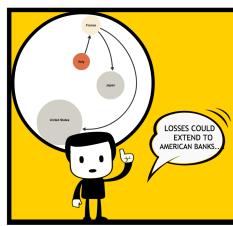
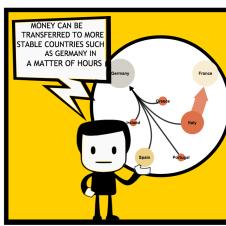
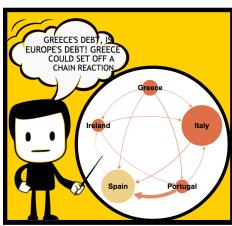
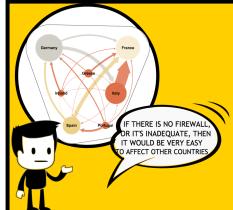
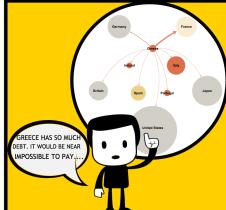
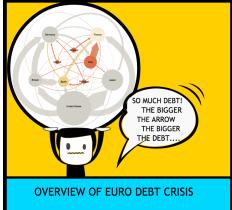


# Data Comics: Sequential Art for Data-Driven Storytelling

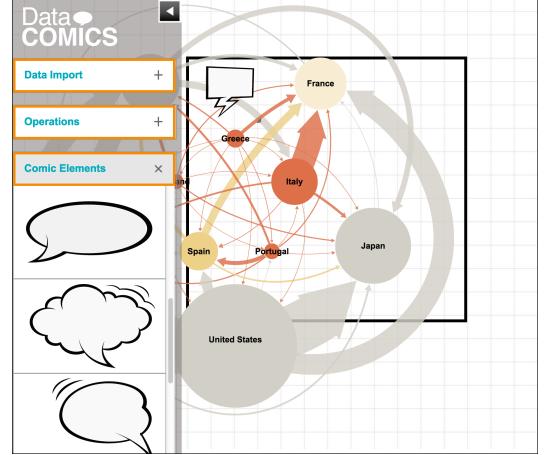
Zhenpeng Zhao, Rachael Marr, and Niklas Elmquist

University of Maryland, College Park, MD, USA

zhaoz@umd.edu, rachael.marr@gmail.com, elm@umd.edu



(a) Data Comic of the European debt crisis.



(b) DATACOMICSJS tool used to construct the comic.

Figure 1. A Data Comic built using our DATACOMICSJS tools that is based on datasets, visualizations, and images clipped from the web. The Presenter mode allows for viewing the comic as a whole, or viewing the panels in sequence. Comic-style symbols and annotations provide emphasis.

## ABSTRACT

We present Data Comics, a novel method for storytelling using sequential art—also known as *comics*—constructed from data-driven visualizations. This allows for building narratives using comic layouts of panels containing both snapshots and live visualizations. Each panel in a comic layout can be decorated with visual comic symbols—such as captions, speech and thought bubbles, directional arrows, and motion lines—to augment the narrative. To validate our method, we implemented a web-based Data Comics application that consists of (1) a Clipper for capturing data-driven content from the web, (2) a Decorator for creating panels and adding comic symbols, (3) a Composer for arranging clips into comic strips, and (4) a Presenter for viewing a finished comic. We compared the method to PowerPoint slideshows in a qualitative study, and found that participants perceived Data Comics as more engaging, efficient, and enjoyable.

## Author Keywords

Comics; storytelling; visual communication; narrative visualization; sequential art; evaluation; presentation.

## ACM Classification Keywords

H.5.2 Information Interfaces and Presentation: User Interfaces—*Interaction styles*; I.3.6 Computer Graphics: Methodology and Techniques—*Interaction techniques*

## INTRODUCTION

Comics—also known as *sequential art* [29] or *sequential images* [9]—represent a casual yet effective and engaging way of telling stories using a juxtaposition of illustrations, text, and visual annotations [8, 9]. By now an established—if not always respected [15]—narrative medium, comics combine the powerful visual language of film and pictograms with the evocative and imaginative nature of books to tell stories on any topic [6]. Because of their popularity, many people around the world have experience with comics from childhood, and sometimes even from adult graphic novels later in life. This means that they already have a common ground [7] with the visual language of comics for conveying narrative structure through time, motion, and action. For this reason, comics have also been harnessed for more “serious” applications beyond entertainment, including statistics [18], physics [17], and even the study of comics itself [29].

In this paper, we advance the idea of using the comics medium as a vehicle for telling stories about data (Figure 1):

*Data Comics* (DC),<sup>1</sup> for data-driven comics. Pioneered by Gershon and Page’s seminal paper on the topic [16], storytelling has recently received significant attention in the field of visualization and sensemaking, including workshops at IEEE VisWeek 2010 and 2011 [11, 12], surveys [20, 33], and actual techniques (e.g., [14, 25, 28]). Even the commercial tool Tableau includes a presentation feature called story points since 2013 [26]. Using comics as the narrative medium for storytelling builds on this trend, and has the benefit of leveraging the visual language of comics that many of us already have learned by heart, as well as bringing its engaging, vivid, and often whimsical nature to sensemaking.

But what is a Data Comic, and, more importantly, how does one create one? In general, comics depend on mastery of visual artistry and design as well as intimate knowledge of narrative techniques, so creating them often falls outside the skills of the average person. For Data Comics, however, the visuals are understood to be mainly data-driven, with only comic-style visual elements added for effect, engagement, and annotation. This implies that a Data Comic can be created by juxtaposing multiple visualizations into comic strip layouts consisting of a sequence of panels, each appropriately annotated and decorated with both visual and textual elements, and arranged into a sequence that progressively develops the overarching story told in the comic.

To facilitate the creation of Data Comics, we present DATACOMICSJS (Figure 1(b)), a Google Chrome extension that consists of four components: (1) the Clipper, for collecting both snapshots of visualizations and images as well as raw data from any webpage viewed in the browser; (2) the Decorator, for editing the visual design of an individual panel, including clips, images, captions, and comic-style visual elements; (3) the Composer, for managing the layout, size, and position of panels making up the comic; and (4) the Presenter, for ultimately allowing a viewer to navigate in a finished Data Comic, including viewing the entire comic as a whole, as well as view single panels in sequence. Note that DataComicJS is **not** a replacement for PowerPoint or any other presentation software. Our purpose is merely to show that Data Comics is an effective mechanism for telling stories about data.

We validate our ideas on Data Comics in two separate ways: (1) using three examples of Data Comics we have created using our tool based on real-world datasets and visualizations from the Internet, and (2) using a qualitative evaluation comparing the Data Comics method to standard PowerPoint presentations. Our results from the evaluation study do not show any significant difference in knowledge acquisition between the two presentation methods, but does indicate that participants were significantly more engaged and interested when using Data Comics compared to the PowerPoint slideshows. This confirms our intuition about the Data Comics method: that it is a more compelling and exciting narrative mechanism than traditional slideshows.

---

<sup>1</sup>The acronym’s similarity to the name of influential comic book publisher DC Comics—steward of well-known comic book characters such as Superman, Batman, and Wonder Woman—is coincidental but fortuitous.

The remainder of this paper is organized as follows: We begin by reviewing the literature on narrative visualization. We then present our design space for Data Comics. This is followed by a description of the implementation of the Google Chrome extension we built to realize the new method. We then review our example Data Comics, followed by our evaluation methods and our results. We close with a discussion, our conclusions, and our plans for future work.

## RELATED WORK

Our data-driven storytelling method using comics lies at the intersection of visual communication, storytelling, and sequential art. Below we review relevant work in these areas.

### Visual Communication and Visualization

Visual aids such as images, signs, typography, icons, and drawings have long been used as a particularly effective medium for communication [35]. Beyond human perceptual factors, part of the reason for this effectiveness is the mutual knowledge, mutual beliefs, and mutual assumptions that visual communication enjoys. Commonly dubbed “common ground” [7], these mutually agreed-upon conventions allow a particular medium—such as the visual—to encode significant amounts of information given minimal resources.

Visualization, defined as the use of graphical representations of data to augment cognition [3], is a particular form of visual language traditionally used for solitary sensemaking. The notion of communication-minded visualization (CMV) [36] builds on ideas from visual communication by noting that visualization can often be used for more than just individual insights. A precursor to storytelling with visualization, several examples of CMV systems exist, including Themail [37], the Baby Name Explorer [38], and Isis [31].

### Visual Storytelling

Storytelling, where a sequence of events are conveyed using plot, locations, and characters, is a particularly effective communication medium because of the typically high degree of common ground shared between narrator and listeners. Visual storytelling draws upon imagery—both static and dynamic—for this purpose, and includes media such as film, television, animation, design, and even art.

Already in 2001, Gershon and Page [16] suggested using storytelling in visualization to improve its use for visual communication. In fact, the newly popular infographics practice on the web builds on many such storytelling principles. Despite this, it is only recently that data-driven storytelling was fully embraced by the visualization community, with a survey by Segel and Heer in 2010 [33], and successful workshops at the annual conference in 2010 and 2011 [11, 12]. Hullman and Diakopoulos followed this up by studying how framing, context, and design impact the rhetoric of a narrative [20]. Since then, several practical methods and techniques have been proposed, including using free-form sketching for narration [28], story points in Tableau [26], and automatic spatialization for visual exploration [23]. Most recently, Hullman et al. [21] studied sequence in narrative visualization, proposing a graph-driven approach for transitioning between views to minimize load on the viewer.

## Comics as a Communication Medium

Comics are often defined as sequences of images—“sequential art” [29] or “sequential images” [9]—that combine to tell a story using graphical means [15]. It is therefore a visual communication medium, and, by virtue of being popular particularly for children and teenage audiences, but also for adults, enjoys a high level of common ground for conveying information. Furthermore, the visual language of comics is often clear, concise, and intuitive [8, 9].

How comics affect the reader has long been a topic of study. Dorfman et al. [13] discussed the ideology in Disney comics from the perspective of culture and economics. McCloud [29] built on these ideas by suggesting that the engagement in comics mainly arises from the simplified and non-photorealistic appearance of faces and characters, increasing recognizability and facilitating imagination. In a way, this “vague and unspecific” nature and lack of fixation in comics helps bridge the gap between books and film [6].

Some efforts have tried to harness comics for computer graphics and visualization. The non-photorealistic rendering community uses cartoon-style rendering for artistic effect—examples include motion emphasis [10], squash and stretch deformation [4], and cartoon-style animation [5]—and it has since been employed in several computer games. For visualization, Jin and Szekely [24, 25] proposed a visual query environment that uses a comic-strip metaphor for querying and presenting temporal patterns. However, their system solely uses comics strip for layout, and does not leverage the full potential of comics as a visual communication medium.

## DESIGNING DATA COMICS

Data Comics is a visual storytelling method based on sequential images consisting of data-driven visual representations. Its purpose is to build engaging narratives about data. Our inspiration for this method came from a conflux of factors, including the recent focus on storytelling for visualization [27], the increasing use of comics for “serious” applications (e.g [17, 18, 29]), and the EuroVis 2011 keynote by Scott McCloud on comics as a medium in transition [30]. Our motivation is to both take advantage of the plethora of existing visualizations on the web as well as the familiar visual language of comics, including layout, characters, and comic elements such as motion lines, speech bubbles, and arrows.

To explore the design space of this idea, below we review an operational model for data comics and then discuss each of its aspects in turn, including creating panels, managing their layout, and letting a viewer navigate the final product.

### Basic Model

For the purposes of this paper, a comic consists of a sequence of *panels* organized into one-dimensional *tiers* (or strips) and separated by *gutters*, or spacing, between the panels [15, 29]. The panels in a tier are typically organized to be read from left to right to form a narrative (at least in Western cultures). Tiers can in turn be organized into *pages*, where each tier becomes a row separated by a vertical gutter, and several pages can be linked together into a *book* (or comic book).

## Panel Content

Unlike a normal comic, most panels in a Data Comic consist of visualizations that convey information using graphical means.<sup>2</sup> These could be simple and familiar statistical graphics such as barcharts, time-series charts, and piecharts, or more advanced visualizations such as treemaps [34], node-link diagrams, or even parallel coordinate plots [22], all depending on the visualization literacy of the intended audience and the instructional annotations in the panel.

Because of this focus on data-driven graphics, the designer is largely relieved of creating artistic content, which requires drawing skills that only few people have. Instead, the visual content can be constructed by either creating entirely new visualizations in a panel using raw data, or by clipping a snapshot from an existing visualization into the panel.

### Characters, Annotation, and Effects

A Data Comic would not truly be a comic if it did not also leverage the visual language of comics. Designers creating Data Comics can be given access to this in several ways:

- **Comic-style rendering and layout:** To emphasize the comic medium, clipped content can be redrawn using non-photorealistic rendering techniques (e.g. [10]).

*Benefit:* This puts the user in a comics state of mind, allowing them to engage their preexisting knowledge of comics.

- **Characters:** To drive the narration in a Data Comic, the designer should include characters that complement the data-driven visualizations. Because this requires artistic talent, Data Comic designers should probably be furnished with a visual library of existing characters.

*Benefit:* Putting a human face on the story makes the narrative more engaging and personal.

- **Comic elements:** The above library mechanism can also be used to provide designers with access to common visual elements used in comics, such as motion lines, highlights, or even onomatopoeia (words that mimic sounds).

*Benefit:* Comic elements direct users’ attention to specific parts of a visualization and propel the narrative.

- **Captions, speech, and thoughts:** The visuals in sequential art is typically scaffolded by text in both captions, speech balloons, and thought balloons [29, 15]. A Data Comic should support these same mechanisms.

*Benefit:* Narrative visualizations need captions and explanations to drive the story and guide the reader.

### Layout Management

McCloud [29] postulates that it is *closure*, the mental process that the reader performs to connect images in one panel to the next, that is the real narrative “magic” of comics. This is akin to the closure that readers routinely perform for books: the fact that not all information is given by the written word (or the comic) means that the reader will fill in the blanks and

<sup>2</sup>For engagement and effect, a few panels may consist solely of artistic content, but this puts corresponding artistic burden on the designer.

make the narrative come alive in their mind’s eye [6]. For this reason, the layout management of a Data Comic—the organization of panels into tiers and pages—is an important consideration in creating a narrative.

Several design aspects are important when creating a Data Comic. In terms of geometric size, we simplify our model somewhat by stipulating that each panel in the same tier has the same height. This leaves the designer with the ability to change the height of an entire tier, as well as the aspect ratio of individual panels. Furthermore, to facilitate easily constructing a narrative, the Data Comics model should allow a designer to easily change the order of individual panels.

### Viewing

Finally, after a Data Comic has been created, its ultimate purpose is to be viewed by its intended audience to convey its designer’s story (and message). Just like a traditional comic, the default view for a Data Comic is to view an entire page, with all of the panels visible. Since screens are different from the written page, however, it also makes sense to support a single-panel navigation mode, where the viewer can sequentially navigate backwards and forwards in the comic.

In a way, this navigation mode is not unlike that of a traditional slideshow, such as PowerPoint or Keynote. However, since our focus is on supporting communication-minded visualization [36] and social data analysis [19], it may be useful to allow viewers to markup, annotate, and comment on individual panels in the viewer.

### IMPLEMENTING DATA COMICS

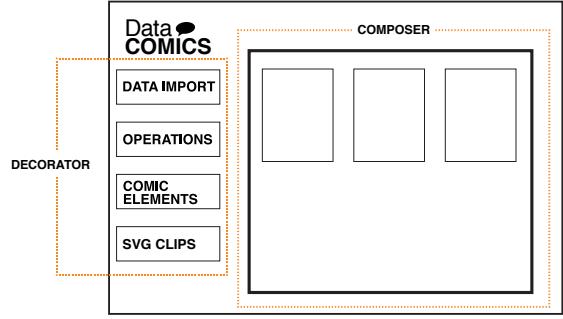
We have implemented Data Comics as a web application called DATAComicsJS. It is a hybrid application consisting of both client-side and server-side components communicating using JSON-RPC:

- **Clipper:** a mechanism for capturing remote content, such as visualizations, images, and raw data, for use in a comic;
- **Decorator:** a visual editor for creating panels consisting of visualizations, images, captions, and comic elements;
- **Composer:** a layout editor for arranging panels of visual content into a comic page; and
- **Presenter:** a viewer to allow the intended audience to navigate and discuss a Data Comic.

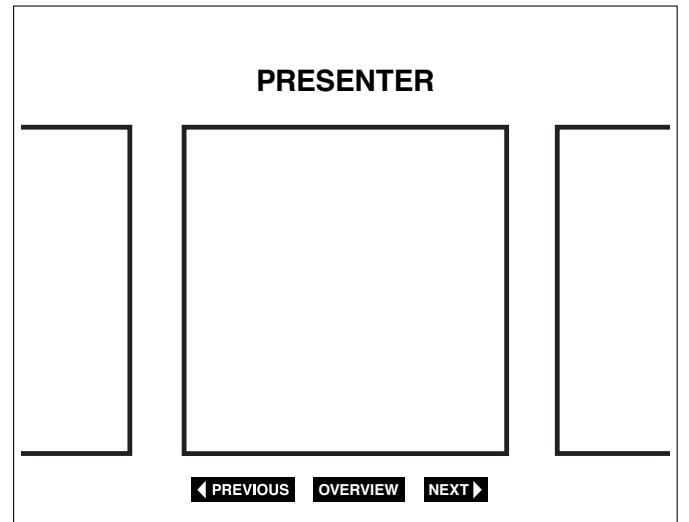
The layout of these components are illustrated by Figure 2 and Figure 3. All of these components are part of a web service that allows for creating, editing, and sharing Data Comics over the web. Below we review each of these components in turn.

### Clipper

The Clipper component of DataComicsJS is implemented as a Google Chrome extension that the user downloads and installs in their local browser. This allows the system to seamlessly integrate with and extend the browser so that users can clip content from any website they visit. This is achieved by traversing the DOM from the element that the user indicates. Three different types of content can be clipped:



**Figure 2.** Basic DataComicsJS interface with the Composer in the middle of the workspace and Decorator on the left side. The Decorator is composed of four expendable menus including data importing, operations, comic elements, and collection of clips.



**Figure 3.** The Presenter in slideshow mode.

- **Raw data:** Structured data, such as in an HTML table element or tab-delimited text, can be parsed and saved;
- **Snapshot (SVG):** Parts or all of an SVG element can be clipped, including any CSS that affect its appearance; and
- **Snapshot (raster):** A specified bounding box of the webpage can also be clipped as a raster screenshot.

The extension is implemented using JavaScript using the JQuery and JQueryUI libraries for DOM manipulation. The captured content is transferred to and stored on the server-side backend, implemented as a simple Python server, using JSON-RPC.

### Visual Editor: Decorator and Composer

After the user has collected the visual content to create a Data Comic, they would open the DataComicsJS client-side editor

itself.<sup>3</sup> The main editor consists of two separate yet interdependent components: the Decorator and the Composer.

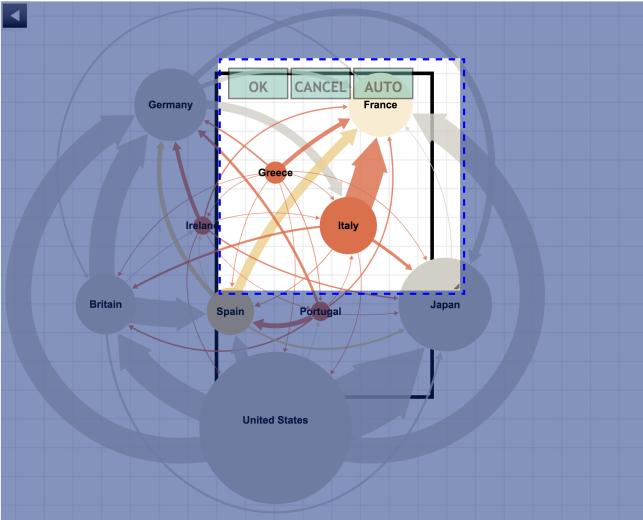


Figure 4. Cropping a portion of a visualization clip to include in a comic panel. A modal window will pop up to help the user control the cropping area. The user can use an auto-fit operation to resize a visual element to best fit the panel size while keeping its aspect ratio.

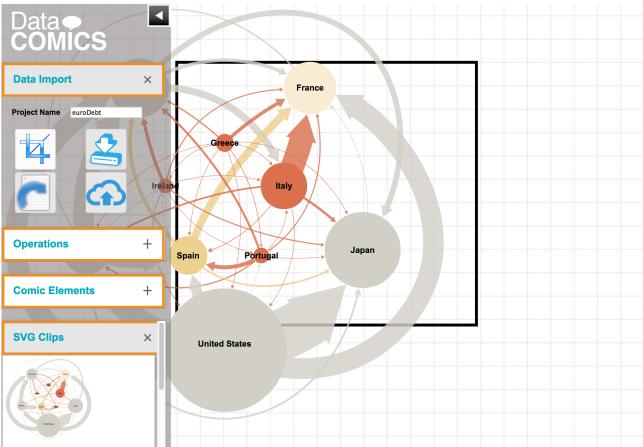


Figure 5. The Data Import section of the interface is used to load and update visualization clips from a server where the user can also save finished work and reload them for presentation. The SVG Clips section is the collection of all content captured by the clipper.

### Decorator

The panel is the basic building block of a Data Comic, and panels are created using the Decorator. The Decorator interface consists of four main components:

- **Toolbox:** The graphical tools that the designer use to create the panel, including transformation tools (translation, scaling, rotation), annotation, cropping (Figure 4), etc.
- **Workspace:** The main panel workspace is a blank area with panel borders. The borders can be dragged to change the aspect ratio of the panel. Any visual content placed inside

<sup>3</sup>The user can at any time use the Clipper—perhaps in a different browser tab—to add more visual content to the editor.

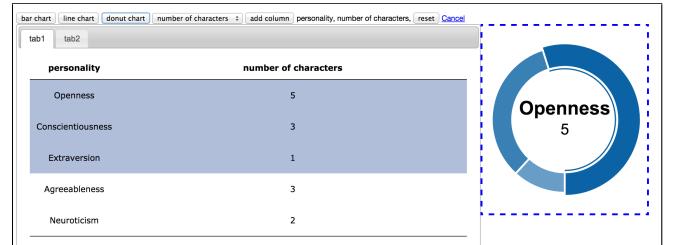


Figure 6. Raw data interface for inserting clipped data from an HTML table form into one of several types of visualizations; here, a donut chart.

the panel will be clipped to the border when viewed outside of the Decorator.

- **Clip collection:** The visual (and raw data) content that the user has collected using the Clipper. The user can simply drag and drop elements from the collection (located in left bottom corner of Figure 5) to the workspace to add them to the panel. Dropping a raw dataset into the panel brings up a dialog where the user selects which built-in visual representation to use to visualize the data; we currently support simple barcharts, donut charts (Figure 6), and line graphs, all implemented in D3 [2].
- **Symbols library:** To give access to a wide range of comic-style visual elements and annotations, we provide a library of such symbols in a visual format (Figure 1(b)). These can be added to the Workspace using standard drag-and-drop. We also provide a general interface for searching and dynamically loading vector graphics element from the free icon library of the Noun Project.<sup>4</sup>

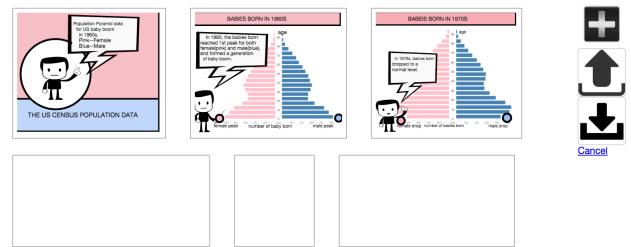


Figure 7. Composer interface allowing the user to add or remove panels using appropriate buttons (right). All panels in the same tier (row) share the same height; they can still have different width. The example shows an intermediate state of creating a Data Comic. Three frames are finished, and the rest of the three empty frames are ready to be filled in.

### Composer

The Composer gives the user control over the layout and organization of panels for a specific Data Comic; Figure 7. The layout flows to fit the page from left to right, top to bottom based on the number of panels and the page width. The interface provides buttons for adding and removing panels as well as dragging and dropping them to change their order. Double-clicking on a panel will open it in the Decorator.

For simplicity, all panels in the same tier (i.e., row of the comic) have the same height. This height can be changed

<sup>4</sup><https://thenounproject.com/>

by dragging on the row itself. Changing the aspect ratio of a panel in the Decorator will update the scaling of the panel on the main comic page.

Finally, the Composer interface also provides tools for providing meta-data about the comic, such as the author, title, and data sources, as well as for exporting the comic for dissemination to its intended audience.

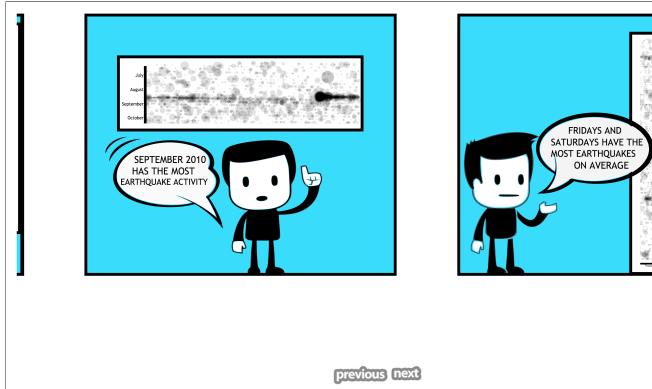


Figure 8. Viewing individual panels using The Presenter in slideshow mode. Buttons allow for navigating slides in sequence.

### Presenter

Once the designer has created and exported a Data Comic, he or she can share it widely using a unique URL that can be sent via e-mail, shared on social media, or posted on a forum. Everyone opening the URL will be given the Presenter view of the comic, which is a read-only viewer. The Presenter supports two viewing (and printing) modes: (1) viewing the entire comic, as is common for normal comics, and (2) viewing individual panels by navigating using on-screen buttons or the arrow keys (Figure 8)). The latter presentation mode is implemented as smooth animated transitions from one panel to the next to reinforce the comic metaphor.

We have yet to add any collaborative features to our DataComicsJS system, but we anticipate supporting two separate mechanisms in the future: (1) commenting and annotation for entire comics as well as individual panels, as well as (2) the ability to fork a comic to allow for further modification beyond the original (similar to GitHub or JSFiddle).

### DATACOMICS EXAMPLES

To validate the Data Comics method and give a concrete idea of what it is capable of, we present a few examples based on real-life data. The inspiration for these Data Comics comes from online data visualizations and infographics we experience on a daily basis. Our work here reformulates some of these insights using the comics medium.

#### Euro Debt Crisis

*Dataset:* The comic is based on a European Debt visualization available from The New York Times.<sup>5</sup> The visualization is a flow chart describing the debt relations between major European countries, the United States, and Japan.

<sup>5</sup><http://www.nytimes.com/interactive/2011/10/23/sunday-review/an-overview-of-the-euro-crisis.html>

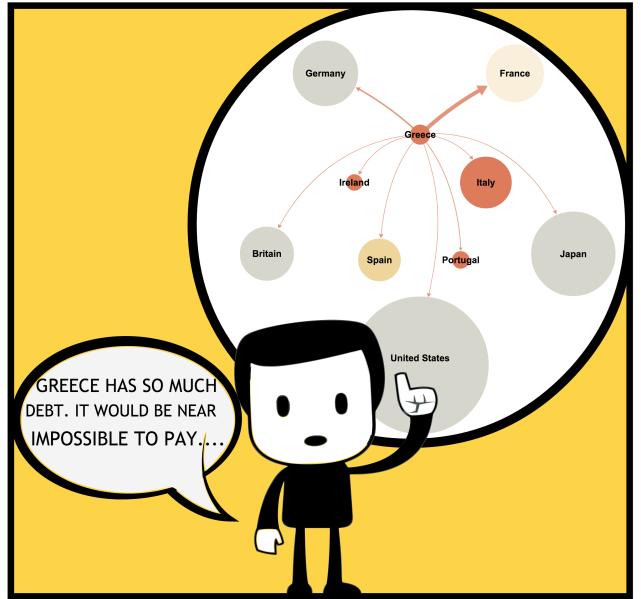


Figure 9. A Data Comic panel highlighting how the Greek debt crisis is the source of the European debt crisis.

*Data Insight:* The original visualization shows the direction and amount of the debt. It is obvious that the debt system is tangled, and the debt is huge. The Data Comic can tell the story in smaller, more manageable chunks.

*Construction and Visualization:* We snapshot the visualization and add comic figure and comic style text diagram to it. Figure 9 shows that the Greek debt crisis has a great negative impact on the economy of the whole Europe. The exaggerated lightning sign and the character figure is intended to show the seriousness of this crisis while adding a human angle. This shows how the Data Comic method can change the style of an existing visualization.

#### U.S. Census Population Pyramid

*Dataset:* Figure 10 is the Data Comics version of a U.S. Census dataset from 1960s to 2000s.<sup>6</sup>

*Data Insight:* The horizontal bar chart illustrates a population distribution based on an age scale and time scale in decades, partitioned by gender. We can clearly see the babies born in each decade and the trend of the population growth. We want to make a story about the population growth during the baby boom period and after to reveal how the babies born are the key factor to shape the population structure.

*Construction and Visualization:* To track the changes of babies born during the period of 1960 to 2000, we captured the bar charts of population pyramid during this time period, and then generated another line chart based on these data. Within each panel, we labeled dimensions such as age, number of babies, and changes in births from previous years. The panels clearly show that the number of babies born decreased

<sup>6</sup><http://vis.stanford.edu/jheer/d3/pyramid/shift.html>

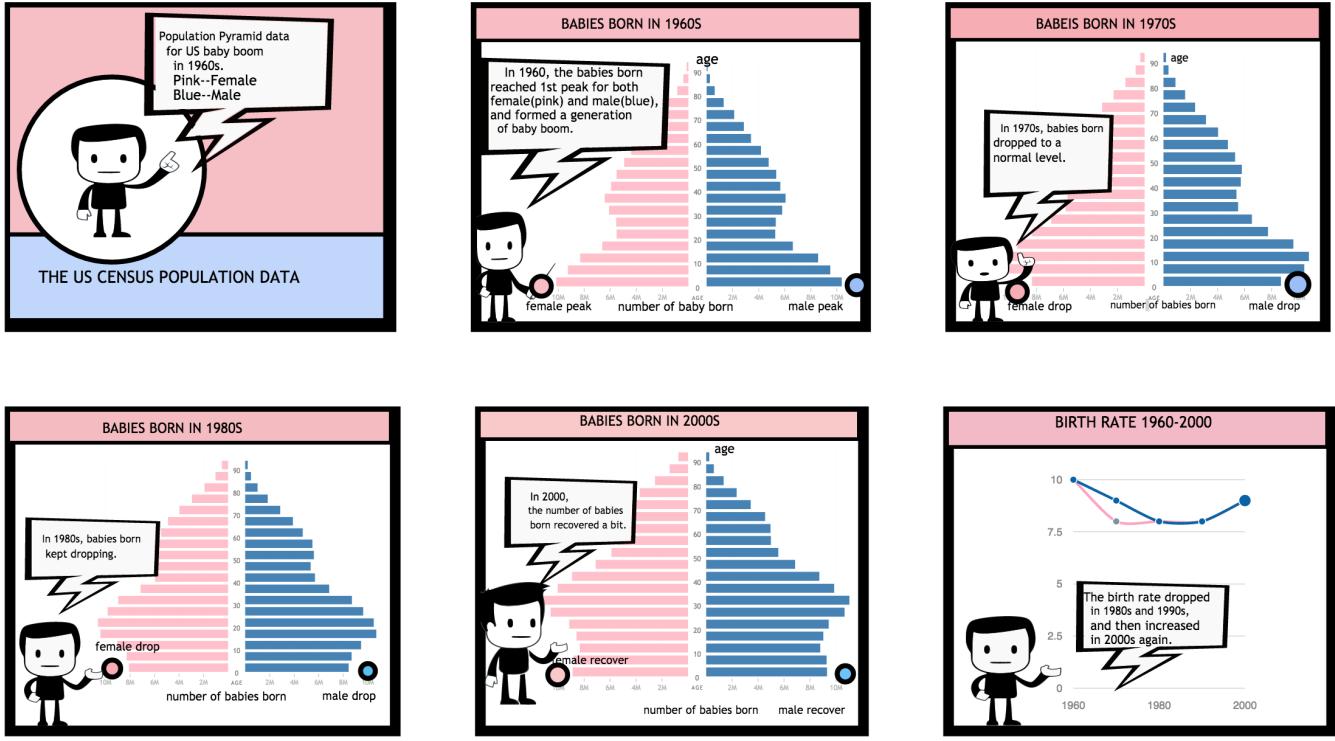


Figure 10. The U.S. Census data shows that there was a baby boom in the 1960s. The panels explain the population in each decade and shows a simple trend with line chart for both male and female babies born.

during the middle two of the four decades. Later, the number of babies born increased in the last bar chart. The comic symbology in this visualization is fairly simple and straightforward, but it can help with the story initialization and transition, which improves the users’ engagement to the system.

### Scientific Journal Comparisons

**Dataset:** The data is from a visualization comparing publication counts over time in scientific journals covering different topics in neuroscience and brain stimulation. Data gathered from the U.S. National Library of Medicine.<sup>7</sup>

**Data Insight:** The comparison among different publications over time is vastly distributed on the visualization space. Through comparison, we found that the Journal Brain Stimulations has steady growth in the field of transcranial magnetic stimulation.

**Construction and Visualization:** We captured the visualization, and then added a hand-shaped indicator and a female figure to help guide the narrative. As part of the whole story of six panels we built for this topic, in Figure 11, the Data Comic performs the role as a personalized guide highlighting the difference for comparison.

### QUALITATIVE EVALUATION

Our working premise in this paper is that Data Comics provide a compelling way of telling stories about data. To empirically explore the virtues of this premise, we conducted

a qualitative user study comparing Data Comics to traditional PowerPoint slideshows. The reason we chose PowerPoint is not to prove that our prototype implementation—DataComicsJS—is in any way superior to PowerPoint or any presentation software, but to pick an application and style of presentation widely used in the real world. Here we describe the methods and results from this evaluation.

### Participants

We recruited 12 paid participants (6 male, 6 female) to participate in the user study. The participants were self-selected from the student population at our university, were aged between 20 and 31 years of age, had normal or corrected-to-normal vision, and were proficient computer users (all demographics were self-reported).

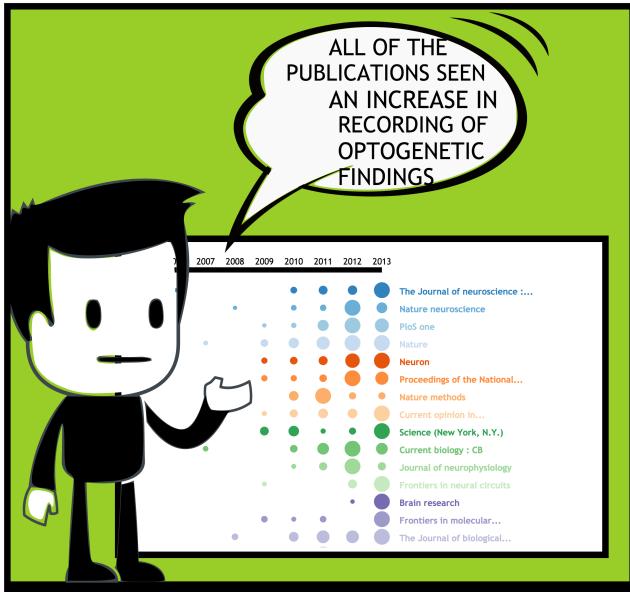
### Apparatus

We conducted the experiment on a standard laptop computer equipped with a 15-inch LCD screen (resolution 1280×800), a standard keyboard, and a three-button mouse. Both the Data Comics prototype and Microsoft PowerPoint was maximized to fill the entire screen during the experiment.

### Task and Datasets

Each trial consisted of the participant using either MS PowerPoint or Data Comics to answer questions about a data story. Each story consisted of several panels of narration. The number of panels were limited to five or six to keep the simplicity of layout while providing sufficient information for the story.

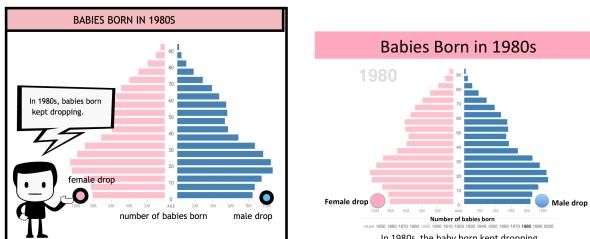
<sup>7</sup><http://neuralengr.com/asifr/journals/>



**Figure 11.** A Data Comic for scientific journals in neuroscience. The character in the Data Comic performs the role of a guide describing trends in the data, such as that there are more publications every year in optogenetics.

Each story presentation focused on a single topic of visualized data and came with an associated list of 7 to 9 questions designed to make the participant focus on details of the visualization. Participants had access to all questions during the time they were interacting with the story.

All in all, we created four separate stories for the evaluation: Twitter heatmap for stocks ( $S_1$ ), the U.S. Census Population pyramid ( $S_2$ ), the world happiness ( $S_3$ ), and Star Wars character fans' personality rating ( $S_4$ ).



(a) Panel in a story made by DataComicJS. (b) Slide in a story made by PowerPoint.

**Figure 12. A comparison of the test story frames with DataComics and PowerPoint.** To make sure the user study is fair, every story is composed with two different methods but with one-to-one correspondence between the visualizations and captions. This frame is about babies born data from the U.S. Census database.

The stories were first created as a Data Comic by clipping from online data sources and visualizations, creating an appropriate number of panels to tell the story, and finally decorating the panels with comic annotations, characters, and captions. We then created a corresponding PowerPoint slideshow with the exact same number of slides as the number of panels in the Data Comic. There was a one-to-one correspondence

between the visualizations and captions between panels and slides as Figure 12; the only difference between the two versions was that the PowerPoint slideshow only used visualizations and text, whereas the Data Comic included characters and comic-style symbology. Furthermore, while the text explanations were identical across versions, the Data Comic version integrated them in comic-style captions or speech and thought balloons.

### Metrics

Our focus with the evaluation was not to primarily study quantitative metrics, such as time and accuracy, but to collect subjective and qualitative feedback on the difference between Data Comics and traditional slideshow presentations. For this reason, we developed a questionnaire polling participants on their subjective experience of a story. This was administered to participants directly after each story, and consisted of the following 1–5 Likert-scale questions:

- Q1 Engagement:** How engaged and attentive did the participant feel by the story?
- Q2 Speed:** How fast did the participant feel they were in finding answers to questions?
- Q3 Space-efficiency:** How well does the story make use of visual space in the presentation tool?
- Q4 Ease of use:** How approachable and easy to use was the story to view and understand?
- Q5 Enjoyability:** How enjoyable was the experience of viewing the story to the participant?

Beyond the Likert-scale questions, we also asked participants for their freeform feedback on the story and the tool.

### Factors

We included two factors in the experiment, described below.

#### *Presentation ( $P$ )*

This factor modeled the presentation technique  $P$  given to the participant for solving questions in a particular trial:

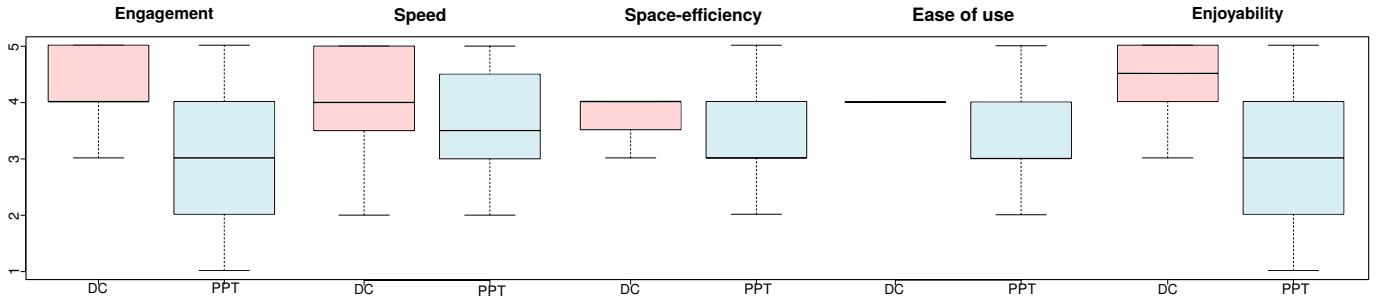
- **Data Comics:** The narrative visualization is presented as a Data Comic in the Presenter in our prototype implementation. Participants were able to view the entire comic, as well as navigate panel by panel in the comic.
- **PowerPoint:** The narrative is presented as a PowerPoint slideshow. Participants can navigate backwards and forwards in the slideshow. They were also able to view all of the slides at once in the “slide sorter” view.

#### *Story ( $S$ )*

We also hypothesized that the specific story and topic of the data visualization may impact our outcome. Thus we added a factor  $S$  to model the different stories.

### Procedure

An experimental session started with the participant arriving, reading and signing the consent form, and being assigned an identifier and story order. The administrator then explained



**Figure 13. Subjective ratings (Likert 1-5 scale) for engagement, speed, space-efficiency, ease of use, and enjoyment.**

the general goals and task for the experiment. Each trial started with the administrator demonstrating how to use a Data Comic. The participant was then given two examples, one Data Comic and one PowerPoint, and was allowed to ask questions about the examples and task during this time.

When the participant finished the training, they were given a story opened in the appropriate tool as well as a paper sheet with the corresponding questions. They were given up to 10 minutes to answer the questions, and were encouraged to use all of the time. After answering all questions, the participant was given the subjective questionnaire polling their experience in the trial. This was repeated for all four stories—two using Data Comics, two using PowerPoint.

A full experimental session lasted approximately 50 minutes, including training and questionnaires.

## Results

We organize our reporting of the results into quantitative versus qualitative aspects.

### Quantitative Results

Figure 13 depicts boxplots of the subjective ratings for both Data Comics and PowerPoint on engagement, speed, space-efficiency, ease of use, and enjoyability (Q1 through Q5). We analyzed the 5-point Likert scale of subjective ratings for effects of presentation technique  $P$  (Data Comic vs. PowerPoint), and found that the engagement (Q1), efficiency (Q3), and enjoyment (Q5) were significantly different between the two techniques (Friedman tests,  $p < .05$ ), but the speed (Q2) and ease of use (Q4) had no significant difference (Friedman tests,  $p = .51$  and  $p = .08$  separately). We also found no significant effect of story  $S$  on any of the metrics.

### Qualitative Feedback

**Inviting Reading:** Nine out of twelve participants mentioned that the comic-style rendering helped them view the materials as a whole story from the very beginning without any explicit direction. They noted that the speech balloon helps focus by creating a feeling that there is a virtual conversation going on, and the comic figure helped them more involved in the scenario of the story. All these comments from the participants suggest that Data Comics invite reading, even when incorporating only simple and trivial comic elements.

**Viewing as a Story:** The Presenter is organized to show not just the current panel, but also the two surrounding ones (Figure 3). While this nominally is a waste of visual space—

a slideshow shows each slide in full-screen mode, yielding more pixels to complex visualizations—participants seemed to enjoy this view, presumably because it suggests continuity and story flow (“there is more to see beyond this panel”) and it evokes the “comic state of mind” we seek. Our observations and interviews confirmed this fact; the extra context panels seem to encourage participants to keep reading. Also, our Likert scale results show that participants actually felt that comics were more space-efficient (Q3) than the slideshows.

In fact, we observed that all but two of our participants would start each Data Comic task by first reading through the entire comic from beginning to end. Thus, the comics format seems to invite reading. This is in contrast to the PowerPoint slideshows, which no participants were observed to read fully before answering questions. Several participants remarked that the slideshows did not “feel” like stories, but rather information sheets that they were just flipping through to find information. While there is no intrinsic value to this story aspect of Data Comics, we do think it increases user engagement, as evidenced by our quantitative results.

**Facilitating Memory:** Eight out of twelve participants mentioned that the comic version of each story helped them remember the contents, even down to the individual panel for specific information. Figures are naturally memorable, even when they are not relevant to the current topic; this mirrors findings by Bateman et al. [1] on the beneficial effect of “chart junk” on recall in visualizations and infographics. We also noted that a Data Comic does not need to be designed in a very artistic way, but can incorporate basic clipart-like imagery and graphics. Participants also mentioned that even a little variation of comic figures can help distinguish frames.

## DISCUSSION

Our qualitative evaluation indicated that Data Comics were significantly more engaging, space-efficient, and enjoyable to use than PowerPoint slideshows. While this may partly come from a novelty effect, this is still an intuitive result since many people associate comics with entertainment and fun. It confirms our belief that comics are generally a more compelling medium than traditional slideshows. Furthermore, this is supported by qualitative feedback from our participants that was overwhelmingly in favor of the comic representation.

We explicitly chose not to measure the completion time or correctness in our user study. We do not think that there is much difference in the communicative power of Data Comics

versus PowerPoint slideshows, and this perception was also confirmed by participants in our experiment. While significant differences in objective measures would have provided powerful evidence in favor of Data Comics, such results are notoriously difficult to attain for high-level sensemaking tasks because of individual differences between participants such as existing knowledge, varying interpretations of the task, and mixed visual and cognitive abilities [32]. We also think that such objective measures miss the point—our focus is on storytelling and engagement. Our approach is further validated by Lee et al. [28], who chose to only collect subjective ratings from participants in their SketchStory evaluation.

Our implementation of Data Comics is only one point in the solution space, and the general idea of sequential art for storytelling can be applied to other settings. For example, one approach may be to add comic storytelling features, such as panel layouts, visual annotations, or speech and caption, to existing visualization systems. This may lead to more enjoyable and more compelling storytelling mechanisms than traditional ones, such as integration with PowerPoint slideshows. Furthermore, while our DataComicsJS system is aimed at allowing data analysts with little to no artistic talent to create Data Comics, we think that the basic method of using comics to tell stories can be used to great effect by real designers and comic book artists. We hope to see many such Data Comic infographics in the future.

It is important to note that all of the functionality of the DataComicsJS tool can be replicated in presentation tools—such as Microsoft PowerPoint—and drawing tools—such as Adobe Illustrator—with sufficient time and effort. While DataComicsJS makes constructing Data Comics easy with its clipping functionality, comic layout and elements, and Noun Project integration, each of the Data Comics showcased in this paper can be built using other tools. However, our contribution in this paper is not the tool, but the Data Comics idea of using sequential art for data-driven storytelling.

Our work in this paper has several limitations. First of all, much of our argumentation of using sequential art for data is based on two assumptions: that the audience has (a) prior experience, and (b) a favorable opinion about comics. With no prior experience, much of the benefit of an established common ground in the visual language of comics is lost. Furthermore, given the sometimes questionable respectability of comics [15, 29], its use as a communication medium may be problematic. For example, it can be argued that a Data Comic may not be the best vehicle for presentations in very formal settings, such as a boardroom meeting. Similarly, the intrinsically light-hearted nature of comics may be inappropriate for sensitive or difficult topics, such as natural disasters, emergency situations, and other types of crises or stories on the loss of lives or livelihoods.

Another weakness with the Data Comic method is that it does require a certain time investment to create; typically more than would be required for a comparable PowerPoint slideshow. Even if our DataComicsJS tool is designed to facilitate this process, there are certainly more steps involved in creating a comic than a slideshow because of the need for

narrative structure and comic-style annotation, steps that cannot easily be automated. In fact, to take full advantage of the comic medium, the designer may want to find (or draw!) customized characters and annotations to fit the narrative, which can be very time-consuming. Furthermore, while a Data Comics authoring tool such as ours can help suggest appropriate ideas, the process also requires some knowledge of storytelling techniques on behalf of the designer.

Finally, it is also important to note that adding multiple visualizations into the same Data Comic may yield high visual complexity. This can be a problem both for viewers as well as the visualization system that is rendering the resulting comic. These issues should be considered and appropriately balanced by the designer of the Data Comic.

## CONCLUSION AND FUTURE WORK

We have presented Data Comics, a new approach to capturing and editing existing online data visualization to tell stories about data. We have implemented a hybrid web application that consists of a Google Chrome extension for clipping content, a visual editor to create panels consisting of visualizations and comic-style symbols, a composer to arrange the panels into entire comics, and a viewer for disseminating the results to a wider audience. Results from a controlled user study indicate that Data Comics significantly outperforms PowerPoint in terms of engagement, space efficiency, and enjoyability. This is consistent with the power of using the comic medium for narration in general.

Our future work will be to continue exploring narrative visualization using sequential art. For example, perhaps there are situations when it is appropriate to combine dynamic visualizations with a static comic. Furthermore, we also want to incorporate more collaborative features into the system. We also plan to add SVG editing functionality to DataComicsJS to make the system even more versatile. Finally, we are also interested in exploring other storytelling media beyond comics for narrative visualization.

## ACKNOWLEDGMENTS

[Anonymized for double-blind review.]

## REFERENCES

1. Scott Bateman, Regan L. Mandryk, Carl Gutwin, Aaron Genest, David McDine, and Christopher A. Brooks. 2010. Useful junk?: the effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, 2573–2582. <http://doi.acm.org/10.1145/1753326.1753716>
2. Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. 2011. D<sup>3</sup>: Data-Driven Documents. *IEEE Transactions on Visualization & Computer Graphics* 17, 12 (2011), 2301–2309.
3. Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman (Eds.). 1999. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann, San Francisco, CA.

4. Stephen Chenney, Mark Pingel, Rob Iverson, and Marcin Szymanski. 2002. Simulating Cartoon Style Animation. In *Proceedings of the International Symposium on Non-Photorealistic Animation and Rendering*. 133–138.
5. Sung-Bae Cho, Kyung-Joong Kim, and Keum-Sung Hwang. 2007. Generating Cartoon-Style Summary of Daily Life with Multimedia Mobile Devices. In *New Trends in Applied Artificial Intelligence*. Vol. 4570. Springer Berlin Heidelberg, 135–144.
6. Hans-Christian Christiansen. 2000. Comics and film: a narrative perspective. In *Comics & culture: Analytical and theoretical approaches to comics*, Anne Magnussen and Hans-Christian Christiansen (Eds.). Museum Tusculanum Press, University of Copenhagen, 107–122.
7. Herbert H. Clark and Susan E. Brennan. 1991. Grounding in Communication. In *Perspectives on Socially Shared Cognition*, Lauren B. Resnick, John M. Levine, and Stephanie D. Teasley (Eds.).
8. Neil Cohn. 2013. Beyond speech balloons and thought bubbles: The integration of text and image. *Semiotica* 197 (2013), 35–63.
9. Neil Cohn. 2014. *The Visual Language of Comics: Introduction to the structure and cognition of sequential images*. Bloomsbury, London.
10. John P. Collomosse, D. Rountree, and P. M. Hall. 2005. Rendering cartoon-style motion cues in post-production video. *Graphical Models* 67, 6 (2005), 549–564.
11. Nick Diakopoulos, Joan DiMicco, Jessica Hullman, Karrie Karahalios, and Adam Perer. 2011. Telling Stories with Data: The Next Chapter—A VisWeek 2011 Workshop. (2011).
12. Joan DiMicco, Matt McKeon, and Karrie Karahalios. 2010. Telling Stories with Data—A VisWeek 2010 Workshop. (2010).
13. Ariel Dorfman and Armand Mattelart. 1984. *How to Read Donald Duck: Imperialist Ideology in the Disney Comic*. Intl General.
14. Ryan Eccles, Thomas Kapler, Robert Harper, and William Wright. 2008. Stories in GeoTime. *Information Visualization* 7, 1 (2008), 3–17.
15. Will Eisner. 2008. *Graphic Storytelling and Visual Narrative*. W. W. Norton & Company.
16. Nahum D. Gershon and Ward Page. 2001. What Storytelling Can Do for Information Visualization. *Commun. ACM* 44, 8 (2001), 31–37.
17. Larry Gonick and Art Huffman. 1990. *The Cartoon Guide to Physics*. HarperPerennial, New York.
18. Larry Gonick and Woollcott Smith. 1993. *The Cartoon Guide to Statistics*. HarperCollins, New York.
19. Jeffrey Heer, Fernanda B. Viégas, and Martin Wattenberg. 2007. Voyagers and voyeurs: Supporting asynchronous collaborative visualization. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1029–1038.
20. Jessica Hullman and Nicholas Diakopoulos. 2011. Visualization Rhetoric: Framing Effects in Narrative Visualization. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (2011), 2231–2240.
21. Jessica Hullman, Steven M. Drucker, Nathalie Henry Riche, Bongshin Lee, Danyel Fisher, and Eytan Adar. 2013. A Deeper Understanding of Sequence in Narrative Visualization. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2406–2415.
22. Alfred Inselberg. 1985. The plane with parallel coordinates. *The Visual Computer* 1, 2 (1985), 69–91.
23. Waqas Javed and Niklas Elmquist. 2013. ExPlates: Spatializing Interactive Analysis to Scaffold Visual Exploration. *Computer Graphics Forum* 32, 3pt4 (2013), 441–450.
24. Jing Jin and Pedro A. Szekely. 2009. QueryMarvel: A visual query language for temporal patterns using comic strips. In *Proceedings of the IEEE Conference on Visual Languages and Human-Centered Computing*. 207–214.
25. Jing Jin and Pedro A. Szekely. 2010. Interactive querying of temporal data using a comic strip metaphor. In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology*. 163–170.
26. Robert Kosara. 2013. Story Points in Tableau Software. Keynote at Tableau Customer Conference. (Sept. 2013).
27. Robert Kosara and Jock D. Mackinlay. 2013. Storytelling: The Next Step for Visualization. *IEEE Computer* 46, 5 (2013), 44–50.
28. Bongshin Lee, Rubaiat Habib Kazi, and Greg Smith. 2013. SketchStory: Telling More Engaging Stories with Data through Freeform Sketching. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2416–2425.
29. Scott McCloud. 1994. *Understanding Comics: The Invisible Art*. William Morrow Paperbacks.
30. Scott McCloud. 2011. Comics: A Medium in Transition. *Computer Graphics Forum* 30, 3 (2011), xiii.
31. Doantam Phan, Andreas Paepcke, and Terry Winograd. 2007. Progressive Multiples for Communication-minded Visualization. In *Proceedings of Graphics Interface*. 225–232.
32. Catherine Plaisant. 2004. The challenge of information visualization evaluation. In *Proceedings of the ACM Conference on Advanced Visual Interfaces*. 109–116.
33. Edward Segel and Jeffrey Heer. 2010. Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization & Computer Graphics* 16, 6 (2010), 1139–1148.

34. Ben Shneiderman. 1992. Tree Visualization with Tree-Maps: A 2-D Space-Filling Approach. *ACM Transactions on Graphics* 11, 1 (Jan. 1992), 92–99.
35. David Sless. 1981. *Learning and Visual Communication*. Wiley.
36. Fernanda Viégas and Martin Wattenberg. 2006. Communication-Minded Visualization: A Call to Action. *IBM Systems Journal* 45, 4 (2006), 801–812.
37. Fernanda B. Viégas, Scott Golder, and Judith Donath. 2006. Visualizing email content: portraying relationships from conversational histories. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 979–988.
38. Martin Wattenberg. 2005. Baby Names, Visualization, and Social Data Analysis. In *Proceedings of the IEEE Symposium on Information Visualization*. 1–7.