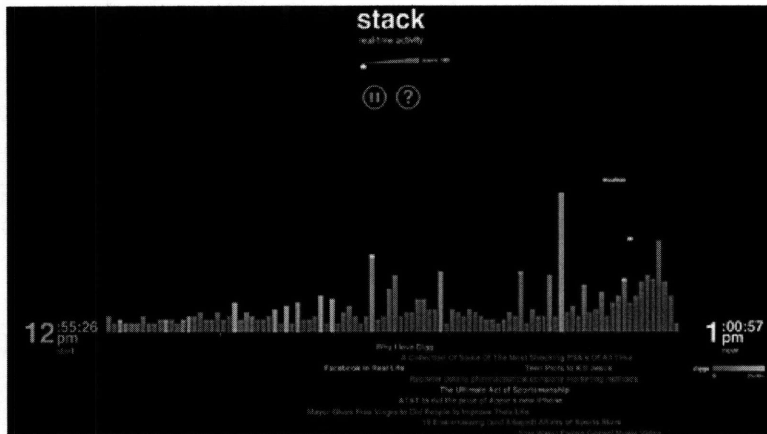


Figure 34. Screenshot from *Digg Stack* (Stamen Design).

This example again reiterates the slippery nature of aesthetics and affect suggested in the previous section. Though aesthetically pleasing visuals and affective interaction design are clearly influenced by a variety of components, it may not make sense to try to deconstruct them as such. Typically, from the perspective of the user, enjoyment is derived from interacting with these systems simply because they are “beautiful” or “fun.” These notions may connect to a definable set of tropes (the formalism of various schools of design certainly supports this idea) but it seems equally likely that the infinite variety of aesthetic experience, highly dependent on context, will always resist such categorization. Thus, in trying design aesthetically or affectively for a mass audience, we may simply need to accept that ambiguous level of definition – something that doesn’t sit well with quantitative-minded designers, but has certainly driven artistic practice since day one. Instead, in aiming for infovis that is both aesthetic and supporting of analysis, we should focus on balancing aesthetics with other considerations, such that one element does not overpower the other to the detriment of the overall design.

Social Infovis Design

The explosion of interest in the “social graph” and its related technologies (including information visualization) over the last few years, in addition to reaffirming that humans are, in fact, social creatures, has suggested another avenue from which to approach the design of popular visualization systems. Social media, whether in the form of social networking websites, blogs, or other Web 2.0-related implementations, have provided (at least) two compelling features to their users: first, the ability to simply connect with one another, at internet scale, around ideas, activities, artifacts, or otherwise, extends our analog social existence in to the digital space. Second, the same media technologies often facilitate more complex forms of collective intelligence, based on large scale social interactions, which can be leveraged as a means to foster increased insight in to the issues around which they operate.

With regard to the first function, as the cliché above suggests, the type of experiences these features provide fulfill some of our fundamental emotional needs and desires.

Patrick Jordan, reiterating the work of Lionel Tiger, identifies “socio-pleasure” as one of the “four pleasures” that we experience in our interaction with the world (2000, p. 13):

[Socio-pleasure] is the enjoyment derived from relationships with others. This might mean relationships with friends and loved ones, with colleagues or with like-minded people. However, it might also include a person’s relationship with society as a whole – issues such as status and image may play a role here.

In this sense, though addressed separately here, social experiences are highly affective. Social media designers have already discovered this in droves, but this type of design is only starting to be applied to information visualization systems, particularly ones targeted at non-expert users. Some of the most popular websites on the internet today are either explicitly about socialization, or incorporate socially designed features as a fundamental part of the user experience. Social networking sites like Facebook and MySpace, and services such as Second Life, represent the former category; these sites have amassed tremendous followings simply by virtue of the fact that they provide a space for users to be social and interact with one another. They are fundamentally about supporting the type of behavior that we thrive on in the “real world” at internet scale. Sites such as YouTube, Flickr, or del.icio.us provide non-social services (aggregation of videos, photographs, and hyperlinks), but are greatly enhanced by social elements focused around user-generated content, such as the sharing, tagging, and commenting of user content hosted by the services. The phenomenon of blogging also falls in to this category, in the way that it has re-conceptualized the authoring and publication of content as a conversation rather than one-way dissemination. Again, one of the fundamentally important elements of this kind of social design is the way in which it models our usual social relationships; users are given an expressive identity that simultaneously distinguishes them from, and serves as the basis for their connection to, the rest of society (in this case, the services and their other users). This, in turn, humanizes the technological mediation of information.

The second function, the facilitation of collective intelligence, is also widely supported by other types of social media on the web. The most explicit example of this would be Wikipedia, which literally uses the combined knowledge of its contributors to develop a comprehensive, digital encyclopedia. On a smaller scale, collaborative tools such as Google Docs allow users to simultaneously edit the contents of a text document or spreadsheet. Less formally, the ubiquitous blog or forum-style commenting systems described above have been deployed not only on blogs, but also attached to stories on news sites, product pages on commerce sites (Amazon, Best Buy, etc.), and almost anywhere else where a user might want to contribute an opinion. Often, the discussion generated in these forums has the potential to act as a knowledge-aggregation system; while individual participants may have only a partial understanding of the topic being discussed, the combined knowledge of many participants, each contributing their own observations and expertise to the discussion, can serve to produce a more comprehensive and nuanced description of the issue.

Similarly, the statistical aggregation of user choices in social media systems can provide a streamlined perspective on the content of those systems: commerce sites like Amazon, in addition to providing space for comments, also employ powerful recommender systems that help users find the products they're interested in based on the choices of previous users (i.e. Amazon's "People who liked *X* also liked *Y*"). Social news sites like *Slashdot.org* or *Digg.com* provide similar functionality as a means to sort news stories and user commentary by popularity. Once again, ephemeral real-world social activities (the exchange of recommendations between friends, for instance) are scaled up and re-mediated by the internet. When formally codified and indexed by social applications, these sorts of behaviors add value not only through the process of discussion and deliberation, but also as useful archived artifacts that accumulate around a topic of interest. They become a type of "discourse history" that can be valuable to future users (see Wexelblat and Maes on "history-rich footprints" (1999)). This type of functionality seems well suited to infovis design; embedding past analysis within the application itself could provide compelling benefits to long-term use.

Visualization has often worked in service of other social media, typically in the context of visualizing social networks. *visualcomplexity* currently lists 69 such projects (the most of any category), including several network graphs associated with popular websites Facebook and Flickr, as well as many examples that visualize personal email collections. However, while these tools allow users to explore their "social graphs" in compelling new ways, it is usually the subjects of the visualizations that are "social," and not the visualizations themselves. Furthermore, while they might be embedded within a larger social media system, they typically do not embody its functionality; they are often presented as a view of such systems rather than as an interface to them. How, then, can social design be applied to visualization systems themselves, regardless of the data they represent, to capture both the fun and added insight of social data exploration?

Much of the current thought on this question has arisen from the observation of public "free form" visualization usage, which is significant in itself, given the traditional model of infovis design and evaluation that usually restricts its exposure to small-scale research use and focused usability testing. Martin Wattenberg's *NameVoyager*, which visualized the historical popularity of baby names, became an early successful example of public-facing infovis when it spawned an unexpectedly large amount of conversation across the web—even amongst users not expecting children—that often explicitly referenced specific slices of data within the visualization (Viégas and Wattenberg, 2006). Utilizing unassociated social media (primarily blogs), users were compelled to share their discoveries, propose explanations for trends in the data, and pose data-driven challenges to one another. Fernanda Viégas, Jeffrey Heer and Danah Boyd describe similar types of interaction around visualizations of email and social networks: Viégas notes the unexpected degree to which users were inclined to share visualizations of private email (Viégas and Wattenberg, 2006), and Heer and Boyd observed *NameVoyager*-style social interactions in the context of local, face-to-face collaboration in front of a single instance of a social network visualization (2005). What each of these examples imply, and what was observed explicitly in the latter case, is that users are more inclined to interact with a visualization, and spend more time with it, when it is a shared or collaborative

experience. However, this fairly obvious conclusion is not at all specific to infovis systems, but has already been borne out in the wide array of social media design described above. The social visualization research that has resulted from these early observations further confirms this observation – many subsequent social infovis design considerations are either generalized, application agnostic guidelines, or suggested methods of implementing such guidelines in the context of an interactive visualization system (Heer and Agrawala, 2007). This is actually quite comforting, as it suggests that social infovis design is perhaps more intuitive and straightforward than it might seem – after all, new social media sites are popping up almost daily!

That being said, the intersection of social media and visualization remains relatively unexplored. However, in the last two years several public-facing projects have emerged, the two most prominent being the start-up *Swivel* (<http://www.swivel.com>) and IBM's *Many Eyes* (<http://www.many-eyes.com>). Though they take slightly different approaches to social visualization—*Swivel* seems to focus more on being a data repository with relatively simple visualization functionality, while *Many Eyes* provides more robust visualization and socialization features with less focus on data aggregation – both support a wide variety of the functionality referred to above. *Many Eyes* in particular, being the brainchild of Viégas and Wattenberg, stands as the current benchmark for socially designed information visualization in its incorporation of these social media standards at all levels of the system. At the most general level, it looks like a social media site: users log in with their own identity that gives them a sense of ownership over their contributions (data sets, visualizations, comments, etc., through the “My Stuff” page) and the experience of using the site. It is also clear from the front page that the content of the site is user driven, with the latest featured contributions on display. Drilling down to the content level, the social nature of the site is reinforced: though the content itself (data sets and visualizations) is somewhat unusual for a social media website, the surrounding conventions are all familiar. Blog-style comments are attached to each visualization and data set (private messaging is supported as well), as are links to “share this,” “watch this,” and “rate this.” Mirroring similar functionality on other social sites, these links allow you to embed a visualization in your own blog (otherwise known as “blog this”), track the discussion surrounding the content (through the “my stuff” page), and rate the content as a means to allow for interesting material to become more prominent on the site. Finally, and most importantly, these standard conventions are extended in to the visualizations themselves: facilitating what Viégas and Wattenberg call “deep sharing,” (Viégas and Wattenberg, 2006) comments attached to a visualization can be associated with a specific view of the data that reflects any interaction the commenter had with it (filtering, zooming, etc.). This enables the kind of linking, or “pointing,” that is integral to the process of information foraging (Pirolli and Card, 1999) and referencing that constitutes collaboration and discussion on the web, thereby encouraging the same type of critical analysis of visualization as is applied to more standard information channels.

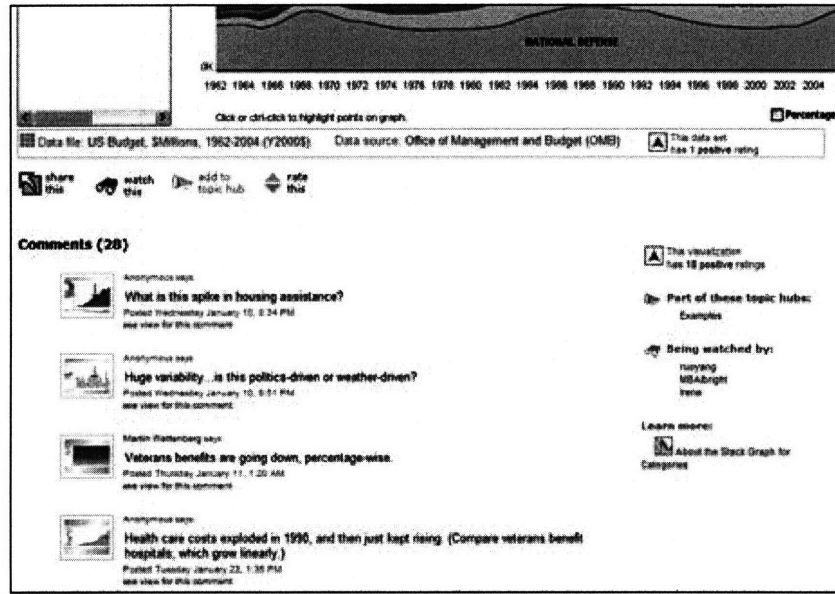


Figure 35. Screenshot (detail) from the *Many Eyes* website showing the social media features associated with each visualization (IBM).

This final point is worth reflecting on in its significance to bringing visualization to the public, as it arguably represents the lynchpin to the crucially important function that *Many Eyes* performs as a whole. In the previous discussion of the semantics of information visualization, it was suggested that one of the fundamental issues facing non-expert infovis users is one of interpretation. Unfamiliarity with the visual languages being employed in infovis design leaves users wondering what it is that they're looking at. By the same token, it's important to remember that the same type of confusion exists at a higher level that concerns the context in which information visualization is encountered. In the current experimental phase of public-facing visualization design, infovis often appears with very little meaningful context. Individual designers post their work on personal web sites, design firms post them on portfolio pages, or visualizations appear as novelty side-bar items on a variety of web sites without much qualification. In a general sense, this results in a disjointed experience of information visualization that potentially obscures both its purpose and its accessibility. In contrast, casting visualization as social media provides an instant framework to contextualize an encounter. Though information visualization often presents as an utterly unfamiliar interface, interaction with infovis systems can be made to fit existing, recognizable models of interaction deployed in more familiar contexts. By couching visualization use and analysis within the trappings of ubiquitous social media systems, a context is defined that might make sense to an otherwise overwhelmed user. Significantly, *Many Eyes* succeeds in this regard because of its thoroughness. While it would be easy to contextualize infovis (or anything else) as social media in a very superficial way, the incorporation of social media affordances down to—and within—the visualizations themselves results in a genuinely familiar experience that takes advantage of existing user skills and lowers the barriers to entry. In using social media as a vehicle, it also reiterates the notion that infovis is something we can experience with the same degree of social and cultural

fluency as we do more traditional blogs and social networks. In other words, it suggests that information visualization really is for the people.

Narrative Infovis Design

The final design paradigm discussed here involves the conceptualization of information visualization as a form of narrative. It addresses the longstanding but often ignored suggestion that data implicitly represents some kind of story that infovis provides an opportunity to tell. This, of course, follows from a large body of thought that has addressed the central role of narrative underlying the ways we exchange information in society more generally, but which is beyond the scope of this discussion. Nevertheless, while the design strategies explored thus far mostly focus on improving the visual display of information (by highlighting issues of visual salience, affect, and usability in an attempt to improve intuition and engagingness), they have only indirectly addressed a second important function of infovis for the people, which is the explanation of the data itself. It is one thing to be able to read a visualization and enjoy interacting with it, but this type of experience is arguably only relevant, or even possible, when the significance of the data being represented is clear to the user. Providing context for the data is important, and becomes increasingly relevant as instances of information visualization become more robust and self-contained. As the model of the infographic as supporting element to a written story or article is supplanted by one in which complex, interactive, technology-enabled infovis applications are the stories themselves, they can no longer rely on external explanatory text to describe the meaning of the data they represent. Being able to convey that context through the visualization itself will become a significant component in the design of legible public-facing infovis, as a means to increase both the coherence and memorability of the data being conveyed.

It is important to note that one of the likely reasons why this line of design has remained largely unexplored in traditional infovis research is because of the distinction often made between “presentational” and “exploratory” information visualization. Presentational infovis refers to visualizations that represent information that the designer already understands; they are typically pre-selected slices of the data that show a trend or feature that the designer is trying to convey to the user. Most traditional infographics that appear in newspapers are examples of presentational infovis, and so far have best incorporated elements of narrative structure. Exploratory information visualization refers to visualization systems that are designed to operate on raw data, and are used to explore a data set to discover trends or features within it for the first time; they are also generally more robust in their functionality for filtering or reorganizing the data display to facilitate this exploration process. Most traditional research-driven infovis falls in to this category. Accordingly, this type of visualization has received more attention, and in its characterization as a research tool has not been the focus of much explicit narrative design. However, the distinction between these two types of infovis is starting to become blurred. In many ways, it is a distinction based on scale; exploratory visualizations have traditionally operated on large unwieldy data sets, while presentational ones focused on

smaller “presentable” chunks. The technological affordances of modern infovis, providing the ability to handle more data interactively, make the notion of presenting larger amounts of information less difficult. When the data being presented is scaled up, and functionality is provided to search through it efficiently, it effectively becomes exploratory. Similarly, in situations where time-sensitive live data needs to both be explored and presented immediately (stock market data, for instance), these styles of representation begin to converge. Finally, we can imagine future infovis designs in which the incorporation of social elements, such as the tagging of specific views on the data as featured on the web-based social visualization service *Many Eyes*, use user exploration itself as a means to build up a presentational structure around the data.

So, accepting that most infovis systems can support a narrative component, how might such a component be designed, and what value does it add? Two recent very public deployments of information visualization have explicitly married data display to storytelling: Al Gore’s use of information visualization in his film *An Inconvenient Truth*, and health expert Hans Rosling (of *Gapminder.org*) in his TED Conference presentations on global health statistics. Gore uses visualization not only to support his argument, but as a central feature of the film, literally using the trends shown in his data graphics to contextualize his warnings about global warming. Though his use of a crane to emphasize the severity of a trend in one of his graphics might be overly dramatic or rhetorical, it nevertheless helps to activate the ability of the graphic to speak to the significance of the data it represents.

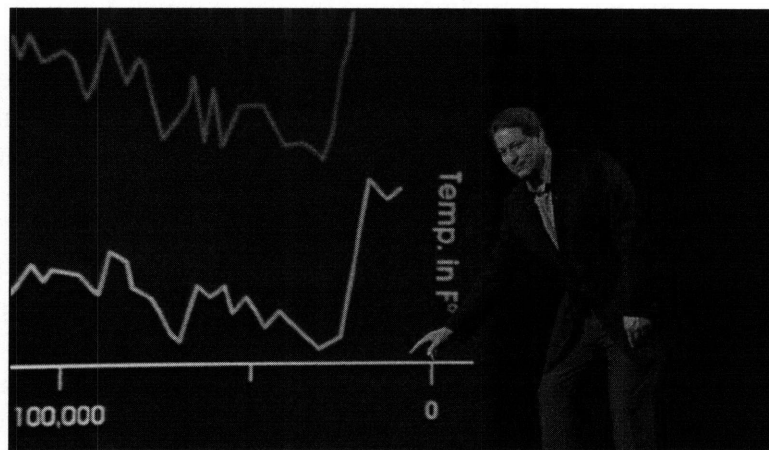


Figure 36. Al Gore presents data in *An Inconvenient Truth* (2006).

Rosling similarly motivates his data through lively narration of visualizations created with Gapminder’s (now Google’s) *Trendalyzer* software. As data points representing mortality rate and per capita income of various countries over the past 100 years move around the graph, Rosling’s hyper-active sports-like commentary provides additional context and draws the viewers’ attention to specific trends occurring over the course of the animation. The meta-data and “meta-presentation” that his commentary provides helps engage his audience by drawing connections between different elements of the visualization.

Incorporating this type of commentary in to a visualization system without the presence of a commentator poses challenges, but is certainly feasible. Traditional designers, often producing graphics for widely distributed newspapers, have often relied on the inclusion of explanatory text within their graphics as a means to guide the user through the data, but have always had to work around the space constraints of a static image. The advent of digital infovis erases this limitation, allowing for annotations to be included as a toggle-able layer on top of the rest of the graphic, or implemented as pop-up messages that appear when the user moves the mouse cursor over the relevant parts of the graphic. Reiterating the need to guide users (readers) through a potentially complex visualization (Ericson, 2007), and reflecting the recent “computational journalism” movement, the New York Times have successfully employed this technique in a variety of their recent interactive graphics: In “Paths to the Top of the Home Run Charts,” a graph plotting the total number of homeruns versus age of the top hitters in the MLB, text inserts are included to explain certain anomalous data (such as Ted William’s apparent three year homerun slump, explained by the fact that he had left baseball to fight in World War 2 during those years). In “Casualties of War,” a robust infovis system tracking deaths and related statistics associated with the Iraq War, a similar method of annotation is used. On the histogram counting the number of deaths per day, certain days are tagged with additional information that explains the circumstances of those particular deaths, appearing when the user places the mouse over the appropriate histogram bar (see Figure 38).

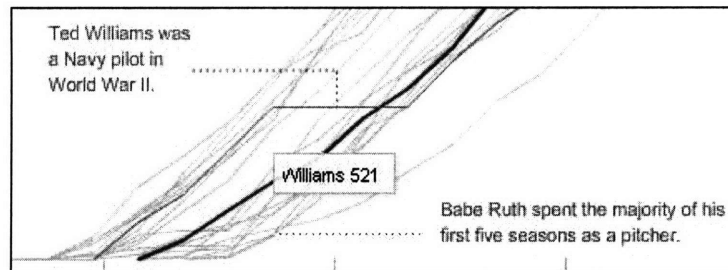


Figure 37. Screenshot (detail) from *Paths to the Top of the Home Run Charts*, an interactive graph of the home run progress for top hitters in the MLB annotated with descriptive text (New York Times).

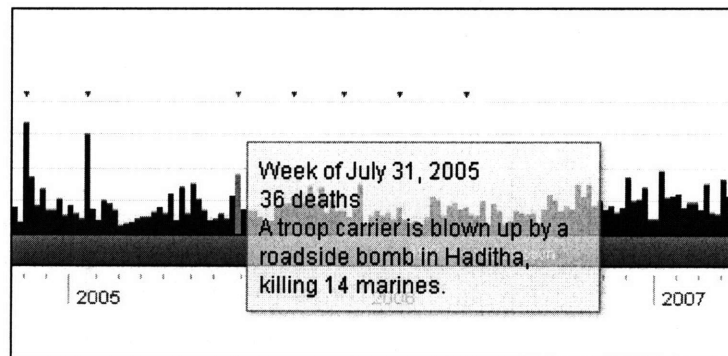


Figure 38. Screenshot (detail) from *Casualties of War*, and interactive visualization showing American troop deaths associated with the Iraq War. Significant data points are annotated with descriptive roll-over text (New York Times).

Of course, text is not the only way to provide narrative annotations. Related to the use of icons and images as discussed earlier in the section on semantics, Rosling's *Dollar Street* visualization annotates family income histograms with photographs of the types of houses a given income can afford. What emerges is a very clear story about the distribution of wealth around the world that may not have the same impact as the statistical graphics alone.

This notion of annotation or “tagging” of data or data views again suggests the possibility of emergent narratives arising from the contributions of users in a socially designed infovis system. This already occurs on *Many Eyes*, where the functionality to comment on specific views of a visualization gives rise to discussion of the various properties of a data set. This discussion in turn becomes a narrative around the data visualization that not only enriches the informational content of the system, but serves as guide for its future users. Similar functionality was deployed (in a much more tongue-in-cheek fashion) by The Coca-Cola Company in its now defunct 2005 *WorldChill* visualization system. Marketed as a type of meteorological map for human emotions, the web-based system encouraged users to provide their location on a world map, along with the degree of “chill” (i.e. stress) that they were feeling at the time and a brief text message, which was then appeared as a colored dot on the map whose text would appear when moused-over. The result, though often absurd and nonsensical, was a visualization that sometimes exhibited narrative structure that was emphasized by the addition of the text messages. For example, episodes of high stress in the Boston area in October and December were perhaps inexplicable until a survey of user messages indicated that these months corresponded with midterm and final exams at the many colleges in the area.

Finally, different visualization techniques support narrative components in less explicit ways. The news aggregation site *Silobreaker* (<http://www.silobreaker.com>) uses semantic network graphs to contextualize specific news stories within a larger web of related topics. So, for instance, stories about individual presidential candidates are linked to those about other candidates, as well as to stories that cover campaign issues and other aspects of the election process. Users can use the initial article as an entry point in to a more nuanced, contextualized reading of the topic that allows for a higher level perspective on an issue. The network graph provides a sort of “big picture” narrative to the reading of news that is becoming increasingly valuable amidst growing concerns of media bias and the “echo chamber effect” on the web. We might imagine further extending this functionality by including the sorts of user-generated commentary discussed above, or by placing a greater emphasis on the timeline associated with the semantic network. As Rosling demonstrates when he characterizes his animated graph of changing health and economic conditions in the US, Japan, and Sweden as a “race” between a “yellow Ford,” “red Toyota,” and “brown Volvo,” time variation alone provides an implicit basis for contextual narrative. Returning to Stamen Design's *Digg Stack* visualization, we note that it is not only fun to watch real-time returns on the popularity of *Digg.com* stories, but the evolution of the data over time provides a basis on which to construct a story about the activity on the website that helps us make sense of the information being presented. For non-expert users, stories about data are likely to be more salient than their statistical characteristics.

CONCLUSION

Information visualization will undoubtedly play a major role in the way we communicate data in a society that is becoming increasingly digital. As the “standard rationale” suggests, the Information Age is characterized by a scale of information production that has not only dwarfed that of previous eras in a matter of years, but promises to only expand further in the wake of technological advancement. In response, visualization has established itself, in principle, as an effective counter-measure to this information overload – but in what context?

The preceding discussion has been an attempt to draw attention to the discrepancies within the theory and practice of information visualization design that are adversely impacting its broader reception. These discrepancies concern the ways in which infovis has been conceptualized as a means to attend to this information overload; the traditional utilitarian view reflected in the “standard rationale” conceives of visualization as a purely functional operation, designed as a tool with which rigorous analysts do rigorous analysis. The more recent aesthetic view, empowered by the increasing availability of flexible, user-friendly visualization design tools and techniques, conceives of infovis as an expressive art form. Various other views lie in between. The underrepresented perspective, however, is one that considers the corresponding changes in “data culture” that accompany the Information Age. As information becomes more ubiquitous in a variety of forms (though primarily as digital content delivered via the internet), we, as a general public, are becoming more comfortable with our ability to navigate these information spaces as part of our everyday lives: we use Google to search the web, we look for audiovisual content on sites like YouTube and Flickr, we make informed purchasing decisions based on sophisticated analysis of information-rich commerce sites such as Amazon, we exchange information and ideas across complex networks of blogs and news aggregators. These activities point to the development of a generalized information literacy deployed against the information deluge; as such, the Information Age is not merely about the imposing presence of information at internet scale, but also the way that we respond to it socially and culturally. Accordingly, we the people are ready to engage with information visualization as a means to mediate some of this information exchange—not necessarily as rigorous analysts or art appreciators, but as regular people with a continuing need to make sense of the information environment that we live in. Unfortunately, information visualization design isn’t quite ready to engage with us on these terms.

Interestingly, infovis often appears as the face of a technologically-enabled information culture: it features prominently in our representations of our future, distant or near, as seen in film and television (think of the now famous computer interface from the sci-fi film *Minority Report*, or the images that appear on computer screens in any number of technologically-oriented films). Perhaps more telling, it has become a marketable image, used to sell everything from iPhones to Coca-Cola.



Figure 39. Information visualization in *Minority Report* (2002).

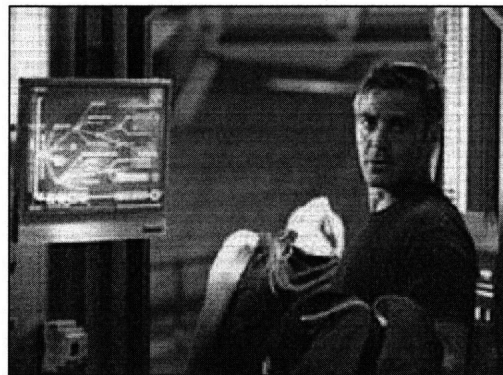


Figure 40. Information visualization in *Solaris* (2002).

And yet despite the fact that the technology to produce these types of visualizations exists today in very accessible forms, its appropriation in the design of real-world infovis that can be used by ordinary people as functional systems for navigating information lags far behind these media representations. In other words, we like the idea, or aesthetic, of ubiquitous information visualization, but we don't know how to use it in practice. While there is arguably an art-theoretical claim to be made about the role of artists in defining the aesthetic of new media technologies before their full impact can be felt (McLuhan, 1964, p. 9), does this view undermine the functional value of visualization? The discrepancy between the popular "image" of visualization and its potential function suggests itself as one in need of reconciliation. This discussion has been heavily critical of the traditional view of information visualization in its resistance to understanding infovis as an accessible aesthetic artifact, but this is not the type of infovis that often makes it on to our web browsers. As suggested earlier, the artistic appropriations of infovis are, in fact, increasingly defining its public face. Accordingly, as Stephen Few argues, it is just as important to be critical of them, so that the public perception of the affordances of infovis is not disproportionately influenced by its aesthetic appeal. The world is ready for information visualization that is simultaneously aesthetically compelling, analytically useful, and contextually meaningful. This synthesis needs to be reflected in the real-world applications that contribute to its popular image.

Towards this end, existing real-world examples of public-facing infovis need to be studied in the contexts in which they appear, including the work that comes out of non-

academic channels, in order to better understand how infovis functions as a popular artifact. The current fragmentation of the infovis design community has quarantined the critical analysis of visualization systems in a way that severely limits what can be learned from their design. In academic circles, the “official” infovis community primarily discusses bleeding edge implementations that rarely exist outside of the research lab. As such, the discussion is typically limited to self-evaluation by the designers according to standardized evaluation parameters that reflect a level of specificity that may not be relevant to a system’s use in a real-world context. As Kosara suggests, this is a view that focuses on “implementation details and single mouse clicks,” rather than the “meaning of a visualization” (Kosara, 2007). On the other hand, the vast majority of public infovis on the web is the product of individual designers or design firms working outside of any context that supports formalized critical analysis of their products. Additionally, much of this work is either presented as “art,” and thus considered unsusceptible to criticisms of its utility, or determined to be outside the scope of serious infovis, and therefore not worthy of analysis.

Nevertheless, these real-world examples are the only ones that are potentially reaching a wide audience, and therefore represent the only data we have on the popular experience of information visualization. Recent discussion surrounding the release of Stamen Design’s *Twitter Blocks*, for example, provides a compelling look in to how one form of high-profile infovis was received (Arrington, 2007, Carden, 2007, and Migurski, 2007). While the discourse has largely played out in the form of the designers defending the intent of their work against (mostly unconstructive) blog-comment criticisms, there is a lot of interesting information to be gleaned from it. For instance, it suggests quite a bit about the impact of user expectations on their experience of an infovis system, regardless of designer intent. How do Twitter users expect a Twitter visualization to function? How do non-Twitter users interpret it? How do its aesthetics influence this? To what degree do its stylistic affordances encourage or discourage use? What about its utilitarian ones?

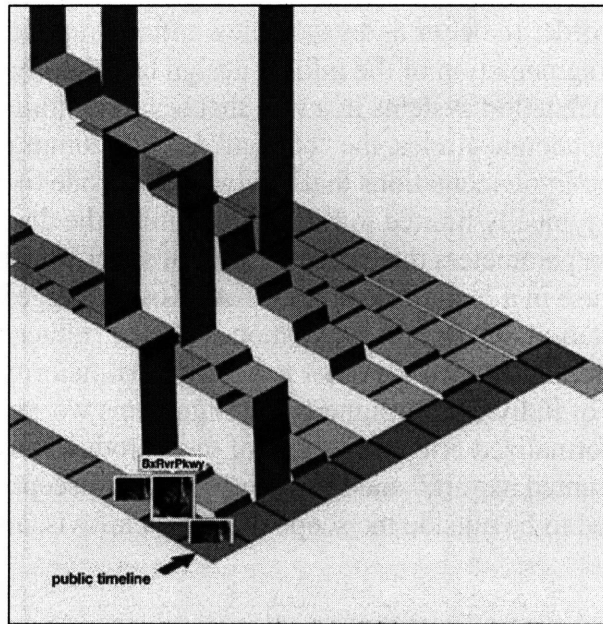


Figure 41. Screenshot (detail) of *Twitter Blocks* (Stamen Design).

There are broader questions that can be asked about public infovis in the wild: how does the degree of integration with other site content (including physical placement of visualization links on the page) impact its use? Is there a certain level of functionality that needs to be made available (and obvious) to a user in order to convince them to interact with the visualization when the content can be retrieved through other channels? These types of questions cannot be adequately addressed when analyzing an infovis system outside of its natural habitat. Accordingly, the answers to these questions, and the design successes and failures they suggest, cannot inform future design if infovis systems that embody them remain unanalyzed. In other words, in order to understand how popular infovis works in the world, we need to study popular infovis in the world. Because the people who think formally about information visualization are not always the ones creating instances of it that other people actually use, we need to expand our notions of who—and what—constitutes infovis design. In surveying the current body of public-facing infovis work on the web, it is clear that the knowledge of each group can inform the designs of the other.

Not surprisingly, this discussion has only scratched the surface of the issues surrounding information visualization design for the people. The goal here was to provide an overview of the state of information visualization today, and describe some of the central issues revolving around its use in a public, non-expert context. However, there are a number of related directions—some of which are already being actively pursued in the design community—that this type of analysis could move towards:

Further reflection on the interdisciplinarity of infovis design. Information visualization, as a field, is often described as being highly interdisciplinary, requiring both technical expertise and creative ability, and whose development is informed by work in a variety of domains, from graphic design to cognitive psychology. While this is certainly true

theoretically, is perhaps less true in practice. First, there are certainly more related fields that could contribute to a robust understanding of infovis theory and practice; in this discussion, there has been a focus on incorporating principles from the realm of user-centered design (itself a somewhat vague and interdisciplinary field) in to our thinking about infovis design, but there are more. A media studies perspective, for example, could bring new insight in to the function of infovis as a representational system. Also, as has been suggested, the contributions from the “aesthetic” dimension of infovis design need to be better included in its design discourse at a level more practically useful to designers than the art theoretical one.

Second, it is apparent that these disciplines are rarely well integrated in to an overarching theory of infovis design, insofar as such theory exists. The push for a more holistic understanding of information visualization design and reception is an important one, as it would emphasize the necessity of connecting specific disciplinary contributions to one another. Specificity makes research easier, but the practical usefulness of that research is increased when it can be reconciled with—or at least address—the larger body of work that defines the field. For instance, how do we reconcile cognitive psycho-physical findings on color perception with an aesthetic color theory? Perhaps a trivial example, but it is meant to reflect the ways in which apparently conflicting findings from different disciplinary approaches often remain in conflict. Without meaningful qualification, these discontinuities offer little practical value. This is further reflected in the tendency of the field to create taxonomies as a means of cordoning off design practices that are ostensibly outside the scope of particular disciplinary frameworks. A holistic view suggests that all practices are part of a more broadly defined conception of information visualization, and should be engaged as such. Further thought in to how this could be accomplished seems appropriate.

Investigation of new evaluation techniques for information visualization. Though only briefly suggested in this discussion, the current standard methods of infovis usability evaluation are of questionable relevance to the design of visualization systems for the public. While measurements of efficiency and effectiveness are certainly important to the development of infovis systems, their valuation arguably implies a type of focused, expert-level interaction characteristic of utilitarian software. While this is appropriate for utilitarian infovis used in research or business applications, it does not represent what interaction with public-facing information visualization looks like. An average user encountering a visualization while browsing a news website during their lunch break is not the same as a network administrator trying to make sense of traffic through the company intranet. Accordingly, their expectations and experiences are not the same.

In order to better understand how users experience infovis systems, and better capture the relevant design components that contribute to that experience, it seems appropriate to consider a more ethnographic approach to usability testing. As has been demonstrated by the research surrounding examples such as Wattenberg’s *NameVoyager* or Heer’s *Vizster* (Viégas and Wattenberg, 2006, Heer and Boyd, 2005), understanding how users engage with visualization in a more natural context (in the “real world”) leads to insights that are rarely discoverable within the controlled context of a usability lab. This is, of course,

easier said than done. Ethnography requires that there be a well defined population to study, and public visualization is not yet mainstream enough to support one (for instance, there is no YouTube for information visualization). Building in systems that automatically record detailed usage statistics is one possible direction, but even detailed statistical data often fails to capture the nuances of software use. In any case, further thought is needed to figure out how we can leverage ethnographic methods to improve infovis design.

Discussion of the role of education with respect to visualization literacy. This is arguably one of the most important components of an effort to popularize information visualization. As was briefly suggested earlier, a major difficulty in presenting complex information visualization to non-expert audiences is the gap between the semantics of modern infovis and those that are employed in the design of comparatively simple charts and graphs. Though literacy for these simple forms (“graphicacy”) is taught at a young age in school, it is not only treated very superficially, but is based on outdated conceptions of how even simple infographics are designed. This results in situations like the previously described anecdote from the New York Times graphics team, where basic graphical literacy cannot be counted on.

The ability to read (and write) information visualization is very much a new media literacy whose relevance is obvious as we move forward in to the Information Age. As our sci-fi movies suggest, the types of visualization being developed today will become the basis for much of our information exchange in the future. These are digital artifacts – animated, interactive, and potentially massive in terms of the scale of information they represent – and need to be taught as such. “Time goes on the horizontal axis and the variable goes on the vertical axis” doesn’t cut it anymore, and while we can count on “digital natives” to absorb some of the required knowledge through technological familiarity, this new literacy needs to be explicitly addressed in the classroom. It is also extremely important to go beyond functional literacy, and address critical literacy of visualization as well. As discussion around *Many Eyes* has shown, the increasing popularity and “democratization” of visualization as an information channel brings with it concerns of misuse. Though the formality of information graphics has often been used to imply accuracy or truthfulness in the data they represent, we are now starting to recognize that such graphics are just as susceptible to manipulation as any other form of communication. Teaching critical reading of visualization will therefore become just as important as teaching critical reading of text. Of course, all of this also easier said than done, particularly in an era of dwindling educational budgets and curriculum concerns, but it cannot be ignored. Furthermore, it is the responsibility of the information visualization community, in designing visualization for the people, to help establish what graphicacy should look like in the Information Age. This will require continued exploration and discussion.

In conclusion, the field of information visualization is on the cusp of a transition to mainstream relevance that will not only revitalize the field itself, but potentially revolutionize the way society as a whole engages with information on a scientific, social,

and cultural level. It now represents a sufficiently matured means of communication that should no longer be characterized as “by experts, for experts.” As has been shown, the pieces necessary to popularize information visualization are already in place; the necessary technology and visualization expertise exists and is readily available. Accordingly, this discussion has presented very few, if any, genuinely new theoretical or methodological design paradigms, but has rather attempted to gesture towards an alignment of existing approaches as a means to conceptualize and facilitate the design of *information visualization for the people*.

LIST OF FIGURES

1. Playfair, William. "Chart Showing at One View the Price of the Quarter of Wheat & Wages of Labour by the Week," 1786. From "Milestones in the History of Thematic Cartography, Statistical Graphics, and Data Visualization: 1700 – 1799." Accessed 5 May 2008.
<<http://www.math.yorku.ca/SCS/Gallery/milestone/sec4.html>>
2. Screenshot from 1992 Ross Perot television ad. Retrieved from [YouTube.com](http://www.youtube.com), 20 April 2008. <http://www.youtube.com/watch?v=ERlGndQ_xtM>
3. Histogram of energy use. [NSTAR](http://www.nstar.com). Accessed 20 April 2008.
<<http://www.nstar.com>>
4. "All of Inflation's Little Parts," [New York Times](http://www.nytimes.com), 3 May 2008.
<http://www.nytimes.com/interactive/2008/05/03/business/20080403_SPENDING_GRAPHIC.html>
5. Levin, Golan, et al. "The Dumpster," 14 Feb. 2006. [Whitney Artport](http://artport.whitney.org).
<<http://artport.whitney.org/commissions/thedumpster/>>
6. "Ambient Orb," [Ambient Devices](http://www.ambientdevices.com). Accessed 2 May 2008.
<<http://www.ambientdevices.com/cat/orb/orborder.html>>
7. Minard, Charles. "Napoleon in Russia," from Tufte, 1983.
8. Screenshot from "Tableau," [Tableau Software](http://www.tableausoftware.com). Accessed 25 April 2008.
<<http://www.tableausoftware.com/>>
9. Screenshot from "Spotfire," [TIBCO](http://www.tibco.com). Accessed 25 April 2008.
<<http://spotfire.tibco.com/index.cfm>>
10. Dragulescu, Alex. "malwarez," Personal website. Accessed 25 April 2008.
<<http://www.sq.ro/malwarez.php>>
11. On, Josh. "They Rule," [theyrule.net](http://www.theyrule.net). Accessed 25 April 2008.
<<http://www.theyrule.net/>>
12. Screenshot from "Digg Arc," [Stamen Design](http://labs.digg.com). Accessed 25 April 2008.
<<http://labs.digg.com/arc>>
13. "Things Meat Loaf Would Do For Love," designer unknown. Accessed 2 May 2008.
<http://www.thehumorarchives.com/joke/Things_Meat_Loaf_Would_do_for_Love>

14. "Relationship Between Money and Problems," designer unknown. Accessed 2 May 2008. <<http://jamphat.com/rap/>>
15. "America's Most Popular Charts," 17 Jan. 2007. The Onion. <http://www.theonion.com/content/statshot/americas_most_popular_charts>
16. Silobreaker. Silobreaker, Ltd. Accessed 2 May 2008. <<http://www.silobreaker.com>>
17. visualcomplexity.com. Accessed 25 April 2008. <<http://www.visualcomplexity.com>>
18. Examples of "Isotype," Isotype Institute. Accessed 12 Jan. 2008. <<http://www.fulltable.com/iso/>>
19. Screenshots of Macintosh and Windows Desktops, from "History of the graphical user interface," Wikipedia. Accessed 5 May 2008. <http://en.wikipedia.org/wiki/History_of_the_graphical_user_interface>
20. Diagram of conceptual models in design, from Norman, 2004, p. 76.
21. Mondrian-inspired visualization of bus schedules, from Skog, et al., 2003.
22. Diagram of the visual variables, from Bertin, 1983, p. 43.
23. Screenshot from "Map of the Market," SmartMoney.com. Accessed 3 May 2008. <<http://www.smartmoney.com/marketmap/>>
24. Example of pie chart and tag cloud, images accessed 2 May 2008 from "Google Analytics," <<http://www.google.com/analytics>>, and "jill/txt," <<http://jilltxt.net/?m=200610>>
25. Lins, Adriana and Alicia Cheng. "31 Days in Iraq," 6 Feb. 2006. Op-chart, New York Times. <http://www.nytimes.com/2006/02/06/opinion/06chart.html>
26. Warren, Jeffrey. "ARMSFLOW," armsflow.org. Accessed 12 Jan. 2008. <<http://www.armsflow.org>>
27. Harris, Jonathan. "Choose Your Weapon," International Networks Archive. Accessed 12 Jan. 2008. <<http://www.princeton.edu/~ina/infographics/weapons.html>>
28. Ropeik, David and Nigel Holmes. "Never Bitten, Twice Shy: The Real Dangers of Summer," 9 Aug. 2003. Op-ed, New York Times.

- <http://www.nytimes.com/2003/08/09/opinion/09OPAR.html?ex=1375848000&en=77d4b622b1181cf8&ei=5007&partner=USERLAND>>
29. Comparison image of Edward Tufte's critique of the iPhone. 25 Jan. 2008. Graphpaper.com. <http://www.graphpaper.com/2008/01-25_edward-tuftes-iphone>
 30. "Categories within the model of information aesthetics," from Lau and Vande Moere, 2007.
 31. Screenshot from "Visual Thesaurus," Thinkmap. Accessed 2 May 2008. <<http://www.visualthesaurus.com/>>
 32. Google Analytics. Google. Accessed 2 May 2008. <<http://www.google.com/analytics>>
 33. Screenshot from "VisitorVille," World Market Watch, Inc. Accessed 5 May 2008. <<http://www.visitorville.com/>>
 34. Screenshot from "Digg Stack," Stamen Design. Accessed 25 April 2008. <<http://labs.digg.com/stack/>>
 35. Screenshot from "Many Eyes," IBM. Accessed 6 May 2008. <http://services.alphaworks.ibm.com/manyeyes/view/S62J0EQ9mVa6O_kEoi71E2~>>
 36. Screenshot from An Inconvenient Truth. Dir. Davis Guggenheim with Al Gore. Lawrence Bender Productions, 2006.
 37. Screenshot from "Paths to the Top of the Home Run Charts," 31 July 2007. New York Times. <http://www.nytimes.com/ref/sports/20070731_BONDS_GRAPHIC.html#>
 38. Screenshot from "Casualties of War," 28 Dec. 2006. New York Times. <http://www.nytimes.com/ref/us/20061228_3000FACES_TAB2.html>
 39. Screenshot from Minority Report. Dir. Steven Spielberg. 20th Century Fox, 2002.
 40. Screenshot from Solaris. Dir. Steven Soderbergh. 20th Century Fox, 2002.
 41. Screenshot from "Twitter Blocks," Stamen Design. Accessed 25 April 2008. <<http://explore.twitter.com/blocks/#>>

WORKS CITED

Arrington, Michael. "New Twitter Visualization Tools coming: First is Twitter Blocks," Weblog entry, TechCrunch. 30 Aug. 2007.

<<http://www.techcrunch.com/2007/08/30/new-twitter-visualization-tools-twitter-blocks-on-friday-morning/>>

Benway, J. P., and D.M. Lane. "Banner Blindness: Web Searchers Often Miss 'Obvious' Links," Internetworking: ITG Newsletter, Dec. 1998.

<http://www.internettg.org/newsletter/dec98/banner_blindness.html>

Bertin, Jacques. Semiology of Graphics: Diagrams, Networks, Maps. Madison, WI: University of Wisconsin Press, 1983.

Burkhard, Remo, et al. "Visualization Summit 2007: Ten Research Goals for 2010," Information Visualization. Oct. 2007: 169-188.

Card, S. K. and J. Mackinlay. "The Structure of the Information Visualization Design Space." Proceedings of the 1997 IEEE Symposium on Information Visualization. Oct. 1997: 92-99.

Card, Stuart K, J. MacKinlay, B. Shneiderman (editors). Readings in Information Visualization: Using Vision to Think. San Diego: Academic Press, 1999.

Carden, Tom. "Criticism for Twitter Blocks," Weblog entry, Random Etc. 2 Sep. 2007. <<http://www.tom-carden.co.uk/2007/09/02/criticism-for-twitter-blocks/>>

Cawthon, N. and A. Vande Moere (2007), "The Effect of Aesthetic on the Usability of Data Visualization," IEEE International Conference on Information Visualization (IV'07). 2007: 637-648.

Chevallier, V. "Notice necrologique sur M. Mindard, inspecteur general des ponts et chaussées, en retraite," Annales des ponts et chaussées. 2, 1871: 1-22. Translated by Dawn Finley. <<http://www.edwardtufte.com/tufte/minard-obit>>

Ericson, Matthew. "Visualizing Data for the Masses: Information Graphics at the New York Times," IEEE Symposium on Information Visualization, 2007. Sacramento, CA. 28 Oct. 2007.

Ferreira, Jennifer, James Noble, and Robert Biddle. "A Case for Iconic Icons," Proceedings of the 7th Australasian User Interface Conference – Volume 50. 2006: 97-100.

Few, Stephen. "InfoVis as Seen by the World Out There: 2007 in Review," IEEE Symposium on Visual Analytics Science and Technology, 2007. Sacramento, CA. 1 Nov. 2007.

Fishwick, Paul. "Software Aesthetics," preprint. Feb. 2008.
<<http://www.cise.ufl.edu/~fishwick/aescomputing/manuscript.pdf>>

Gitelman, Lisa. Always Already New: Media, History and the Data of Culture. Cambridge, MA: MIT Press, 2006.

Heer, J. and D. Boyd. "Vizster: Visualizing Online Social Networks," IEEE Symposium on Information Visualization, 2005. Oct. 2005: 32-39.

Heer, J. and Maneesh Agrawala. "Design Considerations for Collaborative Visual Analytics," Proceedings of the 2007 IEEE Symposium on Visual Analytics Science and Technology. Nov. 2007: 171-178.

Holmquist, L. E. "Evaluating the comprehension of ambient displays," Extended Abstracts of CHI 2004. 2004: 1545.

"InfoVis for the Masses," Panel discussion. IEEE Symposium on Information Visualization, 2007. Sacramento, CA. 28 Oct. 2007.

Isen, A.M. "The influence of positive affect on cognitive organization: some implications for consumer decision making in response to advertising," Advertising Exposure, Memory, and Choice. Ed. Mitchell, A.A. London, UK: Lawrence Erlbaum, 1983.

Jankun-Kelly, T.J., Robert Kosara, Gordon Kindlemann, Chris North, Colin Ware, and E. Wes Bethel. "Panel: Is There Science in Visualization?" IEEE Visualization 2006 Conference Compendium. Oct. 2006.

Jordan, Patrick. Designing Pleasurable Products: An Introduction to the New Human Factors. London, UK: Taylor & Francis, 2000.

Kelly, James D. "The Data-Ink Ratio and Accuracy of Newspaper Graphs," Journalism Quarterly. 1989: 632-639.

Khaslavsky, Julie and Nathan Shedroff. "Understanding the Seductive Experience," Communications of the ACM. May 1999: 45-49;

Kosara, Robert. "Visualization Criticism – The Missing Link Between Information Visualization and Art," Proceedings of the 11th International Conference on Information Visualization (IV), 2007. 2007: 631–636.

Lau, A. and A. Vande Moere (2007), "Towards a Model of Information Aesthetic Visualization," IEEE International Conference on Information Visualization (IV'07). 2007: 87-92.

Lorensen, B. "On the Death of Visualization," Position Papers NIH/NSF Proc. Fall 2004 Workshop Visualization Research Challenges. Sept. 2004.

Manovich, Lev. Info-Aesthetics. 27 Oct. 2001. <<http://www.manovich.net/IA/>>

McLuhan, Marshall. Understanding Media: The Extensions of Man. 1964. Cambridge, MA: MIT Press, 1994.

Migurski, Michal. "uselessness," Weblog entry, tecznotes. 1 Sep. 2007. <<http://mike.teczno.com/notes/uselessness.html>>

Neurath, Otto. "Visual Education: A New Language," Survey Graphic. Jan. 1937: 25.

Norman, Donald. "Human-Centered Design Considered Harmful," Aug. 2005. jnd.org. Accessed 21 April 2008. <<http://www.jnd.org/dn.mss/human-centered.html>>

Norman, Donald. Emotional Design: Why we love (or hate) everyday things. New York, NY: Basic Books, 2004.

Pirolli, Peter and Stuart K. Card. "Information Foraging," Psychological Review. 106(4), 1999: 643-675.

Plaisant, Catherine. "The Challenge of Information Visualization Evaluation," Proceedings of the Working Conference on Advanced Visual Interfaces. 2004: 109-116.

Pousman, Z., J. Stasko, and M. Mateas. "Causal Information Visualization: Depictions of Data in Everyday Life," IEEE Transactions on Visualization and Computer Graphics. Nov. 2007: 1145-1152.

Prensky, Marc. "Listen to the Natives," Learning in the Digital Age. Dec. 2005: 8-13.

Rossi, Sandra. "What users hate the most about Web sites," Computerworld Today (Australia). 14 June 2006.

Shapely, Deborah. "The da Vinci of Data," New York Times 30 May 1998.

Shneiderman, Ben. Panel discussion. IEEE Symposium on Information Visualization, 2007. Sacramento, CA. 28 Oct. 2007.

Silobreaker.com. Silobreaker, Ltd. Accessed 21 April 2008. <<http://www.silobreaker.com>>

Silva, P. and A. Dix. "Usability – Not as we know it!" Proceedings of BCS HCI 2007, People and Computers XXI, BCS eWiC Sep. 2007
<<http://www.hcibook.com/alan/papers/HCI2007-YouTube/>>

Smartvote.ch. Politools.net. Accessed 21 April 2008. <<http://www.smartvote.ch>>

Spiderman. Dir. Sam Raimi. Columbia Pictures, 2002.

Stasko, John, et al. "InfoCanvas: Information Art," 2003. Georgia Institute of Technology. Accessed 20 April 2008.
<<http://www.cc.gt.atl.ga.us/gvu/ii/infoart/approach.html>>

TIBCO Spotfire. 13 April 2008. TIBCO. Accessed 13 April 2008.
<<http://spotfire.tibco.com/>>

Tractinsky, N., A. Shoval-Katz, and D. Ikar. "What is Beautiful is Usable," Interacting With Computers. 12(2), 2000: 127-145.

Tufte, Edward R. The Visual Display of Quantitative Information. Cheshire, CT: Graphics Press, 1983.

Ware, Colin. Information Visualization: Perception for Design. San Diego, CA: Academic Press, 2000.

Wattenberg, Martin. "Introduction to Visualization: From data to pictures to insight," IBM.com. 16 May 2005. IBM Visual Communication Lab. Accessed 10 Dec. 2007.
<<http://www.alphaworks.ibm.com/contentnr/introvisualization>>

van Wijk, J.J. "The Value of Visualization," Proc. IEEE Conf. Visualization 2005. 2005: 79-86.

Vande Moere, A. "Form follows Data: the Symbiosis between Design and Information Visualization," International Conference on Computer-Aided Architectural Design (CAADfutures'05). 2005: 31-40.

Viégas, Fernanda and Martin Wattenberg. "Communication-Minded Visualization: A Call to Action," IBM Systems Journal. 45(4), 2006.

visualcomplexity.com. Accessed 21 April 2008. <<http://www.visualcomplexity.com>>

Wexelblat, A. and P. Maes. "Footprints: History-Rich Tools for Information Foraging," Proceedings of ACM Conference on Human Factors in Computing Systems. May 1999: 270–277.

Woolman, Matt. Digital Information Graphics. New York, NY: Watson-Guption Publications, 2002.

Zuckerman, Ethan. "Global Attention Profiles," 8 Aug. 2003. Berkman Center for Internet and Society at Harvard Law School. Accessed 10 April 2008.
<<http://gapdev.law.harvard.edu/>>