Scientific Storytelling Using Visualization

Kwan-Liu Ma and Isaac Liao
University of California, Davis

Jennifer Frazier
The Exploratorium

Helwig Hauser
University of Bergen

Helen-Nicole Kostis
NASA Goddard Space Flight Center

> I like a good story well told. That is the reason I am sometimes forced to tell them myself.
> —*Mark Twain*

What springs to mind when you hear the word "storytelling"? For most of us, it conjures up images of children gathered in front of a rocking chair, rapt with attention as an elder narrates a fairy tale. Unencumbered by the inhibitions of older years, they aren't afraid to interrupt and ask for details to satisfy their curiosity, or clamor for more when the story ends. How can we, as visualization researchers and practitioners, elicit this same engagement and wonder in our viewers? How can we aid Mr. Twain in his plight and ensure that he isn't the exclusive purveyor of good stories well told?

All stories are sequences of causally related events. However, the good ones tend to share several important features. First, they take time to unfold, and their pacing matches the audience's ability to follow them. Second, they hold the audience's attention by having interesting settings, plots, and characters. Finally, they leave a lasting impression, either by piquing the audience's curiosity and making them want to learn more or by conveying a deeper meaning than your everyday run-of-the-mill sequence of causally related events.

In using visualization to tell a story, what does "good pacing" mean? A well-paced story exhibits deliberate control over the rate at which plot points occur. However, any given pace might feel too fast or too slow to different audiences, depending on their attention spans and personal preferences. Similarly, in designing visualizations, it's crucial to gauge the intended viewers' familiarity with both the subject matter and visualization conventions. For a given dataset, distributing data across multiple line charts might prove the most suitable approach for a general audience. However, domain experts might prefer to combine data into a single parallel-coordinate diagram to facilitate comparison.

What are the settings, characters, and plots of visualizations? The setting is all the background information a viewer needs to know to contextualize and comprehend the visualization. In theatrical productions, the stage is generally set before the curtain rises; similarly, viewers should be introduced to the subject matter before seeing a visualization of it. In addition, visual elements representing data points are the characters and centers of attention in visualizations—they're the stars of the show. Finally, a visualization's plot, and dramatic tension, arises from the juxtaposition of its visual elements, how they interact and compare with one other, and how they evolve over time.

Armed with these notions, let's discuss how to use visualization to tell a good story, and tell it well. In particular, we emphasize *scientific storytelling*—telling stories using scientific data—which is a topic that the visualization research community has paid little attention to so far. In contrast, storytelling in information visualization has been

Published by the IEEE Computer Society

the topic of several recent workshops and panels, and provides a starting point for the discussion of scientific storytelling.

Storytelling in Information Visualization

VisWeek 2010 in Salt Lake City featured the day-long workshop, "Telling Stories with Data: Using Visualization to Create Narratives and Engage Audiences" (http://thevcl.com/storytelling). Hosted by Matt McKeon (IBM Research), Joan DiMicco (IBM Research), and Karrie Karahalios (Univ. of Illinois), the workshop featured a range of speakers including journalists, bloggers, literary analysts, and developers of information visualization software. Throughout the day, we saw numerous examples of how stories are told with data, including

- maps showing US and coalition casualties in Iraq and Afghanistan ("Home and Away"; www.cnn.com/SPECIALS/war.casualties/index.html);
- a political blogger (Matthias Shapiro, aka "10,000 Pennies") using pennies to explain that a budget cut of $100 million, although sounding impressive, is actually a tiny fraction of the US national deficit; and
- CommentSpace (www.commentspace.net), a collaborative-visualization website that lets users create, share, and comment on dataset views.

Visualization creators shared their goals and design decisions, and breakout sessions allowed for small-group discussions.

Several interesting points emerged over the course of the day. The general consensus was that framing data as a narrative makes it more interesting and memorable. Why might this be? Cognitive science postulates the existence of two types of memory: semantic memory, for remembering disconnected facts, and episodic memory, for remembering sequences of events. By presenting themselves as narratives, visualizations can tap into episodic memory and establish themselves as cohesive entities.

In addition, the issue of interactivity in visualizations came up repeatedly. The style of storytelling in static visualizations, such as infographics, differs fundamentally from that in interactive visualizations, which let users navigate and modify views of data. Making a visualization more interactive gives users more freedom to explore but lessens visualization designers' control over how the story is told. In the end, the participants concluded that a visualization's interactivity should be carefully balanced against the need to guide the viewer through the data. A useful compromise might be to start the visualization in a noninteractive mode, ensuring that it presents the dataset's most salient features, and then let users explore the rest of the dataset.

The visualizations presented and discussed in this workshop fell squarely in the domain of information visualization, which tends to use more abstract representations that are usually targeted toward more general audiences. In contrast, what challenges does scientific visualization face in storytelling?

How can we aid Mr. Twain in his plight and ensure that he isn't the exclusive purveyor of good stories well told?

Scientific Storytelling

Visualization has become an important tool for scientists in their daily work. Scientists create visualizations for various purposes: to validate experiments, explore datasets, or communicate findings to others. If appropriately presented, such visualizations can be highly effective in conveying narratives. So, using the criteria we mentioned earlier, let's explore the possibility of telling stories using scientific visualizations.

Information visualization's narrative impact stems from visual comparisons using simple, abstract representations of data: bar charts show differences in length, scatterplots show differences in position, treemaps and pie charts show differences in area, and heat maps show differences in color and intensity. As such, information visualization stories are about comparison or change: "Look at how much bigger A is than B," or "Look at how C has grown over time."

In contrast, much of scientific visualization's narrative impact comes from being able to see real data that are normally invisible. At its best, scientific visualization extends our senses, letting us perceive and manipulate data at otherwise impossible scales and perspectives, such as vector fields in weather systems, isosurfaces in supernova simulations, and layers of human anatomy rendered semitransparently. Whereas information visualizations are allegories—abstractions and summaries of raw data—scientific visualizations are more literal; they strive for realism and spatial accuracy, sacrificing details only to facilitate understanding.

In some ways, scientific visualization has it easy. Usually, the intended viewers are the scientists who generated the data, and others in the same field. So, they need little introduction—in terms of our storytelling metaphor, they're already familiar with the setting, and all that's left is identifying the characters (for instance, what glyphs represent and how color is used). In fact, when we design scientific visualizations, the scientists are usually the ones setting the stage for us! Additionally, the fact that the data are already highly relevant to them increases the likelihood that visualizations will leave a lasting impression in their minds.

Whether the form is literary, performance based, aural, visual, or interactive, a storyteller should know the story's audience and take ownership of the story.

However, difficulties arise when introducing scientific visualizations to broader audiences. Even the best visualizations are incomprehensible if their concepts are alien, and scientific visualizations often assume viewer familiarity with the subject matter. Moreover, time constraints and limited attention spans often preclude the possibility of full explanations. How can we address these issues?

In 2010, a one-day workshop on scientific storytelling took place at the University of California, Davis. Participants included visualization researchers and practitioners as well as experts in animation, scientific journalism, and science museum exhibition. The rest of this article presents highlights and findings from this workshop.

Production Visualization at a Scientific Research Center

Using visualizations to tell scientific stories is a routine practice at NASA. Observational data—data that can be recorded by instruments and sensors—are continuously collected, archived, and processed from NASA airborne missions and experiments. As of 2011, 64 airborne missions are operating within NASA's Science Mission Directorate (19 in the Earth Science Division, 16 in Heliophysics, 15 in Astrophysics, and 14 in Planetary).[1] Each mission usually involves multiple sensors and instruments that aim to acquire and transmit datasets daily, hourly, or even every few minutes. Data acquisition is ongoing and lasts for the mission's duration. Most airborne missions are operational for more than a year, and some can be operational for more than a decade (for example, Landsat satellites).

NASA scientists, who are sometimes the missions' principal investigators, need to process and visualize data acquired from airborne science missions to advance their research and support outbound communication and scholarly work, such as publishing in scientific journals. NASA also needs data visualization to engage and educate the public about its research and science efforts. Scientists and mission teams have their own tools to process and analyze data but can't easily develop and produce high-quality visualizations for three reasons:

- the data's complexity and volume,
- the complexity of the tools and technology necessary to produce high-quality visualizations, and
- the lack of expertise in visualization techniques and storytelling production.

The Scientific Visualization Studio (SVS; http://svs.gsfc.nasa.gov) at NASA's Goddard Space Flight Center (GSFC) facilitates scientific inquiry and outreach in NASA programs through visualization. The SVS works closely with scientists to create visualization products, systems, and processes to promote greater understanding of Earth and space science research at GSFC and in the NASA research community. The SVS also provides expertise in data visualization and science storytelling and is part of the larger Earth Science Storytelling team, which comprises three entities: the SVS, the Conceptual Image Laboratory (concept animators producing non-data-driven products), and Goddard TV Multimedia (a team of producers, science writers, video editors, camera crews, and Web and social-media experts).

Data visualizations produced and developed at the SVS are cinematic-quality computer graphics short films, similar to productions by Hollywood computer animation studios. The visualizations' main characteristics are scientific integrity, data preservation, seamless blending of multiresolution data from different sources, aesthetics, and a solid story that engages the public. The successful production of such visualizations depends on free-form collaboration among members from all three teams and requires

- communication between all the parties involved, including scientists;
- data availability and transparency regarding dataset limitations or problems;
- a context that makes the science story relevant and interesting to the public;

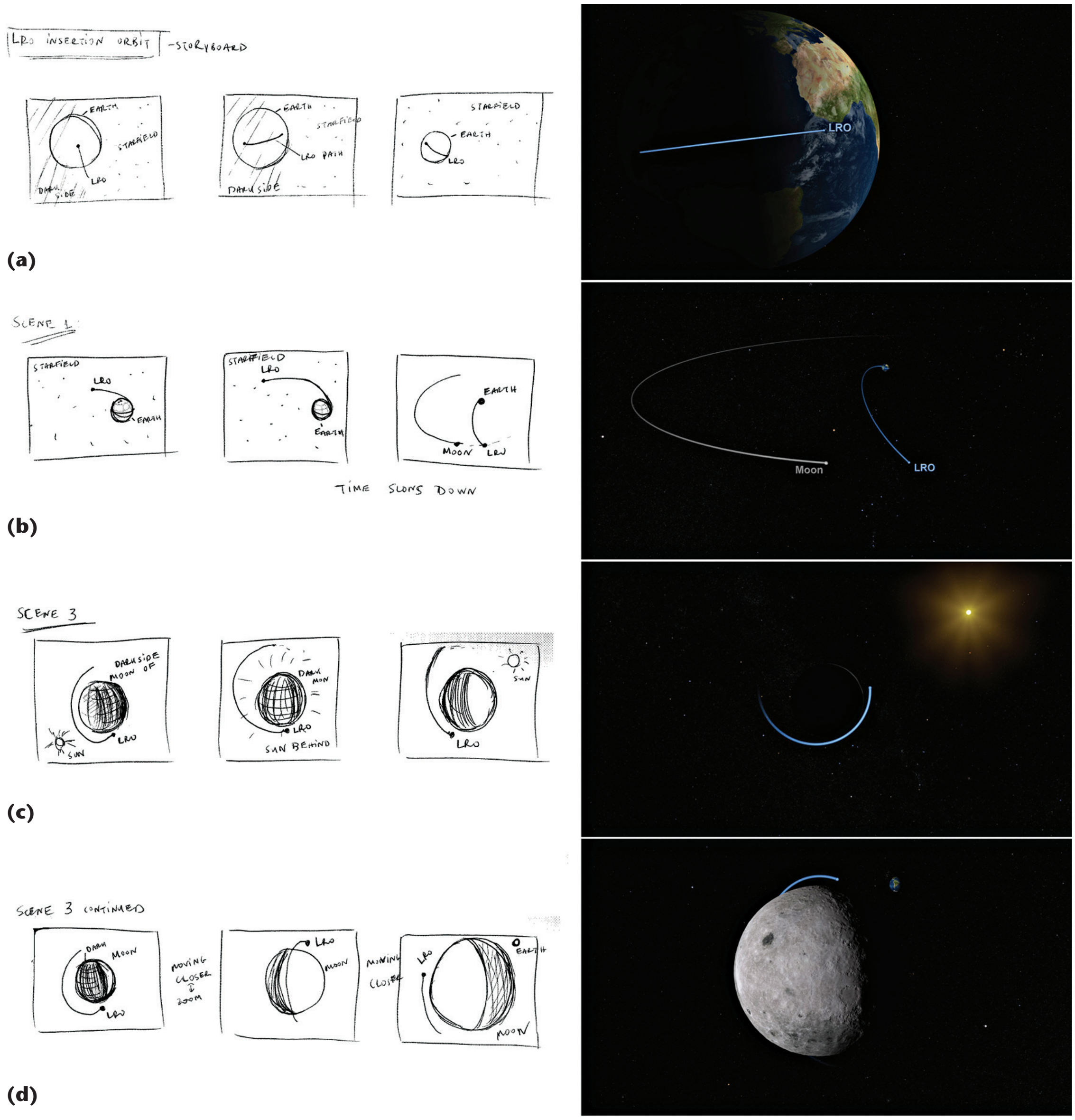

Figure 1. From storyboard to visualization: the story of NASA's Lunar Reconnaissance Orbiter (LRO) mission, told as a sequence of images. (a) The LRO launched from Cape Canaveral. (b) The LRO approaching the moon. (c) The LRO orbit trail shown with the sun and the dark side of the moon. (d) The LRO moving into orbit around the moon.

- resources for producing visualization stories; and
- the science visualizers' abilities to shift roles, wear multiple hats, and collaborate.

Storytelling is a key component of every SVS visualization. Although storytelling manifests itself differently in various art forms, whether literary, performance based, aural, visual, or interactive, a storyteller should know the story's audience and take ownership of the story. In general, all forms of stories have four ingredients: perspective, characters, imagery, and language. These ingredients are combined in a structure that defines the story from beginning to end. Visual storytelling, and specifically, storytelling for animation, borrows from the conventions of photography, cinema, episodic comics, and the performing arts. The structure in storytelling for animation is established by camera work (visual perspective, time and space of framing, composition, point of view, lighting, color, form, and style); audio work (use or nonuse of sound, and timbre); and the animation's visual, aural, and editorial rhythm.

Figure 1 captures stages of the SVS production of a visualization of NASA's Lunar Reconnaissance Orbiter mission (http://svs.gsfc.nasa.gov/goto?3603). Visualization-driven end products are archived in the SVS repository (http://svs.gsfc.nasa.gov), which is a free, publicly accessible database

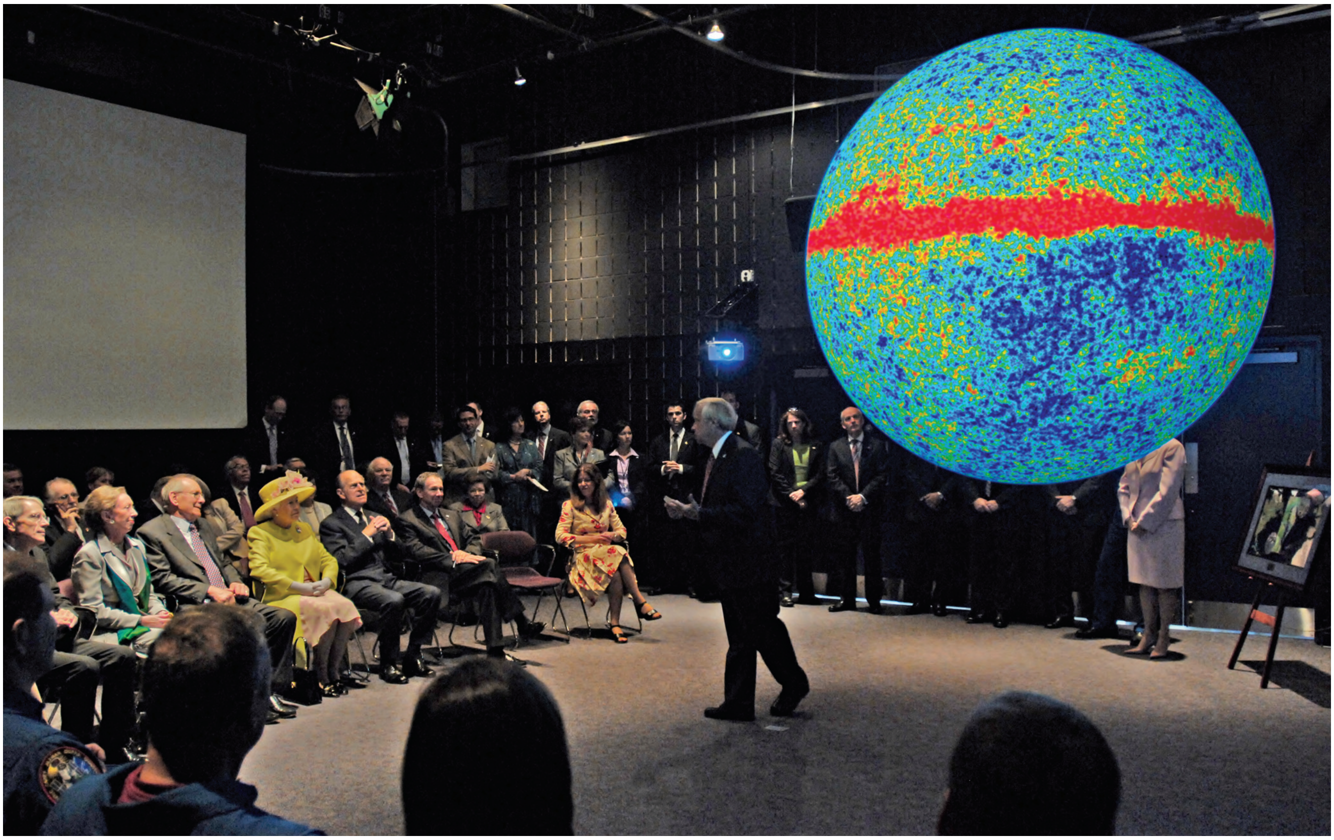

Figure 2. Science on a Sphere, a presentation to Queen Elizabeth II and Prince Philip at NASA's Goddard Space Flight Center. This system displays data onto a 6-foot-diameter sphere.

with more than 3,800 entries (as of Sept. 2011). The products span many visualization forms, including 2D, 3D stereoscopic, Science on a Sphere (see Figure 2), hyperwalls (http://svs.gsfc.nasa.gov/vis/a000000/a003700/a003793/nccs_1024x576.jpg), dome shows, and even touch displays (for example, NASA's Viz iPad application; http://svs.gsfc.nasa.gov/nasaviz). Each production includes various formats including frame sets, still images, movies, and, when appropriate, data in a wide gamut of resolutions. Upon release, the products might take on lives of their own because the public can use them freely.

Although a streamlined process is in place for producing visualizations, there are always challenges that might compromise the end product's quality, structure, and story. These challenges are often rooted in data issues—for example, data gaps, insufficient data, low resolution, or even data that don't show the expected phenomena. At other times, new visualization techniques are required in order to highlight important, necessary information. The need for new techniques can occur within either the technical infrastructure of the visualization production pipeline (for example, modified shaders, transitions between different coordinate systems, or the development of a new pipeline) or the design domain (for example, finding the best ways to map complex data to visual models). In short, the resources and effort required to produce high-quality visualizations for scientific storytelling can overwhelm any individual scientist.

The SVS is one example of a successful scientific storytelling and visualization studio. Creating visualizations suitable for consumption by the general public poses unique challenges and requires a dedicated team of versatile, talented individuals. In short, scientific storytelling isn't a trivial endeavor, and creating successful visualizations requires the collective effort of many specialists working together.

Production Visualization at a Science Museum

At science museums, people can experience science in ways they can't at school or home. Visitors can swing on a giant pendulum, stand under a life-size *T. rex* fossil skeleton, or watch the birth of a galaxy in 3D. Museums tell the stories of science, and—perhaps more important—provide a unique venue for people across generations to play together, interact with scientists, and use scientific tools.

Museums have long used visualizations to show the public things they can't normally see, such as evolutionary relationships or DNA structure. But visualizations are an increasingly critical medium for science museums. As the volume of data collected by scientists expands exponentially, visualization is the tool that lets them make observations or detect patterns. Whether comparing genomes, mapping a virus's structure, or developing new models of Earth's climate, most scientists now do some—if not all—of their work using visualized data. To tell the stories of modern science, museums must use visualizations.

Visualizations' growing importance in science presents an exciting opportunity for museums. Scientific visualizations can provide stunning images, engaging the public with phenomena they've never seen before. Visualizations can be displayed on large, dynamic interfaces, providing new ways for the public to participate in interactive, social learning. Museums can also use visualizations to create authentic tools for the public to make their own discoveries, analogous to microscopes or telescopes.

But for scientific visualizations to have any significant meaning for the public, they must be carefully interpreted and designed. Visualizations often show complex, abstract phenomena at extreme size scales using colors that have no inherent meaning. For instance, a research study at the Exploratorium (www.exploratorium.edu), a science museum in San Francisco, showed that many visitors grossly misinterpreted the scale and use of color in a nanoscale image.[2] Similar studies have documented learners' difficulty in interpreting visualizations from fields as disparate as genetics and astrophysics.

As science museums increase their use of scientific visualizations, they're providing more interpretation through labels, videos, and live explanations. A complementary, although less common, strategy is to redesign scientific visualizations with the public in mind. Molly Phipps and Shawn Rowe conducted a study showing that students better interpreted visualizations of oceanographic data that had been redesigned with more intuitive color schemes and recognizable (although unscientific) landmarks.[3]

The most significant challenge for museums is finding ways to transcend the use of visualizations as explanatory animations or pretty pictures. In many museums, visitors can watch stunning simulations of Earth's climate or the collapse of a star, but they can't control or explore them. Such direct interaction would let visitors control their experience and make discoveries with data the way scientists do.

To address this challenge, the Exploratorium is creating visualization tools with which the public can ask and answer their own questions with real scientific data. In a pilot project funded by the US National Science Foundation (NSF), the

Figure 3. The Bay Model at the Exploratorium (www.exploratorium.edu/outdoor/#/exhibit/bay-model) lets visitors interact with a scientifically accurate model of how tides, currents, and rivers combine to create the complex water flows of the San Francisco Bay estuary.

Exploratorium is collaborating with visualization researchers. By tailoring the development process to different end users (for example, the public) and iterating through intensive prototype testing with visitors, the Exploratorium hopes to pioneer this new genre of museum exhibits.

One precursor of this project is the Bay Model (www.exploratorium.edu/outdoor/#/exhibit/bay-model). It lets Exploratorium visitors interact with a scientifically accurate model of how tides, currents, and rivers combine to create the complex water flows of the San Francisco Bay estuary (see Figure 3). Using a touch screen, visitors place virtual floats into a video image projected onto a 3D topographic model of the Bay Area. After launching a float, the visitors watch how currents move it to different locations according to predicted tide and river flow cycles. Color-coding highlights varied water conditions during tidal phases.

In summary, visualization's increasing role in scientific discovery presents a tremendous opportunity for science museums to engage the public with stunning images, novel interfaces, and authentic tools. However, transforming the rapidly growing number of scientific visualizations into meaningful experiences for the public requires thoughtful interpretation, design, and collaboration.

Storytelling Using Interactive Visualization

Following the publication of the NSF's report *Visualization in Scientific Computing* in 1987, the early development of the field of scientific visualization was driven largely by the need to gain insight into large, complex scientific and medical datasets. This led to many new visual abstractions, rendering methods, and interaction techniques. However, the visualizations used in scientific storytelling are generally created after the fact, separately and independently from data exploration. This is because the visualization process has no built-in storytelling model; that is, stories based on visualization, data exploration, and knowledge discovery must be manually constructed by scientists.

The concept of incrementally creating a story by depicting the visualization process's progress is intuitive and powerful. The scientist is immersed in the data domain and assembles pieces of the story as he or she learns more and more about the data. A visualization system called AniViz realizes this concept by letting users incrementally build a story and present the story as an animation.[4] As users interactively explore data, they can locate interesting views, specify views as animation keyframes, review the animation constructed so far, add annotations and voice-overs, and edit keyframes and transitions until the exploration is complete and the resulting animation is satisfactory.

Although storytelling by nature isn't completely interactive, we ponder how we can facilitate interactive storytelling.

You can present the keyframe approach as a story model if that's more intuitive to users. Michael Wohlfart and Helwig Hauser described just such a story model, consisting of two types of components.[5] *Story nodes* are major steps or milestones in which a story briefly halts, perhaps for interactive exploration by the story consumer, and then resumes. *Story transitions* smoothly connect story nodes, leading from one node to the next.

Using Wohlfart and Hauser's model, you can create several types of visualization stories. For instance, you could tell a story that conforms to Ben Shneiderman's Visual Information-Seeking Mantra ("overview first, zoom and filter, then details on demand").[6] Such a story would begin with an overview of the data. It would then follow with a focusing transition, leading the user to a more detailed visualization of some particular aspect. It would conclude with a guided sequence of images that substantiate the message to be communicated. Alternatively, you could construct stories aimed at comparative visualization (for example, building a side-by-side comparison during the story). Or, you could base the story on iterative visualization (such as the sequential visualization of all relevant features in a selected region, following a repetitive pattern such as "zoom onto a particular feature, rotate around it, show the context, and then continue to the next feature").

Although storytelling by nature isn't completely interactive, we ponder how we can facilitate interactive storytelling. How can we stimulate the story consumers' participation? Can we let them influence not only how the story is told, but also how the story ends? For example, adventure games let users interact with and affect a premade game story. Also, science museums offer many hands-on activities, which might be considered a form of interactive storytelling. However, once spectators become "spect-actors" (the terminology of Augusto Boal, in *Theater of the Oppressed*[7]), a conflict of control emerges: the spect-actor diverts the course of the story from the original plan. This is also called the *narrative paradox*, and people have suggested different ways to address it (for example, by using emergent narratives, as Sandy Louchart and Ruth Aylett described[8]).

Wohlfart and Hauser proposed a taxonomy of four modes for splitting control between the author and consumer by varying degrees.[5] Traditional *passive storytelling* prohibits interaction on the consumer's part; the author fully controls all domains. In *storytelling with interactive approval*, passive storytelling pauses at certain points and lets spect-actors take temporary control. They can change the visualization's view, representation, and even content. Once they're satisfied with this interactive exploration, storytelling continues as originally intended. In *semi-interactive storytelling*, consumers can take control not just for an interim excursion but for an entire section of the story. Finally, in *total separation from the story*, consumers can completely detach from the story and engage in interactive visualization with total freedom.

In terms of storytelling, interactive visualization could help with three issues that are important to communication: comprehensibility, credibility, and involvement. First, incrementally building a story, enhancing it with labels and annotations, and letting viewers interrupt and control it all reduce the risk of presenting an overloaded visualization that's poorly understood, and thus improve comprehensibility. Second, you can improve a visualization's credibility by letting viewers interact with it and verify that it shows what it claims. Finally, letting viewers interact with visualizations "breaks the fourth wall," transforming them from passive observers to active participants. Consequently, they'll feel a greater sense of engagement with the data being presented.

Clearly, the need exists to consider how storytelling and visualization can make scientific findings more comprehensible and accessible to the general public. Scientific visualization has much to learn from information visualization in this regard. Consider that information visualizations are aimed at the general public and that they draw attention to differences and changes in visual elements. Perhaps scientific visualizations can take a similar approach to reach broader audiences. If we focus on important features by emphasizing how they change across time or experimental conditions, we might be able to tell a compelling story without having to explain extraneous details.

In addition, thinking about visualizations in a narrative context can help make them more comprehensible, memorable, and credible to the general public. Whether we use visualizations to tell a story or use a story model to make visualizations more compelling, we can't neglect the fundamentals of good storytelling. First, know your audience—assess their level of domain knowledge and familiarity with visualization conventions. Next, set the stage—make sure they have enough background on the dataset being visualized to make sense of your visualization. Introduce the characters—show them the visual elements and what they represent. Develop the plot—arrange your visual elements in a way that tells an interesting and compelling story. Finally, leave the audience with a lasting impression by showing them how the story is relevant to them, and its greater implications.

Scientific storytelling using visualization isn't easy; the successful examples highlighted in this article are the exception rather than the rule. Much work remains in establishing guidelines and principles for successful storytelling. As visualization designers, we must ask ourselves how we can better support the scientific community's efforts in reaching out to the general public. Scientists have amazing stories to tell, and we can help ensure that they aren't—to paraphrase Mark Twain—forced to tell them themselves.

Acknowledgments

US National Science Foundation grants CCF-0850566 and DRL-1011084, and US Department of Energy agreement DE-FC02-06ERZ5777 partly supported this research.

References

1. "Missions," NASA, 2011; http://science.nasa.gov/missions.
2. J. Ma, *Visitors' Interpretations of Images of the Nanoscale*, Nanoscale Informal Science Education Network, July 2008; www.nisenet.org/sites/default/files/catalog/eval/uploads/2009/04/1353/visitor_interpretations.pdf.
3. M. Phipps and S. Rowe, "Seeing Satellite Data," *Public Understanding of Science*, vol. 19, no. 3, May 2010, pp. 311–321.
4. H. Akiba, C. Wang, and K.-L. Ma, "AniViz: A Template-Based Animation Tool for Volume Visualization," *IEEE Computer Graphics and Applications*, vol. 30, no. 5, 2010, pp. 61–71.
5. M. Wohlfart and H. Hauser, "Story Telling for Presentation in Volume Visualization," *Proc. 2007 Eurographics/IEEE-VGTC Symp. Visualization* (EuroVis 07), Eurographics Assoc., 2007, pp. 91–98.
6. B. Shneiderman, "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations," *Proc. 1996 IEEE Symp. Visual Languages*, IEEE Press, 1996, pp. 336–343.
7. A. Boal, *Theatre of the Oppressed*, Theatre Communications Group, 1979.
8. S. Louchart and R. Aylett, "Narrative Theory and Emergent Interactive Narrative," *Int'l J. Continuing Eng. Education and Lifelong Learning*, vol. 14, no. 6, 2004, pp. 506–518.

Kwan-Liu Ma *is a professor of computer science at the University of California, Davis, and is on the* IEEE Computer Graphics and Applications *editorial board. Contact him at ma@cs.ucdavis.edu.*

Isaac Liao *is a postdoctoral researcher in the Department of Computer Science at the University of California, Davis. Contact him at ihliao@ucdavis.edu.*

Jennifer Frazier *is a cell biologist at the Exploratorium. Contact her at jfrazier@exploratorium.edu.*

Helwig Hauser *is a professor of visualization at the University of Bergen. Contact him at helwig.hauser@uib.no.*

Helen-Nicole Kostis *is a science visualizer and the project manager of the NASA Viz iPad application at NASA's Goddard Space Flight Center. Contact her at helen-nicole.kostis@nasa.gov.*

Contact department editor Theresa-Marie Rhyne at theresamarierhyne@gmail.com.