The Social Life of Visualization

Jeremy Yuille
RMIT University | jeremy.yuille@rmit.edu.au

Hugh Macdonald
RMIT University | hugh.macdonald@rmit.edu.au

Data.gov is an open government initiative of President Barack Obama's administration, designed to increase the public's ability to find, download, and use high-value, machine-readable data sets generated by the executive branch of the federal government. The site sees public participation and collaboration as one of the keys to the success of Data.gov; it will enable the public to participate in government by providing downloadable data sets to build applications, conduct analyses, and perform research [1].

Data.gov is part of a wider movement called "open data," which various (mainly government) organizations around the world are exploring. In Great Britain, the Power of Information Taskforce has outlined its vision for public sector reform. Of particular note is its concept for open information, whereby "to have an effective voice, people need to be able to understand what is going on in their public services." Also of interest is its vision for open discussion that seeks to promote greater engagement with the public through more interactive online consultation and collaboration [2]. A similar move is under way in Australia with the Government 2.0 taskforce, which is concerned with encouraging online engagement with the aim of "drawing in the information, knowledge, perspectives and even, where possible, the active collaboration of anyone wishing to contribute to public life" [3] .

A signification consideration for these governments might be how to build public participation and collaboration in the process of collecting and sharing data. In order for public data to benefit from the innovation and dynamism of Web 2.0, government needs to change its modus operandi as an information provider [4]. In doing this, it should focus on design sites that create simple, reliable, and publicly accessible infrastructure that expose the underlying data. "Private actors, either nonprofit or commercial, are better suited to deliver government information to citizens and can constantly create and reshape the tools individuals use to find and leverage public data" [4].

As we move into this new age of open public data, what sort of tools should an individual have at his or her disposal to find and leverage this public data? How can these tools be designed so that people can better understand what is going on in their public services and engage with them more fully?

The framework presented here could enable individuals to explore open data, understand what is going on, and engage with it. The framework relies on the use of visualization as an interface to explore data and support social collaboration around it. It allows people to take some data, explore its properties, and present their findings to others. This is the process of shared storytelling that the framework supports: It can be seen as a democratic way of working with the ideals of open data and government 2.0 that allows people to better understand government processes and engage with them more fully.

At the basis of the data-visualization framework is the distinction between an object-centered social network and an ego-centered one. Prominent object-centered social networks are Flickr and del.icio.us, where the network revolves around an activity—sharing photos or sharing links to websites. The best example of an ego-centered social network is Twitter, where people within the network share their thoughts with each other.

The idea of object-centered social networks comes from the work of sociology professor Karin Knorr Cetina's theory of object-centered sociality—the individual and the object as central elements in social interaction. Cetina proposes that objects, around which discussions take place, help focus and

[1] Data.gov; http://www.data.gov/about/

[2] Power of Information Taskforce; http://powerofinformation.wordpress.com/

[3] Government 2.0 Taskforce; http://gov2.net.au/about/

[4] Felten, E., Robinson, D., Yu, H., and Zeller, W. "Government Data and the Invisible Hand." *Yale Journal of Law and Technology* 11 (2008): 161.

start conversations and other social interaction among people. In this case, visualizations are the objects within the network, and because it prioritizes them over the relationships between people in the network, it focuses attention on the process of shared storytelling.
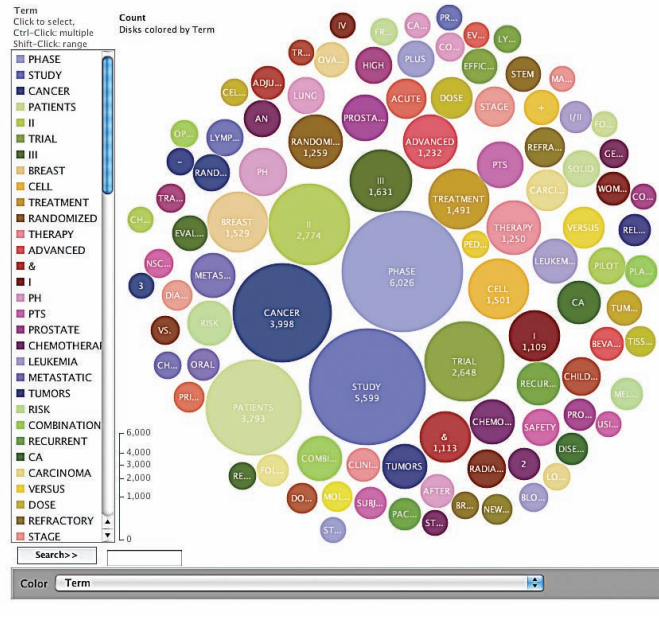
For the process of shared storytelling to occur around an object, such as a visualization, the object needs an identity within the social network so that people within can easily recognize it and its unique properties can be retained. Identity in this context is created in much the same way it is created in other scenarios: through a combination of visual and textual information about the object. Providing these contextual clues about the object ensures that people within the social network can locate it without difficulty and establish whether or not they can make contributions to it. The specifics of giving a visualization an identity within the network is described in the framework as the process of *decoration*, and through its use the object prepares for the process of shared storytelling to occur around it.

Decoration is essential because the process of shared storytelling through visualization can take place only when there is a shared understanding of the medium's properties. Visual communication does not have the level of shared understanding that written communication does. So in addition to presenting the visualization to begin the process of shared storytelling, a well-established visual language has to exist so that the properties of the data can be
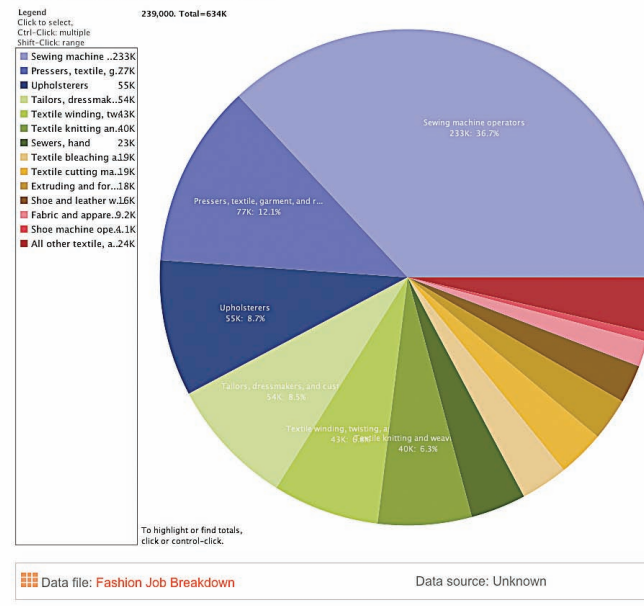
▶ IBM's Many Eyes project uses visualization for data analysis. Data sets are transformed into an on screen image, and the user can then perceive an unexpected pattern.

For the process of shared storytelling to occur around an object, such as a visualization, the object needs an identity within the social network so that people within can easily recognize it and its unique properties can be retained.

communicated. People within a social network may not have the necessary knowledge to grasp what sort of visualization technique to use, but they may still have interesting data to share. This is particularly important for open data that seeks to promote greater transparency for government and increased engagement with democratic

processes through technology, so the entry barrier to using this technology must be low enough.

The shared storytelling process begins with the creation of the visualization itself. People who wish to share their data need help with the process of choosing a visualization technique in case they don't have the necessary knowledge to begin that conversation. This particular process is known as *mapping*, which can help people to visualize a data set so that a shared conversation around it can begin.

The conversation around visualization takes place because it has become a social object within the network. A social object is anything around which discussion takes place. A movie is a social object because it has a plot, cast, crew, and a mise en scène, all of which can be discussed by fans, critics, and other interested parties. On the other hand, a visualization is a social object because it is a representation of a data set.

Consequently, the data set is the true object of discussion, and there must be a set of tools to allow people to explore it

using the visualization as an interface. These tools should transform the data set in some way, whether by flipping the axes to gain a different perspective on the data, or by mapping some new axes onto the visualization to look for correlations in the underlying data. This is the process of *tweakability* that reasons that users must shift and reformat a visualization in order to make sense of the whole data set.

Users within an online social network also need a way of documenting their discoveries, particularly as a way of contributing to the democratic process surrounding a piece of government data. A simple form of annotation is crucial in ensuring that people can employ the visualization interface to highlight the insights they have drawn from the data set and continue the storytelling process. In the framework, it is a non-disruptive form of annotation that will leave the original visualization intact and make the annotated version a derivative. This type of process maintains the analogy to storytelling, in which an original story might be fleshed

CREATE   **INTERPRET**   **CAPTURE**

MAPPING   DECORATION   TWEAKING   ANNOTATION   SNAPSHOT
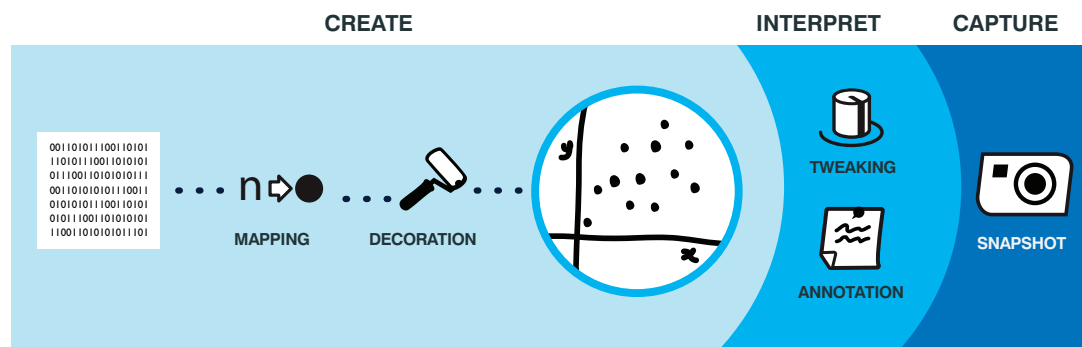
▶ The Social Life of Visualization. The move from raw data to data storytelling, involving processes of: *mapping*—where communication objectives are translated into visualization schemas; *decoration*—creating an identity around the visualization and placing it in the social space; *tweaking* and *annotation*—interfaces to interrogate and mark up the representation of data; and *snapshot*—enabling storytelling to grow around a data visualization.

out on subsequent tellings, the source of the derivation remains clear. This process also supports the democratic notion of backing up a claim with evidence. In exactly the same way, the visualization interface must emphasize the original visualization and its intention, yet support derivations. This process is called *annotation*.

A means of capturing these story annotations will retain the benefit of the community-driven storytelling process around the visualization. Just as the original visualization is a capture of the underlying data set, subsequent versions of the visualization also must be captured and attached to preserve the original message but at the same time allow the breadth that comes of continued discussion. So another part of the model is building processes into the interface that allow annotations to be preserved, commented on, and subsequently reviewed by other members of the community. It is this process within the framework that provides shared storytelling through visualization with the ability to create knowledge artifacts around data. The process is called *snapshot*, and with this, extra data created around the visualization is preserved.

Significant amounts of data can be generated because the entire process of shared storytelling through visualization is a life cycle that repeats infinitely. The framework gives data visualization a life force by providing it with an identity that enables it to exist within an object-centered social network and making it an interface so that people can interact with it

to begin a shared storytelling experience. As data becomes increasingly prevalent on the World Wide Web, the ability to engage with a communal visualization of a data set is a more useful experience than new methods of visualizing data. For designers, this is a change in approach toward visualization; no longer is it about making the most visually appealing and sophisticated representations. Instead, this creativity should be constrained to reverting the control of data to people and providing a good experience along the way. With these technologies, it's impossible to imagine how many important stories will be told by groups of like-minded people in the future.

With the increasing adoption of open data by governments around the world, and in the case of Data.gov especially, the amount of new data sets being added to the site, many stories can come out of this marriage of government 2.0 and the shared storytelling process. The framework presented should offer the flexibility and intuitiveness to enable people to leverage public data for their own purposes. Through this, people should be able to collaborate more closely with their governments and better understand government decisions: how they are made and how they might affect individual citizens.

Postscript

In this article we've discussed how the framework for an interface encourages social interactions around data visualization, which we've proposed, can be used with open data, to encourage transparency and enhance community engagement

with government. If you're interested in finding out about the framework in more detail we've published it as a series of interaction design patterns at http://socialvizpatterns.info.

ABOUT THE AUTHORS
Jeremy Yuille is an interaction designer, digital media artist, and academic with a background in digital art, music, performance, and architecture. He has a bachelor's of design studies from the architecture department of the University of Queensland and a master's of design from SIAL at RMIT University. Yuille is a cofounder of the Media and Communication Design Studio at RMIT, where he undertakes collaborative research with the Australasian CRC for Interaction Design (ACID), supervises postgraduate students, and holds interaction design studios. He is also a certified ScrumMaster and a director of the Interaction Design Association. He infrequently blogs on design and the progress of his Ph.D. at http://isomorpho.us.

Hugh Macdonald is a research assistant on the ACID Loupe Project and has been researching information visualization and theories of social interaction design. He is also currently a Ph.D. student at the School of Media and Communication at RMIT University in Melbourne, Australia. His research involves how the emergence of networked media, along with its technologies and social practices, is changing the way professional sporting organizations and sports fans interact with each other within the media landscape. Macdonald also maintains a keen interest in mobile technologies, having previously worked in this area and received a related master's degree. He is currently looking at some of the social effects of these technologies.