DataTV: Streaming Data Videos for Storytelling

Zhenpeng Zhao, Niklas Elmqvist*
University of Maryland, College Park

March 2016

Abstract

Data videos—motion graphics that incorporate visualizations—have been recognized as an effective way to communicate ideas, but creating such video requires both time and expertise, precluding them from being created and streamed live. We introduce DataTV, a system for combining multiple media sources in real time. We validate our work through an expert review involving researchers using the DataTV prototype to create a one-minute data video for their current project. Results show that the new method facilitates rapid creation and enables users to focus on the narrative rather than mechanics of video production.

Keywords: Data videos, visual storytelling, data visualization, evaluation, video stream, live recording, real-time streaming.

1 Introduction

Applying the age-old practice of storytelling to data visualization [13] is one of the most transformative developments in recent history for visualization research [19]. As this approach matures, practitioners and researchers alike are adding to the types of storytelling media being used for this purpose [26]; beyond the traditional slideshow [18], recent efforts are now exploring the use of infographics [26], hand-drawn comics [4], and interactive sketching [20]. The use of motion graphics, or so-called data videos [1], is particularly interesting as it allows for creating compelling visual narratives using rich multimedia sources, such as live video, graphics, text, sound, and music. However, authoring effective videos that utilize the medium to its fullest is time-consuming, resource-intensive, and requires a significant amount of skill. In particular, this precludes such videos from being created and streamed live on the internet.

In this paper, we study how to enable dynamically streaming visual exploration for storytelling. Such a live “data TV” approach will require a streamlined and fully integrated software platform where a single person can prepare, record, and produce the digital stream with a minimum of time investment. To enable the creation of such live data videos, we propose the DataTV platform, a prototype system for authoring live-streaming data videos using an integrated interface. The prototype supports three separate modes for (1) production, (2) recording, and (3) editing, in a highly streamlined and optimized workflow that allows a single content creator to control the entire process during live streaming.

To validate the DataTV platform, we asked two visualization experts to create a short data video using DataTV on a specific topic to collect real experiences on using the system for storytelling. Results from this expert review (participant-generated clips are included in the companion video for this paper) indicate that these researchers were impressed with the relative ease they could create these videos. Interestingly, our findings suggest that the integrated functionality in the DataTV platform enabled the researchers to focus on storytelling rather than the mechanics of video production. Their feedback informed further improvements to the platform.

2 Background

Motion graphics has recently been proposed as a particularly effective method for data-driven storytelling [1]. Here we review the background on visual communication, visual storytelling, and the state-of-the-art in storytelling for data.

*e-mail: elm@umd.edu
2.1 Production, Presentation, and Dissemination

The research agenda for visual analytics called for a focus on the production, presentation, and dissemination of analytics results [28]:

• The case for communication: Viégas and Wattenberg remark upon the proclivity of visualization for communication by virtue of its graphical form, and encourage focusing on so-called communication-minded visualization [30] where communication enables collaborative analysis.

• Embedded dissemination: To reach its full potential, communication capabilities should be integrated into the visualization tools themselves [28]; for example, Tableau now incorporates the story points feature [18], and most commercial tools supports exporting interactive dashboards to the web.

• Literacy and casual viewers: Presenting insights from data to the masses requires taking the visualization literacy [7] of everyday users into account. Thus, the notion of “casual visualization” [24] is important.

2.2 Visual Storytelling

Storytelling is the conveyance of a sequence of events involving characters and places—stories—and has a history spanning thousands of years [25]. A visual narrative is a story told primarily using visual media, such as illustrations, photographs, animations, video, and—now—visualization [12] 27.

• Richness in visuals: Visual forms of communication, including icons, illustrations, schematics, photographs, and full-motion video, have long been considered to be one of the most compelling storytelling formats [8] 27.

• Video tutorials: Complex skills are today often learned through step-by-step instructions on the internet [15]. Several projects have endeavored to increase the richness of such video tutorials: ToolClips [15] provide contextual access to video assistance, Pause-and-Play [23] links tutorials to the real application, MixT [9] automatically generates video from demonstrations, and a recent approach makes software tutorial videos interactive [21].

• New storytelling formats: Beyond traditional providers (HBO, MSNBC, ABC, etc) as well as the new breed (Amazon Prime, Hulu, and Netflix), open video sharing sites such as YouTube have democratized access to visual storytelling. With such openness comes entirely new storytelling formats that were previously infeasible, such as vlogs (video blogs) [14], reaction videos [3], and video commentaries, including the ever-popular Let’s Play videos [31].
• **Live streaming**: An additional recent development in internet video is the increased focus on live, so-called *streaming*, content, where the media is received and presented to the consumer at the same time it is delivered by the provider. This development was particularly driven by live gameplay content, where prominent gamers—called *streamers*—broadcast their game screen on services such as Twitch.tv. While streaming is also used for heavily produced events such as eSports tournaments, where professional gamers compete over thousands of dollars in prize money in games such as *Overwatch*, *League of Legends*, and *Dota 2*, the vast majority of live streams on Twitch and elsewhere are created using specialized *broadcasting software* such as OBS, XSplit, GameShow, etc. Such software are what inspired Data TV, and are designed for creating high-quality streams using an optimized user interface that can be controlled during live broadcast by a single user.

### 2.3 Data-Driven Storytelling

Combining communication-driven visualization with storytelling yields *data-driven storytelling*: narrative techniques for telling stories about data [25].

• **Storytelling in visualization**: Gershon and Page first proposed using storytelling for visualization [13], and their work has since been followed up by workshops [10][11], surveys [16][26], and even commercial tools [18].

• **Multiple storytelling media**: The purview of data-driven storytelling has quickly grown, from dashboards and slideshow presentations [18] to more esoteric formats such as sketching [20], journeys in time and space, and even comics [4][17]. In their survey of narrative visualization, Segel and Heer [26] identify seven genres—magazine style, annotated charts, partitioned posters, flow charts, comic strips, slideshows, and videos—and also suggest that future storytelling approaches may combine genres.

• **Data videos**: Amini et al. [1] recently identified *data videos* as motion graphics combining both sound and visuals to tell a data story. Pointing to prominent examples from the New York Times and the Guardian, their work encourages professional storytellers to use visuals to craft their narratives. Their follow-up Data Clips [2] work is an authoring tool for creating data-driven clips incorporating visualizations that can be assembled into longer data videos. While obviously deeply influential to our work, their treatment focuses on the careful and deliberate production of data videos through ideation, sketching, storyboarding, capture, and editing; live streaming is not covered or even considered. Furthermore, DataTV provides support for including additional content from the streamer’s computer, such as video, sound, music, and existing applications (such as Tableau or Spotfire), merely using the ability to capture the user’s full desktop.

### 3 Design: Streaming Data Video Production

We claim that there is a need for a multimedia platform for creating such *live data videos* at a pace and scale where they can be streamed and uploaded to an online video sharing service, such as YouTube or Twitch. Here we describe the major design decisions of the DataTV platform we design to meet this need.

**D1 Standalone application**: We design our tool to be a standalone desktop application rather than a web-based one.

- Video production and real-time streaming requires high performance processing and significant storage.
- No existing streaming software is entirely web-based; in fact, some even employ specialized hardware.
- **Alternative**: A web-based tool is platform-agnostic, but real-time video capture is resource-intensive.

**D2 Web integration**: We embed a web browser as a capture source to support web-based visualizations.

- Toolkit such as D3 [6] have made the web a unified platform for delivering visualization to the masses.
- Modern web browsers are full-featured multimedia platforms supporting a wide range of content, including video, sound, speech, vectors, 3D, etc.
- **Alternative**: A web-based tool would trivially support the web, but is impractical due to performance needs.

**D3 Optimized workflow**: The tool supports streaming with a single user acting as talent, engineer, and producer.

- Sustainable workflows for creating streams on a weekly or even daily basis must be time-efficient.
- Most streamers on Twitch—even established ones with thousands of subscribers—operate alone.
- **Alternative**: Abandoning real-time control would prevent streaming.

**D4 Native media support**: We provide source handlers for many media types, such as video, music, webcams, etc.
– Effective data videos incorporate multiple media types beyond “just” the visualizations themselves [1].
– Capturing directly from on-screen windows trivially enables native support for all applications.
– Alternative: Using special software for specific media breaks the workflow, requires expertise, and reduces efficiency, but requires integrating all of the media handlers.

D5 Simplified video production elements: We design the tool to provide simplified video production operations using easily accessible actions.
– Typical analysts do not have a background in video editing, much less storytelling using motion graphics.
– Providing the building blocks of video production will help creators think in terms of storytelling rather than mundane tool operations.
– Alternative: Dedicated video editors have a richer set of video production elements, but break the workflow.

4 DataTV: A Streaming Data Video Editor

We present DATATV, a prototype data video streaming utility. DataTV is a standalone desktop tool for multiple platforms that allows recording multiple video sources from any number of windows on the user’s desktop.

Figure 2: Main user interface of the DataTV prototype tool. A live mode toolbar allows for panning and zooming sources as well as scribbling directly on top of the output. The list panes at the bottom allow for controlling the scenes and sources being displayed.
4.1 Workflow

The DataTV tool is modeled along the standard workflow used in game streaming software such as OBS, XSplit, and GameShow, where the streamed output of the tool is managed using the concept of sources and scenes:

- **Source**: Streaming input such as from a window, specific application, audio source, or multimedia content.
- **Scene**: Composition of sources on an empty display that can be recorded and streamed as output from the tool.

DataTV operates in one of two distinct modes: (a) offline mode, where the user configures sources and scenes, or (b) live mode, where scenes are recorded (and possibly streamed). Most operations (creating, modifying, and deleting scenes and sources) can be performed in both modes to allow for responding to live events (e.g., adding a new web-based visualization to the stream on a spur-of-the-moment idea), but the most common workflow is as follows:

1. **Preparation**: The user prepares all scenes and sources for recording in offline mode. This involves creating the sources, creating the scenes, and composing the sources for each scene on the output display canvas.
2. **Recording**: The user switches to live mode, sets up the stream settings (if enabled), and begins the recording. If the video is streamed live, the user cannot easily switch back to preparation; instead, any changes to scenes or sources must be made while recording. If the video is merely being recorded and not streamed, the user can stop recording and go back to preparation, allowing the clips to be edited together at a later stage.
3. **Storytelling**: The user creates the data video in live mode. This entails switching between scenes (by selecting the scene to display in the scene manager), managing specific sources (transforming, toggling on and off, annotating, etc), interacting with visualizations and other windows, and potentially narrating and/or capturing webcam video of the user.

4.2 Source Management

DataTV maintains a list of currently available sources in the source window (Figure 2). Using this widget, sources can be created, toggled on and off, and deleted in both offline and live modes. Sources are created by selecting the source type and then associating the source with the appropriate object, such as a specific browser window on the desktop. Supported sources include the following:

- **Window capture**: Video output from a specific window on the desktop based on the window title, class, or executable.
- **Video capture**: Video output from an external device, such as a webcam or video camera.
- **Audio capture**: Audio output from a microphone.
- **Media objects**: Sources based on static images, video files, and rich text.

4.3 Scene Management

A scene in DataTV is an empty canvas containing sources that form the current output. Only one scene can be active at a time in DataTV; the active scene is displayed in the main composition window (Figure 2) and governs what is recorded and streamed when in live mode. The user can switch scenes using the scene list.

Managing a scene essentially entails managing the sources involved in the scene. In offline mode, most scene management operations are “heavyweight” in that they require significant setup that is not amenable to live recording. Examples of such operations include adding sources to the scene, managing their depth order (governing the drawing order of the sources), and transforming them (translating, scaling, rotating, and cropping). These operations are achieved by interacting with the sources in the composition window.

In live mode, users should mostly use “lightweight” operations that are designed for easy interaction while recording. This includes toggling the visibility of sources, zooming and panning in a source using mouse dragging and the mouse wheel (such as to zoom in on a particular part of a visualization or window), and scribbling using the annotation feature.

4.4 Video Annotation

Sportscasters regularly use annotation to scribble symbols and highlights directly on the video feed, such as to explain specific events in an instant reply of a touchdown or goal. Presentation software such as Microsoft PowerPoint supports similar “ink annotations” where the presenter can draw pen and highlight strokes directly on a slide to illustrate a specific point. DataTV supports a similar video annotation feature through its live mode interface (Figure 3), which provides a pen, highlighter, and eraser tool. The interface also allows the user to select the drawing color as well as to clear all of the annotation from the output when moving to a new scene.
4.5 Recording and Live Streaming

Our DataTV prototype supports recording to standard native video file formats (FLV, MP4, MOV, etc) as well as live streaming output to services such as YouTube and Twitch.tv. The tool also provides full control over stream settings for both video and audio output.

4.6 Implementation

We implemented our DataTV prototype based on OBS (Open Broadcaster Software) Studio, an Open Source live streaming package for multiple platforms. Our extensions were made in C++ and significantly modifies the workflow of the tool to include a streamlined live mode interface, including scene management, zooming and panning, and live video annotation.

5 DataTV Examples

To validate the DataTV prototype and to give a concrete idea of what it is capable of, we present a few examples based on real-life data and existing visualization systems. The inspiration for these DataTV examples comes from creating and understanding web-based data visualizations and infographics we experience on a daily basis. Our work here reformulates and reimagines some of the insights from the data video creation process by Amini et al. [1, 2].

5.1 Nobel Prize Data Analysis with Keshif

The first example uses the Keshif [32] system to explore the “Nobel Prize Winners” dataset. Keshif is a visualization system designed for interacting with multi-dimensional data using a sophisticated faceted browser. The key interac-
tion in the Keshif system is linked selection, which is a generalization of brushing and linking supporting highlighting, filtering, and comparison.

Figure 4: A live DataTV session using the Keshif system for the Nobel Prize Winners dataset. The user imports an external visualization tool to display the economy of South Africa.

The Nobel Prize Winners dataset is composed of the basic profiles of all winners, including their pictures, year of winning the prize, nationality, etc. The overall story of this example is to briefly introduce South Africa, particularly its former leader and Nobel Peace Prize winner, the late Nelson Mandela. The analyst uses a YouTube video of Mandela (Figure 4) giving a speech as the introduction. Live video used in this way will make the presentation engaging and draw in the viewer. It is easily achieved in DataTV using a YouTube source and controlled live by the user; no specific off-line editing is needed. Then we turn off the video and use a live window with Keshif to lead the story from introducing the categories of Nobel Prize to their age distribution. Again, the storyteller can do this in real-time simply by interacting with the Keshif visualization in a normal web browser window, potentially narrating his findings using the webcam and microphone. The DataTV platform will capture all of these sources, compose them in real-time, and stream them to a remote server. Compared to static media types, such as infographics, the ability to use an interactive interface enables the analyst to change the topic and approach during the storytelling. For example, potentially in response to an audience question through the Twitch chat service, the analyst may decide to give a little history of the Nobel Peace Prize, as well as Nelson Mandela's term as president, culminating in him winning the Nobel Prize. We then conclude the session with a short YouTube video of him giving another speech.

5.2 Stock Market Data Analysis with TimeFork

For this example, we use TimeFork [5], an interactive visual prediction technique to support users exploring the future of time-series data. The TimeFork implementation allows an analyst to explore a multitude of potential futures for specific stocks by initiating a dialogue between the analyst and the user. Here we use TimeFork to create a narrative for tech market stocks (here Apple and Netflix).

Our user starts off the session with an introductory scene involving a YouTube video featuring Warren Buffett talking about the current stock market. Then the user switches to a scene incorporating the TimeFork tool, using it to predict the current trend of hot tech stocks Apple and Netflix by simply interacting with the tool in a web browser. The user can even switch to a window of a desktop visualization tool such as Tableau or Spotfire and include that into
Figure 5: Webcam and Keshif visualization being recorded for the Nobel Prize Winner data video. The author is recording a live video “talking head” view using a webcam and microphone source as input, which is composed into the final video output.

5.3 NY Times Comment Data Analysis with CommentIQ

The CommentIQ [22] system is designed to help community moderators manage large amount of comments associated with online articles by automatically ranking them based on criteria such as relevance, readability, personal experience, and length.

In this example, the user wants to author a streaming data video about the community response on an article from The New York Times titled “City Reacts: State of Emergency” during the 2015 racial unrest in Ferguson, Missouri following the death of Michael Brown at the hands of a white policeman. The user collects two infographics with topics on murder rate across races, and the SWAT deployment rate of different races. The online interactive visualization CommentIQ looks deeper into the comments of the article from The New York Times.

This data video leverages the advantages of each of the three types of media source: video, infographics, and visualization. First, the narrator uses a YouTube video of the Ferguson incident as the introduction. This video shows

http://www.nytimes.com/
the confrontation between the protesters and the police, providing a suitable framing to the video that emphasizes the direness of the situation. The two infographics (Figure 9) give background information by showing an overview of the guns and crimes in the area. Finally, the CommentIQ visualization allows the narrator to discover trends in how the NYT commenter community responded to the article. In all of these cases, the narrator can scribble directly on the composited video output to highlight interesting aspects of the video, such as outliers or trends.

The live streaming functionality of the DataTV platform opens up an entirely new potential for the New York Times to provide a live complement to go with their online comment system. Using the DataTV live stream, community moderators could aggregate and discuss comments in real time, for example when polling voter panels for political debates.

6 Qualitative Evaluation

In order to understand how experts and practitioners from the field of information science and HCI would use DataTV, and to identify potential advantages and challenges in using DataTV, we conducted a usability study. Because there is no comparable tool for creating live-streamed data videos in real time, we opted to not perform a comparative study, but instead to focus on the affordances and capabilities of DataTV in a qualitative evaluation.

Our intent with the evaluation was to understand how DataTV can be used in different scenarios. We thus picked two separate data representations, and designed a set of tasks for each data representation such that the tasks were of similar complexity across the two scenarios. Each participant was assigned one data representation, with the associated set of tasks. During the process we use the basic expert review method proposed by Tory and Möller [29], which includes experts evaluating a tool using pre-defined heuristics. The purpose of the study was to test the usability under the context of the works of the experts.
Figure 7: DataTV recording session for a stock market prediction story involving a YouTube clip of Warren Buffet talking about the current state of the stock market as well as the TimeFork visualization tool [5] for predicting the stock price of Apple and Netflix.

6.1 Participants

We engaged two volunteer participants—Expert 1 and 2—to evaluate our system. The participants were Ph.D. students in the field of HCI or information visualization with at least three years of experience. Both were male.

6.2 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. The built-in camera was used for recording the user’s speech with voice.

6.3 Tasks

Each participant’s task was to create data videos using DataTV and the data visualization randomly assigned to them. Participants were required to use at least one interactive visualization in their video. They were allowed to pick any other appropriate media sources on the web to support their stories. The tasks were devised so that the experts would need to explore the visualizations, select media sources, and sift through information to create a narrative. The data stories we asked them to create were inspired by Section 5:

- **Obama Budget**: A visualization illustrating the components of government budget in the year 2013.
- **U.S. Census**: Demographics for the state of Florida.

A set of tasks was prepared for each scenario, requiring the participants to explore the data visualizations in detail and to look for supporting information online, before finally creating a data video to answer the question. A sample set of tasks for one of the scenarios (Obama budget) is given below:

- Explain roughly how the total ($3.7 trillion) were allocated (using a 30-second video);
- Explain the main types of spending (using a 30-second video);
• Explain how the spending has changed since the last budget (using a 40-second video); and
• Describe how much of the budget was allocated to social security (using a 60-second video).

6.4 Procedure

Participants were shown a demonstration of DataTV and given as much time as they needed to explore and familiarize themselves with the system. We then gave them a sheet of tasks and asked them to create a data video in response to each task. The resulting data video would ideally make use of multiple media types, interactive visualization, and innovative storytelling techniques. During the creation process, which was capped at 60 minutes, the experts were allowed to ask questions about the interface. We followed a think-aloud protocol with the participants, and recorded their behavior via video and written observations.

For the purpose of this study, the participant-generated stories were recorded and saved (instead of streamed dynamically). The final participant-generated videos were limited to a duration of one minute. During the process, the participants used the DataTV tool to design, sketch, record, and produce their videos. After completing their tasks, we followed up with an interview where participants explained their process and provided feedback on the system.

6.5 Results

We collected and analyzed both the products as well as the observations and interview feedback from each expert review session.

6.5.1 Products

Each expert made several data videos, all less than one minute long. Representative data videos are attached as supplemental materials. We also captured screenshots of the ending exact workspace for each expert (Figure 11). All in all, the resulting data videos were all of good quality and suggest that the DataTV platform was instrumental in the process.
6.5.2 Observations

Overall, participants used the tool with little training. Both experts familiarized themselves with the basic operation by first making a few trial videos that were recorded and saved locally. Once satisfied with the basic mental model, they spent a considerable time (20-25 minutes) finding source materials, selecting visualizations, and writing notes. Then they spent an additional 10 minutes to create their scenes, including arranging the layout and size of the different media sources. The actual recording of each video was surprisingly quick; given the preparation, the experts were both able to record their videos in a single take and with few mistakes.

From our observations, it appeared as if the experts did not need much prompting to get familiar with the DataTV interface. The experts seemed to think the interface behaved in a logical and predictable fashion, and they quickly became proficient with very little instruction and within 10 minutes of training. Most importantly, observations and think-aloud remarks seem to indicate that the experts rapidly internalized the DataTV controls and were able to focus on the craft of data-driven storytelling. This was indicated by their utterances increasingly dealing with how to best organize and present insights rather than minutiae of the interface.

6.5.3 Interview Feedback

During the structured interview after completing the tasks, the experts were asked to give feedback on the system, including advantages and disadvantages of DataTV over existing approaches. We summarize their feedback below:

• **Positive:** Expert 1 thought that the DataTV interface was simple and straightforward, with few opportunities to make mistakes. Expert 2 remarked that compared to using multiple software platforms, our system makes the video creation process seamless as it requires less operations and vital tools such as annotation, compositing, and transformations are very accessible.

  Expert 2 also thought that the use of our system was surprisingly easy, particularly the streamlined workflow where very little preparation is necessary. The expert noted that the composite video output was helpful so as to always know what is being streamed and recorded, striking a good balance between real-time control and accuracy.

• **Negative:** Both experts remarked on the lack of video editing capabilities. We responded with the fact that such editing functionality would have precluded real-time streaming of the tool, and they both remarked that...
the compromise was acceptable. The first expert complained that the DataTV interface should provide better control over media sources in the workspace.

7 Discussion

Our informal evaluation indicated that DataTV facilitated live data-driven storytelling. However, the main contribution of this paper is the method of live-streaming data videos, whereas our implementation is merely a prototype to show the validity of the concept.

It is important to note that all of the functionality of the DataTV tool can be replicated in a combination of desktop recording tools—such as VLC—and video editing tools—such as Adobe Premiere Pro—with sufficient time and effort. While DataTV makes constructing data videos easy with its integration of media source picking functionality, media label editing, and video recording, each of the DataTV videos showcased in this paper can be built using other tools. However, the argument for the DataTV platform and related software is two-fold: (1) a single unified platform is needed to allow for live streaming and rapid production, and (2) the integration of all of these data-driven storytelling features in a single tool enables the analyst to think about data videos more in terms of storytelling rather than low-level software, mechanics, and features. Results from our qualitative evaluation support these two arguments.

Having said all this, we believe our work surfaces several new issues that were not considered in the past. For example, while Amini et al. already suggested the data videos concept in 2015 [1], their work still results in a static and prerecorded video, not a live-streamed one. One of the benefits may be that it is easier to quickly produce a data video using DataTV than painstakingly using a suite of tools such as screen recorders, video editors, and audio production tools. However, DataClips [2], presented in 2017, does provide functionality for quickly assembling several clips using predefined visualizations. On the other hand, a live data video can be responsive to an audience, for example in responding to questions or requests for more information. In this way, DataTV is much more of an interactive presentation tool than typical data video production tools (such as DataClips). This is reinforced by the emphasis on live video in DataTV, whereas the narrator is typically disembodied in most existing data videos. We think this suggests that live data videos as those supported in our work is a unique data-driven storytelling medium in its own right. It is also the reason that we found no easy baseline for a comparative evaluation. We leave comparisons
Figure 11: Expert 1 reviewed the U.S. budget trend in 2013. The expert presented their insights using DataTV.

to live presentation software, such as Microsoft PowerPoint, to future work.

Our work in this paper has several limitations. First of all, much of our argumentation of using data video for storytelling is based on two assumptions: that the audience has (a) enough knowledge for understanding the data video, and (b) a favorable opinion about video storytelling. Without enough knowledge, much of the benefit of an established common ground in the visual language of data video is lost. Furthermore, given the sometimes the higher requirement of environment—playing video might be inappropriate in some communication situations—data video might not be the perfect choice for information distribution under situation that noise level is sensitive or displaying device is not well equipped. For example, it can be argued that a DataTV may not be the best vehicle for presentations in very noisy settings, such as a couple people discussing a topic in a train station, where static material might be more suitable. It should also be noted that the content of DataTV, often including personal webcam video, can be inappropriate for public broadcasting.

Finally, the utility of live-streaming data videos as a concept can be questioned. It is certainly true that we do not foresee “Let’s Analyze” videos to dethrone the “Let’s Play” category on Twitch or YouTube anytime soon. However, the power of the internet as both a medium as well as an audience should not be underestimated. There is already a small but growing group of Twitch communities devoted to non-gaming, such as painting, gardening, and programming. The step is not too far from such topics to data analysis. Besides, even if live data videos never become popular, many of the real-time authoring techniques pioneered in DataTV will be invaluable for creating normal, non-streaming data videos, going beyond what even tools such as DataClips can do.

8 Conclusion and Future Work

We have presented a new approach to streaming data exploration in real-time. We have also implemented a prototype desktop application (DataTV) that includes a source picker for managing media sources, a live navigation and annotation toolbar, a composite space for aggregating components, and controls for controlling video streaming and playback. Results from an expert review show that DataTV encourages rapid visual storytelling and helps storytellers to consider narrative rather than video production mechanics.

Our future work will continue exploring other ways to improve narrative data videos based on DataTV. We are
Figure 12: Expert 2 reviewed the population change of Florida. Insights on housing and population gain were presented using DataTV.

also interested in exploring other storytelling media such as sketches, cartoons, technical drawings, infographics, and oral storytelling.

References

[1] F. Amini, N. H. Riche, B. Lee, C. Hurter, and P. Irani. Understanding data videos: Looking at narrative visualization through the cinematography lens. In Proceedings of the ACM Conference on Human Factors in Computing Systems, pp. 1459–1468, 2015. doi: 10.1145/2702123.2702431

[2] F. Amini, N. H. Riche, B. Lee, A. Monroy-Hernández, and P. Irani. Authoring data-driven videos with DataClips. IEEE Transactions on Visualization and Computer Graphics, 23(1):501–510, 2017. doi: 10.1109/TVCG.2016.2598647

[3] S. Anderson. Watching people watching people watching. New York Times Magazine, p. MM60, November 27 2011.

[4] B. Bach, N. Kerracher, K. W. Hall, S. Carpendale, J. Kennedy, and N. H. Riche. Telling stories about dynamic networks with Graph Comics. In Proceedings of the ACM Conference on Human Factors in Computing Systems, pp. 3670–3682, 2016. doi: 10.1145/2858036.2858387

[5] S. K. Badam, J. Zhao, N. Elmqvist, and D. S. Ebert. TimeFork: Interactive prediction of time series. In Proceedings of the ACM Conference on Human Factors in Computing Systems, pp. 5409–5420, 2016. doi: 10.1145/2858036.2858150

[6] M. Bostock, V. Ogievetsky, and J. Heer. D3: Data-driven documents. IEEE Transactions on Visualization and Computer Graphics, 17(12):2301–2309, 2011. doi: 10.1109/TVCG.2011.185

[7] J. Boy, R. A. Rensink, E. Bertini, and J.-D. Fekete. A principled way of assessing visualization literacy. IEEE Transactions on Visualization and Computer Graphics, 20(12):1963–1972, 2014. doi: 10.1109/TVCG.2014.2346984
[8] F. Chevalier, N. H. Riche, C. Plaisant, A. Chalbi, and C. Hurter. Animations 25 years later: New roles and opportunities. In Proceedings of the ACM Conference on Advanced Visual Interfaces, pp. 280–287. ACM, 2016. doi: 10.1145/2909132.2909255

[9] P.-Y. Chi, S. Ahn, A. Ren, M. Dontcheva, W. Li, and B. Hartmann. MixT: automatic generation of step-by-step mixed media tutorials. In Proceedings of the ACM Symposium on User Interface Software and Technology, pp. 93–102. ACM, 2012. doi: 10.1145/2380116.2380130

[10] N. Diakopoulos, J. DiMicco, J. Hullman, K. Karahalios, and A. Perer. Telling stories with data: The next chapter—a VisWeek 2011 workshop, 2011.

[11] J. DiMicco, M. McKeon, and K. Karahalios. Telling stories with data—a VisWeek 2010 workshop, 2010.

[12] W. Eisner. Graphic Storytelling and Visual Narrative. W. W. Norton & Company, 2008.

[13] N. D. Gershon and W. Page. What storytelling can do for information visualization. Communications of the ACM, 44(8):31–37, 2001. doi: 10.1145/581641.381653

[14] M. Griffith and Z. Papacharissi. Looking for you: An analysis of video blogs. First Monday, 15(1), Jan. 2010.

[15] T. Grossman and G. W. Fitzmaurice. ToolClips: an investigation of contextual video assistance for functionality understanding. In Proceedings of the ACM Conference on Human Factors in Computing Systems, pp. 1515–1524. ACM, 2010. doi: 10.1145/1753326.1753552

[16] J. Hullman and N. Diakopoulos. Visualization rhetoric: Framing effects in narrative visualization. IEEE Transactions on Visualization and Computer Graphics, 17(12):2231–2240, 2011. doi: 10.1109/TVCG.2011.255

[17] J. Jin and P. A. Szekely. Interactive querying of temporal data using a comic strip metaphor. In Proceedings of the IEEE Symposium on Visual Analytics Science and Technology, pp. 163–170, 2010. doi: 10.1109/VAST.2010.5652890

[18] R. Kosara. Story points in Tableau Software. Keynote at Tableau Customer Conference, Sept. 2013.

[19] R. Kosara and J. D. Mackinlay. Storytelling: The next step for visualization. IEEE Computer, 46(5):44–50, 2013. doi: 10.1109/MC.2013.36

[20] B. Lee, R. H. Kazi, and G. Smith. SketchStory: Telling more engaging stories with data through freeform sketching. IEEE Transactions on Visualization and Computer Graphics, 19(12):2416–2425, 2013. doi: 10.1109/TVCG.2013.191

[21] C. Nguyen and F. Liu. Making software tutorial video responsive. In Proceedings of the ACM Conference on Human Factors in Computing Systems, pp. 1565–1568. ACM, 2015. doi: 10.1145/2702123.2702209

[22] D. G. Park, S. Singh, N. Diakopoulos, and N. Elmqvist. Supporting comment moderators in identifying high quality online news comments. In Proceedings of the ACM Conference on Human Factors in Computing Systems, pp. 1114–1125, 2016. doi: 10.1145/2858036.285839

[23] S. Pongnumkul, M. Dontcheva, W. Li, J. Wang, L. D. Bourdev, S. Avidan, and M. F. Cohen. Pause-and-play: automatically linking screencast video tutorials with applications. In Proceedings of the ACM Symposium on User Interface Software and Technology, pp. 135–144. ACM, 2011. doi: 10.1145/2047196.2047213

[24] Z. Pousman, J. T. Stasko, and M. Mateas. Casual information visualization: Depictions of data in everyday life. IEEE Transactions on Visualization and Computer Graphics, 13(6):1145–1152, 2007. doi: 10.1109/TVCG.2007.70541

[25] R. C. Schank and R. P. Abelson. Knowledge and memory: The real story. In J. Robert S. Wyer, ed., Knowledge and Memory: The Real Story, pp. 1–85. Lawrence Erlbaum Associates, Hillsdale, NJ, 1995.

[26] E. Segel and J. Heer. Narrative visualization: Telling stories with data. IEEE Transactions on Visualization and Computer Graphics, 16(6):1139–1148, 2010. doi: 10.1109/TVCG.2010.179

[27] D. Sless. Learning and Visual Communication. Wiley, 1981.

[28] J. J. Thomas and K. A. Cook, eds. Illuminating the Path: The Research and Development Agenda for Visual Analytics. IEEE Computer Society, 2005.

[29] M. Tory and T. Möller. Evaluating visualizations: Do expert reviews work? IEEE Computer Graphics and Applications, 25(5):8–11, 2005. doi: 10.1109/MCG.2005.102

[30] F. Viégas and M. Wattenberg. Communication-minded visualization: A call to action. IBM Systems Journal, 45(4):801–812, 2006. doi: 10.1147/sj.454.0801

[31] D. Wadeson. GamerTube: PewDiePie and the YouTube Commentary Revolution. Polygon, 1, September 2013.

[32] M. A. Yalcin, N. Elmqvist, and B. B. Bederson. Keshif: Rapid and expressive tabular data exploration for novices. IEEE Transactions on Visualization and Computer Graphics, 2018. doi: 10.1109/TVCG.2017.2723393