

# Roslingifier: Semi-Automated Storytelling for Animated Scatterplots

ANONYMOUS AUTHOR(S), Undisclosed Institutions



Fig. 1. ROSLINGIFIER automatically generates animated data stories from temporally changing data using animated bubbles, a natural language narrative, visual effects, and temporal branching techniques. The technique is inspired by its namesake, Professor Hans Rosling (pictured in inset; not in the system), who was famous for his flamboyant and animated (no pun intended) data presentations.

We present Roslingifier, a data-driven storytelling method for animated scatterplots. Like its namesake, Hans Rosling (1948–2017), a professor of public health and a spellbinding public speaker, Roslingifier turns a sequence of entities changing over time—such as countries and continents with their demographic data—into an engaging narrative telling the story of the data. From an in-depth analysis of public speakers using data visualization, we derive three specific techniques to achieve this: natural language narratives, visual effects that indicate the entities, and temporal branching that changes playback time of the animation. Our implementation of the Roslingifier method is capable of analyzing any given temporal dataset, identifying and clustering significant regions, automatically generating visual highlighting and a narrative for playback, and enabling the user to customize. In two user studies and an interview with three domain experts, we show that Roslingifier allows users to effectively create engaging data stories.

CCS Concepts: • **Human-centered computing** → **Visualization**; **Empirical studies in visualization**; **Visualization design and evaluation methods**.

Additional Key Words and Phrases: Data-driven storytelling, narrative visualization, Hans Rosling, Gapminder, Trendalyzer.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2021 Association for Computing Machinery.

Manuscript submitted to ACM

**ACM Reference Format:**

Anonymous Author(s). 2021. Roslingifier: Semi-Automated Storytelling for Animated Scatterplots. In *CHI '21: Proceedings of the ACM Conference on Human Factors in Computing Systems, May 8–13, 2021, Yokohama, Japan*. ACM, New York, NY, USA, 33 pages. <https://doi.org/xxxx>

**1 INTRODUCTION**

Standing today at close to 14 million views, Hans Rosling’s TED 2006 talk “*Debunking myths about the ‘third world’*” [41] is perhaps the single most significant promotion of data visualization from the early aughts of the century. Rosling (1948–2017), a professor of public health at the Karolinska Institute in Stockholm, Sweden, would follow this talk up with several other important ones (ten at TED, the most for any single person), and would also go on to co-found the Gapminder Foundation—acquired by Google in 2007—as well as co-author the book *Factfulness* [42] before his untimely death of pancreatic cancer. The common theme for all of his talks and writings was a heavy reliance on data presented using interactive visualization; in the original TED 2006 talk, he used an animated scatterplot to show the progression of various country demographics over time, handily demonstrating how our biases about the world were false. In fact, one of the reasons for Google’s acquisition of Gapminder was to make TRENDALYZER[14]—the underlying software powering these visualizations—as well as the accompanying data, freely available for anyone to use. Google even released a “motion chart” interface gadget in 2008 for inclusion into websites in the hope of broadening its use.

However, while Rosling’s talks are invariably informative and entertaining, anyone who has used an animated scatterplot *à la* Trendalyzer can attest that the experience is hardly the same. In the original TED 2006 talk, Rosling shows the trajectories of the world’s almost 200 countries jumping around on a big screen as the years advanced from early 1900s to present day. Despite the confusion and complexity of so many moving parts, he manages to frame the animation into a coherent and understandable story using three specific techniques that a standard animated scatterplot lacks: by maintaining a running spoken *narrative*, by *highlighting* important events using his hands and (in later videos) a pointer, and by *branching* time through pausing and rewinding the animation. In fact, Rosling’s narrative has been likened to that of a sports commentator given its pitch and speed as well as his ready use of instant replay, and his body language is flamboyant throughout the animation as he points out significant events during the sequence.

In this paper, we propose a data-driven storytelling method designed to augment an animated scatterplot using several methods inspired by Hans Rosling and several other effective data-driven public speakers: *natural language narratives*, *visual effects*, and *temporal branching*. We call this form of data-driven storytelling a *data presentation*: the use of data visualization as the primary driver by an in-person speaker, making it a different category from data videos [1, 44]. While our system is entirely software based, and thus lacks a humanoid robot or agent-based interface that would enable mimicking Rosling’s inimitable in-person delivery (the inset of Rosling in Figure 1 is just an illustration), we nevertheless name the method and its implementation the ROSLINGIFIER, in the spirit of Perin et al.’s Bertifier [37]. In fact, we base our work on in-depth observational coding of several data presentations by recognized speakers, including Hans Rosling himself, yielding a taxonomy of storytelling methods in data presentations.

The textual narratives we produce are presented as subtitles during the animation, and are coupled with *visual effects* and *temporal branching* to direct the viewer’s attention. These narratives, effects, and branching are automatically generated using an algorithm that detects significant trends and events in time-series data. Finally, we provide an interactive authoring interface for a designer to customize and polish a data presentation before presenting it to a stakeholder audience.

The contributions of our paper are the following: (1) a design space of storytelling methods employed by effective public speakers during data presentations derived from a careful survey and coding of many such presentations; (2) a method for automatically generating natural language, temporal branching, and visual effects of significant events in continuous time-series data; (3) a visualization system that implements this method; (4) results from a user study evaluating their utility for creators building data presentations; (5) results from a user study evaluating the utility of these ideas for consumers viewing time-series data; and (6) results from an interview with domain experts evaluating the utility of the system in the field. In the following, we first discuss the related work, and then propose a design space for visual data presentations. We present our implementation and its validation. We close with our conclusion and plans for future work.

## 2 BACKGROUND

Here we review the literature on narrative visualization, including specific methods, techniques, and authoring systems.

### 2.1 Storytelling and Narrative Visualization

As web-based visualization technologies (e.g., D3.js) have become mature and popular over the past decade, mainstream news outlets (e.g., *New York Times*, *Guardian*, *FiveThirtyEight*) have started using visualizations for telling data-driven stories. At the same time, information visualization researchers have found a new opportunity of visualizations as a new medium for communication, proposing a new line of research called “data-driven storytelling” [28, 31, 39]. The following describes theories, frameworks, and techniques for data-driven storytelling.

There have been many research efforts to survey and categorize existing data-driven storytelling techniques. Segel and Heer review 58 visualizations for storytelling and provide a design space of *narrative visualization* that enables both explorative and communicative aspects of visualization. Their design space consists of three dimensions: (i) genres, (ii) visual narratives, and (iii) narrative structures. Hullman and Diakopoulos [19] distinguish visual rhetorics by reviewing 51 narrative visualizations and discuss the effect of applying the rhetorics to four editorial layers in narrative visualization, which are data, visual representation, annotations, and interactivity, respectively. Hullman et al. [21] analyzes narrative sequencing of 45 narrative visualizations, arguing that narrative sequencing is of important factor that affects comprehension and memory. Stolper et al. [45] provide a survey of 45 recent narrative visualization examples. There are many factors that shape the visual narrative flow with on combinations of user input methods, story components, and visual feedback [35]. McKenna et al. [35] investigate how different visual narrative flows impact viewers’ reading experience. Our tool supports efficient story production, allowing narrative sequence directions [21] and highlight techniques with automated event detection methods in the animation genre [44].

### 2.2 Novel Storytelling Media

New storytelling and communication media are beginning to be used for narrative visualization. Data videos are motion graphics that combine pictographic representations and animation techniques into narrative visualization [44]. Amini et al. [1] systematically analyze what elements (e.g., narrative structures) constitute data videos from 50 data videos and discuss data video production approaches, such as strategies for engaging viewers. They later proposed DataClips [2], which is an interactive system for authoring data videos incorporating visualization.

Data comics is an emerging new communication medium and a genre of storytelling [44] that combines aspects of comics and narrative visualization [3, 4, 51]. Bach et al. [3] propose a set of design patterns (e.g., layout) for data comics that can inform design of data comics. Due to lack of effective examples of data comics, the patterns are limited and

need further evaluation with a user study. While relatively new, data comics has the potential to attract the public. To explore the potential of data comics, Wang et al. [49] conduct a study that compares effectiveness and engagement of data comics and infographics (illustrated texts), a wide-spread storytelling medium. Their experiment results indicate that data comics are more fun and engaging and better capture viewers' attention compared to infographics.

### 2.3 Authoring Tools for Generating Data Stories

There are many considerations to make effective storytelling ranging from highlighting for capturing viewers' attention, to design of story structures, interactions and transitions, and supplying appropriate explanations [4, 31, 35, 44]. As such, there have been many visual tools that allow efficient design of stories and narrative visualization. We see three types of approaches in the existing tool for authoring stories. The first type of authoring tools are those that help users to easily or automatically add visual components of storytelling to existing visualizations, such as labels [6, 13, 15, 20, 38, 43]. For example, Ren et al. [38] derive design space of annotations (e.g., shapes) and present ChartAccent, which allows interactive annotations on visualization.

To reduce the burden of manual creation of annotations, Hullman et al. [20] propose Contextifier, which automatically selects features and produces annotations with the features for stock visualizations. Gao et al. [15] showcase NewsViews which provides an automated pipeline for production of custom geovisualization for news. Similar to these tools, our tool also provides automated event detection and caption generation to reduce the burden of manual story curation.

In general, the process for visual analysis and for story creation is separated, so full stories are created after story pieces (e.g., insights, facts) are derived from visual analysis [31]. The second type of visual tools for storytelling are those that help users seamlessly connect the separated tasks [5, 8, 16] by allowing users convert the analysis results into story pieces for presentation [27]. For example, Gratzl et al. [16]'s CLUE system uses the user's visual exploration history to extract and present analysis steps and annotations.

Finally, several visual tools have been proposed recently for creating stories for specific genres or input types [9, 12, 24, 30, 33, 47, 48]. For example, infographics are a popular medium for storytelling with visual elements around text messages. Designing such visual elements often involves difficult tasks in generating, repeating, and editing stages. There are visual tools proposed to help designers in each stage. For example, Kim et al. [22] and Wang et al. [48] propose a technique and system to guide users for easy creation of graphical elements with data. Methods for automating infographics design processes is another popular research topic [9, 12, 47]. Examples include Text-to-Viz [12] for producing infographics design based on natural language statements, and DataShot [47] for creating fact sheets based on tabular data. VisJockey allows users to play animated visualizations coupled with text segments as they read them [29]. Chen et al. [9] propose a deep-learning based automation approach which extracts components of existing timeline infographics for creating improved designs. There also exist additional tools for storytelling media that have gained popularity, such as data videos [2, 10, 34], slideshows [46], and data comics [23, 51].

Most closely related to our work is SketchStory [30], which supports not just off-line authoring of data stories, but also has a pen-based presentation mode. Compared to this all prior work, our work provides a semi-automatic data generation pipeline and is focused on time-series data visualized using animation. Our system generates not just captions, but also textual narratives to augment the animations.

## 3 DESIGN SPACE: VISUAL DATA PRESENTATIONS

Here we present a design space for visual data presentations, including definitions, examples, and narrative mechanisms.

### 3.1 Definition

The goal of this paper is to study *visual data presentations*, which we define as the use of interactive visualization to support in-person presentations. Visual data presentations is a genre of data-driven storytelling [44].

### 3.2 Comparison: Data Videos vs. Data Presentations

The genre of data-driven storytelling most similar to data presentations is what Amini et al. call *data videos* [1], or what Segel and Heer recognized as the *film/video/animation* genre of their seven genres of narrative visualization [44]. We note that while data presentations are similar to data videos, and may in fact often be recorded on video (which is how many of us get to see them), the significant difference is the presence of a in-person speaker—rather than a disembodied narrator—using the visualization as visual aid to convey a message.

Prolific examples of data presentations include Hans Rosling’s 2006 talk on common misconceptions around the so-called “Third World” [41], as well as Al Gore’s presentation on CO<sup>2</sup> emissions in the movie *An Inconvenient Truth*, where the former U.S. Vice President famously used a scissor lift to demonstrate the catastrophic increase in greenhouse gas emissions in the last few decades.

Contrast the above data presentations with videos where either subtitles or a disembodied voice—Hans Rosling has in fact recorded a few of these—narrates animated visualizations, such as in the online documentary *The Fallen of World War II* [17] or in *A Day in the Life of Mister O.* [25], an abstract short film about humanity’s environmental impact on the world’s oceans. The distinction is clear: in a data presentation, the speakers themselves can play a significant role in providing not just an engaging spoken narrative, but can also interact with the visualization by pointing to specific parts, highlighting important trends, and even control the visualization, such as by playing back an animation.

Sometimes the boundary between a data video and a data presentation can be blurred. For example, in the 2019 online political movie *Unbreaking America: Solving the Corruption Crisis* [T2], which very much looks like a data video, actor Jennifer Lawrence actually appears in the video with not just her voice, but also her likeness. As a result, she is able to point to, describe, and explain graph axes, data items, and insights in significant detail. Since Mrs. Lawrence is actually embedded into same space as the visualizations themselves, we tend to think of this more as an example of a data presentation than a data video.

### 3.3 Method

We started the process to derive and explore the design space of visual data presentations by reviewing existing data presentations online. More specifically, we looked for videos presenting data-driven stories using interactive visualization where the speakers—their likeness and not just their voices—are part of the video. To seed our search, we started from the following initial categories:

- **TED talks:** The TED—Technology, Entertainment, and Design—conference and its satellite events include many data-driven visual presentations, including ten of Hans Rosling’s own talks.
- **News and weather reports:** News anchors sometimes use interactive visualization to describe complex events. In particular, weather forecasts are spatiotemporal data stories conveyed by a meteorologist narrating and pointing to specific areas of interest.
- **Data-driven organizations:** Certain organizations, notably the Gapminder Foundation, publish data presentations as part of their mission, and thus constitute a rich source of inspiration.

	Technique	Description
Gestures	Pointing	Indicate an entity with hands to draw the audience’s attention towards an entity.
	Tracking	Pointing and moving at an animated entity to explain changing data and emphasize trends.
	Shaping	Express the shape of data using hands to emphasize spread, ranges, and boundaries.
Visual effects	Labeling	Naming an entity or a group with a word or a phrase.
	Spotlighting	Temporarily change appearance of an entity to emphasize or draw attention to a certain point.
	Tracing	Draw paths of animated entities to emphasize trends or compare different movements.
	Accumulation	Add items to an existing visualization to emphasize change over time.
Animation playback	Pause	Stop the animation for a short time to draw attention to an event at a specific point in time.
	Slowdown	Decrease the animation speed to explain a detailed event sequence in a short time period.
	Speedup	Increase the animation speed to explain trend in a longer period, or skip where little happens.
	Rewind	Play by moving back a few frames to repeat to show the change or to deliver different message.
	Replay	Repeat the entire animation to summarize presentation.

Table 1. Classifying storytelling techniques and intentions in data presentations.

Our search was by no means exhaustive, as we were interested in finding a representative rather than a comprehensive sample. For example, there are thousands of relevant data presentations on Youtube that would fit our general visual data presentation definition above. Thus, we did not endeavor to cast our net too widely, but rather selected a smaller set of videos that fulfilled the following criteria: (1) communicates information about data; (2) includes an interactive/animated visualization; and (3) combines the speaker’s body with the visualization. Furthermore, we curated our selection to capture varied examples.

After having selected our data presentations, we used open-book coding to understand narrative actions employed in the videos. In an initial pass, two of the authors independently coded all of the videos into actions performed by presenters in the videos, including the perceived intentions of each action. We then compared each coder’s results and discussed inconsistent coding before resolving the final version. When there was a coding disagreement, a third author arbitrated the conflict.

### 3.4 Data Presentations Surveyed

We selected 11 data presentations for detailed review based on number of views, quality, and diversity from a larger set of data presentations; see Table 2 for an overview. Since Hans Rosling was a very effective and engaging speaker who often gave data presentations, six of the videos we analyze are his (R1-R6). We also selected 5 other popular data presentation videos to include different stories, visualizations, and environments. R1-R3 are Rosling’s presentation in major news media, where he presented himself on a hologram-style display. R4-R6 are TED talks, where he provides a live demo with a large screen by controlling the system on the stage. T1 shows Al Gore’s CO<sup>2</sup> emission chart in the movie titled *An Inconvenient Truth*. T2 is a political video from an organization called *Represent Us*. N1 and W1 are news broadcasts using interactive visualizations—the EU referendum result and a weather forecast from the BBC. M1 is the stock market analysis from the CNBC.

All of Rosling’s videos (R1–R6) use scatterplots, but often include other charts (e.g., line chart and map in R5 and R6). We find that a line chart is used to show the CO<sup>2</sup> level in T1 and various charts on a map are utilized in T2, N1, and W1, to represent spatiotemporal data.

		Pointing	Tracking	Shaping	Labeling	Spotlighting	Tracing	Accumulation	Pause	Slowdown	Speedup	Rewind	Replay
Talks	[T1]	✓	✓	✓	✓	✓	✓						
	[T2]	✓	✓	✓	✓	✓	✓						✓
News room	[N1]	✓	✓		✓	✓							✓
	[N2]	✓	✓		✓								
Weather report	[W1]	✓	✓		✓	✓		✓					✓
	[W2]	✓	✓		✓	✓	✓	✓		✓		✓	
Market report	[M1]	✓	✓	✓	✓	✓							
	[M2]	✓	✓		✓	✓							
Data presentations	[R1]	✓	✓	✓	✓			✓	✓	✓		✓	✓
	[R2]	✓	✓	✓	✓	✓	✓	✓				✓	✓
	[R6]	✓	✓		✓	✓	✓					✓	✓
	[R7]	✓	✓		✓	✓		✓	✓	✓		✓	✓
		Gestures			Visual effects			Animation playback					

Table 2. Storytelling techniques in data presentation videos.

### 3.5 Narrative Techniques and Intentions

Table 1 shows our derived narrative actions in data presentations, including their intentions. We group these techniques into three categories:

- **Gestures:** The speaker performing gestures using their hands, or sometimes using a handheld pointer, to indicate items of interest.
- **Visual effects:** Graphic effects in the visualization or added in video post-production to highlight entities, trends, or insights.
- **Animation playback:** Changing the speed, position, or direction of the data animation to showcase specific phenomena.

Note that our summary above does not include the ubiquitous narrative tool common to all data presentations: the use of a *verbal narrative* to direct the viewer’s attention, explain specific phenomena, or convey a message. Since our focus here is not specifically on the use of verbal techniques, and since our implementation uses written language and not actual speech, we choose not to delve deeper into this aspect of data presentations. However, it is clear that understanding the verbal delivery of data-driven narrative involves both the *content* of the message—indicators of space, identity, magnitude, effect, causality, etc—as well as its *mode*—speed, pitch, inflection, etc. We leave such expansion of our design space for future work.

### 3.6 In-depth Analysis of Narrative Techniques

Here we present an in-depth analysis of narrative techniques used in our corpus of 11 data presentation videos. Figure 2 shows the temporal event sequence of the selected videos (additional analysis can be found in Appendix B). The color-coded sections indicate the categories of techniques—gestures, visual effects, and animation playback.

*Gestures* are used to indicate or point to entities using hands or handheld pointers, and splits into three sub-categories—pointing, tracking, and shaping. *Pointing* draws the audience’s attention, and is the most common technique throughout the 11 videos. For example, pointing is commonly used to explain the basic features on a chart, such as axes, legends, and

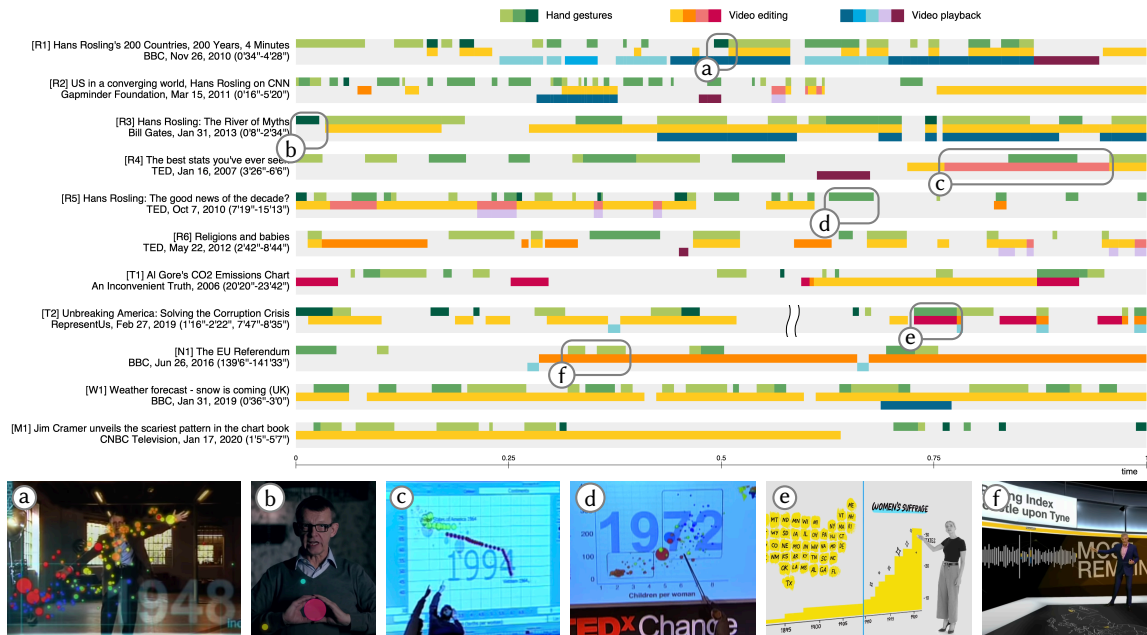


Fig. 2. Analysis of storytelling techniques in 11 data presentation videos. The highlighted time sequences in the various timelines are illustrated using insets at the bottom of the figure.

points. Presenters also use gestures to explain the meaning of the data items on the chart. W1 and M1 are exceptions. The weather reporter in W1 does not explicitly describe the meaning of the map or isotherms. Similarly, the analyst in M1 does not explain the stock chart. In these cases, the presenters seemingly assume that their audience is already familiar with the visual representations used.

*Tracking*, an action of continuously pointing at animated entities, is often used to indicate the trends or movement of data. We distinguish *pointing* and *tracking* as they have different usages and intentions. Tracking gestures usually come with the tracing visual effect on animated scatterplots (Figure 2c) or the accumulation technique on area or line charts (Figure 2e). Tracking gestures are also used to indicate the movement of a group (Figure 2d). *Shaping*, on the other hand, is used to express the geometric shape of data, such as the size, trends, boundaries, or a growing/shrinking movement. Rosling used the shaping technique to describe the size (Figure 2b) and movement of entities (Figure 2a). It is also used in T2 to explain the meaning of the axes and trend lines.

*Visual effects* are techniques that modify the visual appearance of a chart to emphasize entities or trends. Visual effects consist of four sub-categories—labeling, spotlighting, tracing, and accumulation. *Labeling* is used to emphasize important items in a story. Rosling often picks a few representative countries with a big population or data anomalies. He presumably wants the audience to focus more on those countries rather than be confused by the presence of many countries. The stock analyst in M1 labels the peaks and valleys of the line chart to highlight important points. *Spotlighting* emphasizes an entity by temporarily changing its appearance. For example, this could be a brightness change on a black background (R2), or blinking on a beam projector (R6, R7). N1 uses the spotlighting technique (Figure 2f) to emphasize the result of a city by flashing the corresponding point. The *tracing* technique (also discussed by Robertson et al. [40]) is often used on animated scatterplot to draw paths of entities over time. Figure 2c shows an example using tracing, where



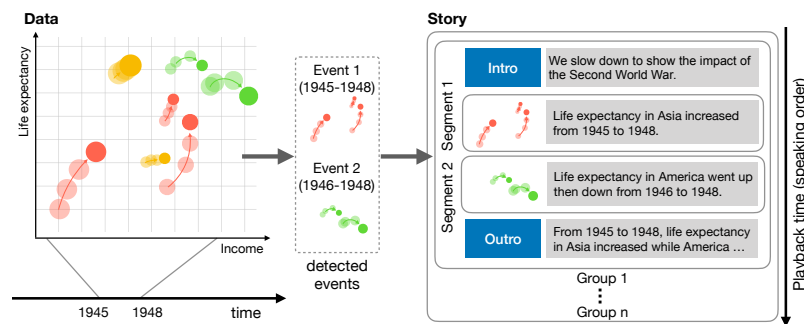


Fig. 3. Creating a story for a sequence of concurrent events. An intro and an outro introduce and summarize the segment group, respectively.

Rosling compares the temporal movement of two entities. Finally, the *accumulation* technique emphasizes temporal changes by gradually adding new items, which is similar to the wipe in movie transitions. This technique works on the area or line chart to draw the audience’s attention to the changes (Figure 2e).

*Animation playback* is employed to emphasize parts of an animation in different ways. *Pausing* is mostly used to explain the trends or reasons on an event, such as World Wars or the Spanish flu. In Figure 2 W1, the weather reporter pauses the map and emphasizes the unusual weather condition of the week. *Slowdown* sometimes replaces the pause, describing an event or a reason for the event in the period. *Speedup* techniques follow after a pause or a slowdown, quickly skipping the less important intervals (T2, N1). *Rewinding* indicates moving back to a few time frames to repeat the interval. Rewind techniques are used to emphasize a certain event or entities or to convey different stories in the same period. *Replay* also goes back to previous frames but has different intention from that of rewind: it repeats the entire animation to summarize a specific animated segment. Rosling often replays the entire animation after explaining the overall trend (R1, R2, R4, R6). The replay technique comes at the end of the presentation (R1) to summarize the talk, or comes in the middle to give the audience time to digest the story before he moves to the next stage (R2, R4, R6). Users can mix the playback techniques based on intentions in presentations. For example, Rosling uses a series of the techniques in a novel way to better present an event in the story. In R2 (Figure 2), he showcases a combination of a pause, replay, and rewind with different playback speeds, to stress how the rest of the world different countries caught up to the U.S. with respect to income and life expectancy in multiple perspectives during 1860–2010. We call this strategy *temporal branching*, where the presenter utilizes rewinds at several time points in a given period, embedding other playback or highlighting techniques in the rewinds to deliver different aspects of an event in detail.

#### 4 AUTOMATICALLY GENERATING DATA-DRIVEN STORIES

Our approach in this paper is to automatically generate a data-driven story from time-series data. The approach includes detecting events in the time-series data into causal sequences that form stories, then using natural language to generate narratives, and finally using our storytelling techniques from Section 3 to enrich these narratives.

#### 4.1 Fundamentals

A *story* is a presentation sequence, which includes *segments* of data in a linear chronological sequence. A noteworthy interval and a group of entities in data is called an *event*. The process of generating a story from a time-series dataset involves creating a segment per event and then serializing these segments into a linear sequence.

For singleton events that have no concurrent events, this is trivial. In situations where multiple events overlap partially or completely in time, we must select a linear sequence for the resulting parallel segments. This can either be done randomly, on the basis of some interest function (e.g., the magnitude of the event), or controlled by the user. However, linearizing such parallel segments results in having to play out the time for one segment, and then rewind in order to start the next segment, etc. To communicate this fact to the user, we introduce intro and outro segments at the beginning and end of each group of concurrent segments. Furthermore, during playback we must convey the time being rewound when switching to another concurrent event.

#### 4.2 Example: Deriving Stories from Time Series

Figure 3 illustrates how we can form a story from data. We use multi-dimensional time-series data that uses income on the X-axis and life expectancy on the Y-axis. Each data point represents a country. The color and size indicate the continent and population of the country.

In the period of 1945–1948, we first detect a group of countries in Asia (red) that is increasing fast in life expectancy. This interval is translated into Event 1, which becomes Segment 1 with the label “Asia.” At the same time, in the period 1946 to 1948, we identify Event 2 where countries in America (green) are also changing fast in life expectancy. As the Event 1 and 2 occurred in the overlapping period, these are linearized into a segment group. A segment group often represents a historical event with a global scope, e.g. World Wars or pandemic. In this case, Group 1 includes events occurring just after World War II. Intro and outro segments summarize what happened in the period and provide reasons for the events.

#### 4.3 Event Detection and Narrative Generation

Creators often face a difficulty in conceptualizing their work when they first see their data and task [36]. When this type of difficulty exists, recommendations can play an important role in guiding and inspiring creators [32]. Roslingifier provides the recommendations that can help creators overview the data and find events that could be used to attract the audience. Many approaches attempt to detect significant events from time-series data [18, 26, 50]. For example, Kong and Agrawala [26] decompose line charts into a set of human perceptual part—such as peaks, valleys, rising, and declining slopes—by calculating the first and second derivatives to identify curvature extrema. There are also tools based on event detection algorithms. For example, Microsoft Power BI automatically detects 17 insights from the Gapminder data on life expectancy data, which focus on showing general trends and outliers. Inspired by the work, we decide to provide event recommendations which can help creators intuitively understand unseen events in a perceptually salient way based on extrema and the size of change.

First, we define an event with a set of data dimension  $D_i$  (e.g., Income) of each legend  $g_j$  (e.g., Asia), a time interval denoted as  $t_s$  and  $t_e$ —the start and end time (e.g., 1945 and 1948)—and a movement pattern (e.g., rise) in the interval. Our event generation algorithm is based on calculating the size of the change for each  $D_i$  of legend  $g_j$  being tracked. If the change size continuously exceeds a threshold, the algorithm detects an event with the time interval,  $t_s$  and  $t_e$ . Depending on the movement of values in the interval, we define four events, RISE ↗, DROP ↘, TROUGH ⬆, and PEAK ⬆. Additionally,

the algorithm detects two more events—PLATEAU→ events represent intervals with no change, and SPREAD✂ represents an interval where the difference between the values is the largest. Finally, we have user-generated events USER🔍. Here we summarize the seven types of events that our system detects, including natural language narratives to describe events using the following rules:

- ↗ RISE: The value increases in the detected interval. *Example:* The increased life expectancy in Europe after World War II. *Narrative:*  $D$  in  $g$  increased between  $t_s$  and  $t_e$ .
- ↘ DROP: The value decreases in the detected interval. *Example:* The decreased life expectancy in the world during the Spanish flu. *Narrative:*  $D$  in  $g$  decreased between  $t_s$  and  $t_e$ .
- ⤴ TROUGH: The value decreases and then increases in the detected interval. *Example:* Life expectancy in the world dropped then recovered from the impact of the World War I. *Narrative:*  $D$  in  $g$  went down then up between  $t_s$  and  $t_e$ .
- ⤵ PEAK: The value increases and then decreases in the detected interval. *Example:* Income in Europe reached the peak in 2005 then decreased in the next year. *Narrative:*  $D$  in  $g$  went up then down between  $t_s$  and  $t_e$ .
- PLATEAU: There is no change in values over a predefined number of time frames. *Example:* Income in Asia did not change from 1800 to 1820. *Narrative:*  $D$  in  $g$  is mostly constant between  $t_s$  and  $t_e$ .
- ✂ SPREAD: The difference between the maximum and minimum values is the largest during the time period covered for a particular data dimension. *Example:* In 1948, the difference between countries was wider than ever (Figure 2a). *Narrative:* In  $t_s$ , the difference between the items was at its widest.
- 🔍 USER: User-generated event created manually. *Narrative:*  $D$  in  $g$ , something happened between  $t_s$  and  $t_e$ .

Narratives for the intro get the information of a set of event segments including the time range covering all member events, from  $T_s$  to  $T_e$ . The member events with the same event type are grouped and are summarized together. For example, from 1945 to 1948, life expectancy in Asia and America increased, income in Europe went up then down. We do not provide narratives for the outro so that users can fine-tune the narrative to make a story by providing reasons or detailed analysis of events. We describe narrative editing in detail in Sec. 5.3.

#### 4.4 Storytelling Techniques

Beyond natural language, we provide visual effects to emphasize important events, as shown in Figure 2 c, e and f. We employ spotlighting and tracing to emphasize legends for an event. When animation is played, the bubbles are colored to stand out while others are grayed out. The traces of colored bubbles are displayed to track changes.

We do not explicitly show a cursor or pointer to implement hand gestures, as data presentation may not have a human presenter. Instead, we provide two labeling features to support indication of an individual or a group of entities, and visual effects to highlight their movement. First, we automatically generate inner clusters of the legend based on the temporal proximity of the entities in the event. Clusters are labeled by summarizing the higher level of information (e.g. subcontinent of counties). We discuss details of the clustering algorithm in Appendix C. Second, we support turning on and off the labels on entities during the animation. Third, we emphasize entities' movement by drawing their traces. Sec. 5.1 describes how the system supports these features to highlight the events.

We employ the animation playback techniques to deliver the story for segment groups. We slow down the playback speed when playing segment groups and increase the speed for other intervals. The chronological sequence of the animation is distorted within a segment group. After playing the intro, the animation goes back to the starting time of the next event; this corresponds to the rewind technique. Finally, the outro is played to summarize the set of events. The implementation of the animation playback is shown in Sec. 5.3.

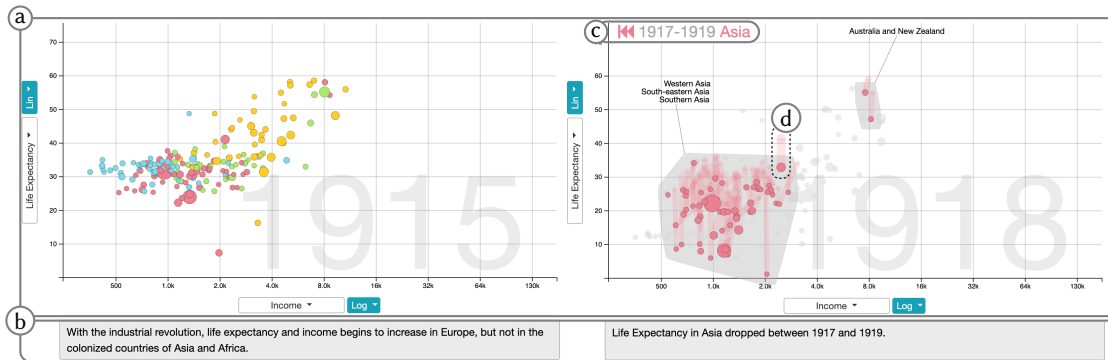


Fig. 4. The presentation output view. (Left) default mode: the chart panel (a) shows the bubble (country) in the X (income) and Y (life expectancy) coordinate. Color and size indicate its continents and population. (Right) highlighting mode: the bubbles are clustered according to their position and movement and labeled. The caption panel (b) presents natural language narratives, describing events at the current time point. The event indicator (c) indicates the period and the legend of the current event.

## 5 ROSLINGIFIER

We propose the ROSLINGIFIER system to enable users to author data presentations from an automatically generated story that they can view and edit by interactively applying various storytelling techniques. Roslingifier consists of three visual components (Figure 1): a presentation output view, an event exploration view, and a presentation editor. We use socio-economic data from Gapminder.org to demonstrate the system below.

### 5.1 Presentation Output View

The presentation output view (Figure 4) is the central component of the system, and is also used for animated presentations. This view supports changes of X and Y data and their scale (linear or logarithmic) and allow users turn on and off the labels on the entities, change the position of the label, and check if narratives are readable. The view consists of two parts: a chart panel (Figure 4a) and a caption panel (Figure 4b). We choose a scatterplot for the chart panel to present multidimensional data in temporal animation, as Rosling did in many of his videos. Bubbles in the plot represent data points on the 2D Cartesian plane where the size and color indicate additional data dimensions, respectively. Users can change the data dimensions of X and Y axes and convert between a linear and log scale. Figure 4a shows Rosling's popular scatterplot where a bubble represents a country on the income (x-axis) and life expectancy (y-axis) coordinates, and the color and size of the bubble indicate its continent and population, respectively. The caption panel (Figure 4b) presents the narrative at the current time point, aligned with the schedule in the presentation editor (see Sec. 5.3).

There are two different modes in the view: default (left) and highlighting (right) modes. The default mode is on when playing the interval outside event segments. All the bubbles in every legend are displayed without any labels unless the user hovers on them. The highlighting mode is activated when playing event groups. In this mode, the event indicator appears in the top-left corner (Figure 4c) to show current event's time period. It also shows whether the animation is going forward or rewinding for presenting an event on the highlighted legend  $g$ . For example, the event indicator in Figure 4c means it is rewinding from 1917 to 1919 to focus on a legend of Asia. For each event, we first color the bubbles where the entities of the bubbles belong to the highlighted legend. Other legends are grayed out to make the legend stand out. The traces of colored bubbles (Figure 4d) are displayed to help audience track the changes.

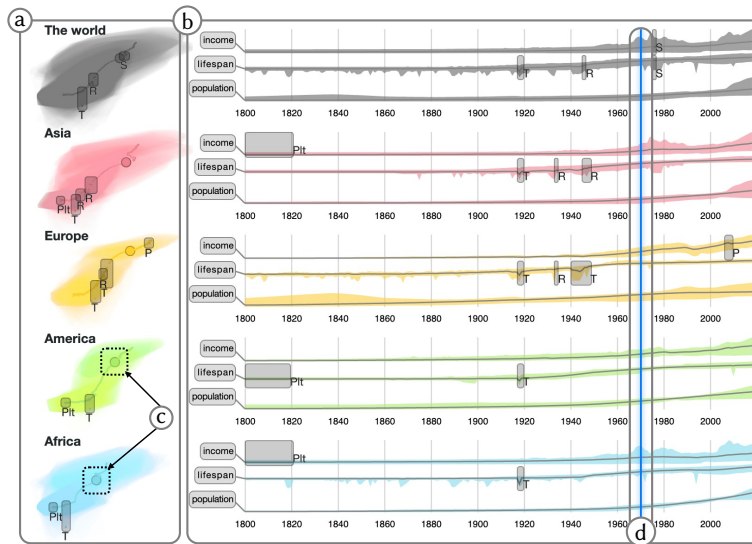


Fig. 5. The event exploration view presents the hull traces (a) to abstract overall movements, and the line charts (b) to show the progression of values. Time indicators (c, d) show the current time frame during the animation.

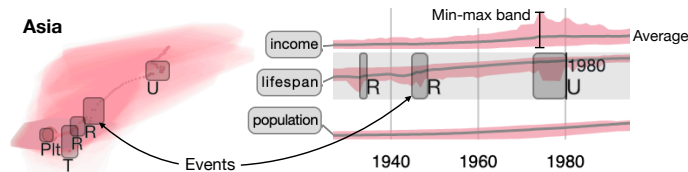


Fig. 6. Close-up on the event exploration view. Shaded rectangles show the position of events on the scatterplot and line chart.

The bubbles of the highlighted legend are clustered based on temporal proximity. We draw a convex hull to connect the bubbles in a cluster. Labels are automatically generated by summarizing the higher level information of the constituent entities (e.g., Southern Asia). We discuss the clustering algorithm in [Appendix C](#). [Figure 4](#) (right) shows an example, where bubbles for other than Asia are grayed out to present an event occurred from 1917 to 1919 in Asia. The clustering algorithm finds two clusters, one for Australia and New Zealand (top-right), and another big cluster for other than the two countries (bottom-left). The label for the bottom left cluster is generated using the subcontinents information of the countries, which are Western Asia, Southeast Asia, and Southern Asia (sorted by the number of countries). The summarized label reduces the amount of information and provides a better understanding of the bigger trends.

## 5.2 Event Exploration View

To help users explore trends and events automatically identified from data, we provide an event exploration view with two visualizations: hull traces ([Figure 5a](#)) and line charts ([Figure 5b](#)). Hull traces show the temporal distribution of bubbles for each legend. [Figure 5a](#) shows five hull traces for five legends—the world, Asia, Europe, America, and Africa—represented by different colors. A hull trace shares the same X and Y axes with the chart panel ([Figure 4a](#)). It consists of multiple convex hulls, each of which covers a group of bubbles at a time frame. The convex hulls are

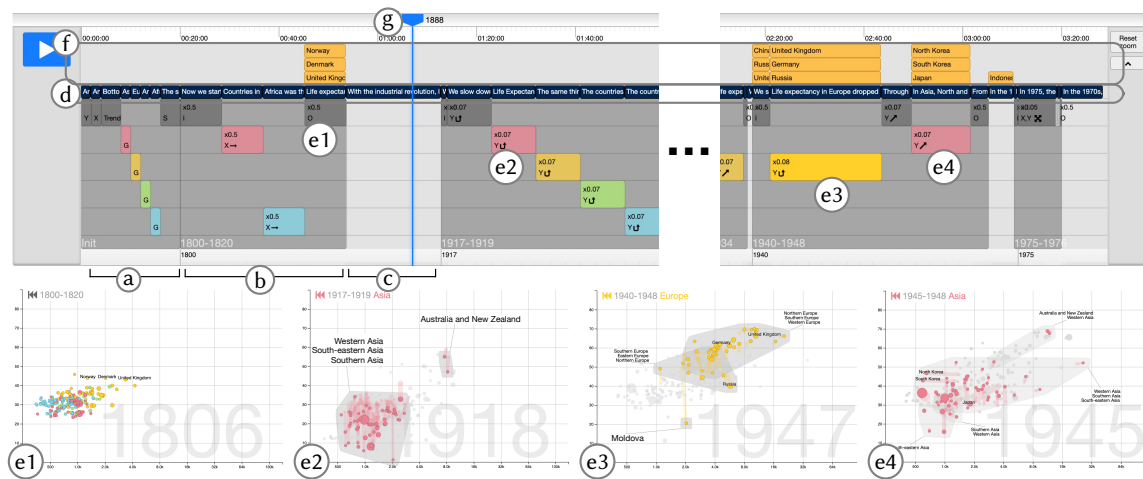


Fig. 7. The presentation editor manages the animation schedule with four types of story blocks: initial segments (a), event segments (b), blank frames (c) and narratives (d). Enabled labels are displayed in (f). The time indicator (g) shows the progression of time. e1–e4 show the chart panels for each corresponding event segment.

translucent to help users to determine how many layers are stacked. For example, in the hull trace of Asia (Figure 6 left), the bottom left corner is darker than other areas by stacking a more number of convex hulls. This implies that Asian countries stayed a long period of time on the bottom left corner. The centroids of the convex hulls illustrate the tendency of the direction that the convex hulls have moved over time, whether from the bottom left to the upper right, or the other way around. The centroid of the current time point is highlighted (Figure 5c) during the animation.

Line charts (Figure 5b) present the progression of values in X (e.g., income), Y (e.g., life expectancy), and bubble size (e.g., population) dimensions of the chart panel (Figure 4a). All line charts share the same time range for the X-axes and normalized (0–1) Y-axis. The shaded areas represent the minimum and maximum values of the data in each dimension. Figure 6 (right) shows three line charts of Asia. In the first line chart (income), the average income of the Asian countries is drawn in a line with a min-max band as a shaded area. By plotting values for the entire time range, the line charts show the changes in values and capture the trends in each data dimension. A time bar (Figure 5d, vertical blue bar) on the line charts goes along the X-axis to indicate the progress of time during the animation.

Events are shown in the event exploration view. Gray rectangles (Figure 6) on both the hull trace and the line chart indicate the period and the type for events. Event types are labeled next to the rectangles (e.g., *P* for PEAK, *Plt* and *U* for PLATEAU and USER). A rectangle on the line chart visually emphasizes the event length, while a rectangle on the hull trace shows the approximate position on the coordinate. We discuss the event detection algorithm in Sec. 4.3. Users can draw a rectangle in the line charts to create user-driven events (i.e., USER).

### 5.3 Presentation Editor

The presentation editor (Figure 7) helps users manage the animation schedule that determines which frames to run in which playback time. There are two timelines at the top and bottom of the view. The top timeline shows story blocks start, duration, and entire playback time information (e.g., “00:00:00”), while the bottom one displays the data time of events (e.g., “1800”). The top timeline shows the entire running time of the presentation with a tick being a time unit

(e.g., a second). The bottom timeline is non-linear to support animation playback techniques (Table 1). The blue line (g) shows time progression, which is linked to the time indicators on the event exploration view (Figure 5 c and d).

Three are four types of story blocks in this view—initial segments (a), event segments (b), blank frames (c), and narratives (d). The initial segments (a) are located at the beginning of the animation to explain basic components in the chart panel (Figure 4a), such as variables assigned to the axes (e.g., income) and legends (e.g., continents). Rosling used to explain the overall data trends on the chart before the presentations began, describing the meaning of entities moving from bottom left to top right corners. We follow Rosling’s order of explanation, constructing a sequence of this segment with 4 steps: we define X and Y axes, explain the corners of the coordinates based on data trends, introduce legends by displaying one color at a time, and describe the bubble size.

After the initial segments, the presentation editor by default creates blank frames (c). At this point, if users produce a result video, it chronologically animates a series of scatter plots on the default mode (Figure 4 left) from the beginning to the end of the data (e.g., from 1800 to 2018). Each snapshot is played for a unit time (e.g., 200 milliseconds in this work). If an event is detected, it replaces the blank frames in the same period. Figure 7b shows a group of segments including PLATEAU→ events in Asia and Africa. The group appears as a gray rectangle and the time range on the bottom indicates its interval, e.g., “1800–1820”. It consists of three components: the intro (I), the events, and the outro (O). Each component includes slowdown rate and data dimension information. Events are marked with a data dimension: X, Y, or S, followed by arrow icons. The intro and outro are labeled as I and O, respectively. Event segments run slower than a regular speed, and different types of events have different slowdown rates. By default, the frames of the intro and outro play 2 times slower than the unit time. We set 15 times slower frames for RISE↗, DROP↘, TROUGH⬆, and PEAK⬆ events, 20 times for SPREAD↔, 2 times for PLATEAU→, and 10 times for USER⊙ events. The presentation editor supports editing events. Users can swap the order within a group, delete events, or edit the playback time of event segments. Throughout the animation, users can (de)activate labels by clicking entities in the chart panel. The label texts are shown in Figure 7f.

## 5.4 Implementation Notes

We implement Roslingifier with Django, HTML, JavaScript, and Bootstrap. We use D3.js for the scatterplot and the line charts, and Python 3 with the pandas library for clustering. The data manipulation and the clustering is done by backend in Python 3 with the pandas library.

## 6 USE CASE

We present two use cases for using Roslingifier. First, we create a story using the relation between life expectancy and income in the Gapminder data. We showcase another story using COVID-19 outbreak data [7] to demonstrate that our tool is generalizable to other datasets.

### 6.1 Life Expectancy vs. Income

A user is asked to give a presentation on the world history to a general audience. As she knows from Rosling’s videos that there are interesting story pieces on the relation between income and life expectancy, she decides to use the data in Roslingifier. In the tool, each bubble represents a country, and the size of each bubble indicates its population. The color of bubbles indicate the continents, as shown in Figure 5(a). The data includes 184 countries from 1800 to 2018.

The story generation starts with event detection. As the user sets the threshold as 3% for detecting event intervals, Roslingifier finds events that show larger differences than the threshold. Figure 5 show the detected events for the five continents. From 1800 to 1820, Roslingifier detects 3 PLATEAU→ event segments—income in Asia, life expectancy in

America, and income in Africa. By using playback, she confirms that the countries gather together and do not move forward (Figure 7e1). Here she deletes the event on lifespan in America in the presentation editor to focus on the income dimension in her story, leaving two segments in the group (Figure 7b, red and light blue legends). The caption panel also reflects the deletion, presenting “*Income in Asia (Africa) is mostly constant between 1800 and 1820.*” To stress Africa as well, she removes “*(Africa)*” and adds “*Africa was the same.*” to the caption. Playing back the video with the captions, she thinks she can make a brief story related to the industrial revolution, when the countries in Europe move toward the upper right while other countries remain the same. In the outro, she turns on the labels of the UK and other countries in Northern Europe to emphasize the movement of those countries (Figure 7e1).

In the next segment group (Figure 7e2), she finds that all continents have a TROUGH↓, experiencing a big drop in life expectancy due to World War I and the Spanish flu epidemic and recovering during 1917–1919. She additionally sees that Roslingifier suggests to narrate these events using the rewind technique, where the chart panel shows how each continent experienced the events in detail. During a test playback, she sees many bubbles dropping and soaring simultaneously. She also notices that the clustering algorithm separates the bubbles when highlighting each continent in the chart panel. She also thinks that it is an interesting point to the audience that Australia and New Zealand form their own cluster, separated from other countries in Asia (Figure 7e2). Roslingifier emphasizes the rapid movement of the bubbles in this period with the traces of the countries. Because she believes that this segment group will engage audience, she decides to keep the group in her presentation. In fact, the story of this segment group is what Rosling also presented [R1].

Lastly, she sees in the last segment group (Figure 7e3) two Rises↗ on life expectancy of the world (1945–1946) and Asia (1945–1948), and one Trough↓ on life expectancy in Europe from 1940 to 1948. Seeing the year information on the timeline, she notices that the detected events are related to the World War II. To see what events Roslingifier finds in the time range, she plays the animation. In the animation, she observes that life expectancy in Europe (Figure 7e3) dropped significantly in 1944, and then recovered when the war ended in 1948. She also sees in the chart panel the movement of major countries—UK, Germany, and Russia. She thinks it is interesting that Moldova is located far from other European countries (Figure 7e3 bottom). While watching the events between 1945 and 1948 in Figure 7e4, she notices long traces that indicate that (1) North and South Korea experienced a huge fall in income in 1945, and (2) both life expectancy and income dropped in Japan, but soon recovered to a greater extent by 1948. Overall, she finds that the events in this group are also worth telling, so she decides to use this part for her presentation after adding labels on the countries for further emphasis.

## 6.2 COVID-19: Number of New Confirmed Cases

We use Roslingifier to create a story on the COVID-19 pandemic during the first three months of 2020. The data includes 170 countries with daily new case counts of 95 days. We map the X-axis to the number of days since 100 cases and Y-axis to new confirmed cases each day. Each bubble represents a country and has the same radius. The threshold is set to 4%.

Figure 8 presents an example segment group with three events from April 6 to April 17. The events are detected due to their rapid change in new confirmed cases: TROUGH↓ in Europe on April 6–16, TROUGH↓ in America on April 13–15, and RISE↗ in Africa on April 16–17. Figure 8 (left) shows the intro of the group with auto-generated narratives that summarizes the interval: “*From 4/6 to 4/17, new confirmed cases in Europe and America went down then up. New confirmed cases in Africa increased.*” We then select a country in each continent to highlight their traces: US, Italy, Egypt, and China. After the intro, the story rewinds to highlight the TROUGH↓ event in America from April 13 to April 15 (Figure 8 (right)). The clustering algorithm detects three clusters. Here we can add enriched data presentations to the



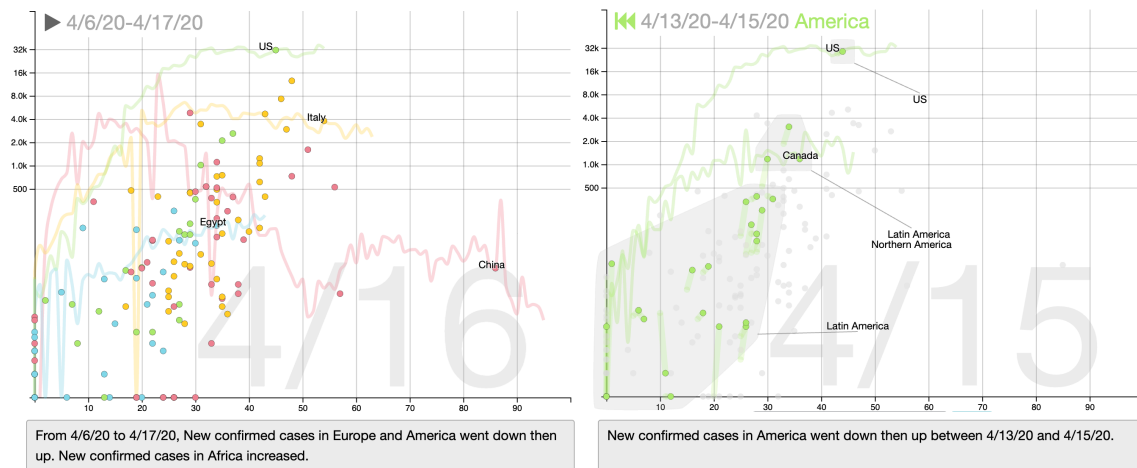


Fig. 8. Roslingifier detects a group of events (Apr 6–17) in COVID-19 dataset. Left: an intro summarizing the interval. Right: detected trough event in America (Apr 13–15). Traces highlight countries' movement for the entire time range.

auto-generated narratives: “It has the highest number of new cases each day, but the slope of the trace has become gentler after the 20th day.” “Latin and south America in the middle (Canada, Peru, Brazil) are 10 days behind, and their slopes are less steep than that of the United States.”

In this use case, we show that Roslingifier can compose a story from other types of time-series data. Figure 13 in the appendices shows a screenshot of the system displaying the entire story for this use case.

## 7 USER EVALUATION

We conducted two user studies to evaluate Roslingifier from both the authoring as well as the audience perspective. We also interviewed three domain experts on using Roslingifier. Here we describe these user studies and their results.

### 7.1 User Study: Authoring

The goal of our authoring user study was to understand how well Roslingifier helps storytellers create a data presentation and deliver the story.

*Method.* We recruited fourteen participants with experience at creating data presentations from a local university. As the participants entered the experiment room, we collected their written consent. We then requested them to fill out a pre-experiment questionnaire on demographic information, including age, gender, education level, and their experience using existing tools. Then we began a training session, which lasted 15–20 minutes on average. During this training, we introduced Roslingifier and demonstrated how to use the system. We then asked participants to perform all basic tasks involved in creating a story, such as adding and deleting events, changing the event order and playback time, editing narratives, moving the position of the cluster labels, and turning on and off the bubble labels. We used the child mortality and babies per woman data in the Gapminder data for this training (Appendix D).

The main task of the study was for participants to create a data presentation using Roslingifier with the data on the relation between life expectancy and income in the Gapminder dataset (Sec. 6.1). The participants' stated goal was to describe trend changes in life expectancy and income over 200 years of world history to a general audience. They were

		Author		Author	Audience
Evaluation of the system	Ease of learning	6.00 (1.11)	Perceived usefulness of features	Country labeling	6.21 (0.89) 6.24 (1.07)
	Ease of use	5.50 (1.02)		Cluster labeling	5.36 (1.39) 5.88 (1.17)
	UI Intuitiveness	5.36 (1.08)		Country tracing	6.57 (0.51) 5.82 (1.34)
	Efficiency	5.71 (0.99)		Slowdown	5.86 (1.03) 6.71 (0.58)
	Fun	6.21 (0.97)		Rewind	4.93 (1.54) 6.09 (1.00)
	Finding new insight	6.50 (0.65)		Comprehension	– 5.79 (1.07)
	Highlighting insight	6.07 (0.73)		Narratives	– 6.65 (0.54)
	Satisfaction	5.86 (0.53)			

Table 3. Summary of post-experiment ratings for authors and audiences (Sec. 7.1 and Sec. 7.2); numbers represent means and (standard deviation) on a 7-point Likert scale.

instructed to include any trend or event that they deemed of interest to such an audience. Because we assumed that participants often acquire background knowledge on given topics from online resources, and need to check whether presented materials are correct or not, we provided two Wikipedia pages (“Timeline of the 19th Century<sup>1</sup>” and “Timeline of the 20th Century<sup>2</sup>”), and allowed them to use the internet to search for additional information. The participants could use a maximum of 1.5 hours to create a presentation. After completing their data presentation, we asked participants to deliver the new data presentation to the experimenter.

To complete the session, we requested participants to fill out a post-experiment questionnaire on their user experience and usefulness of the features of Roslingifier using a 7-point Likert scale (7: the strongest agreement). We also interviewed them with open-ended questions about their opinion on the system. We recorded their screen activities and logged their actions on the interface during the experiment. We recorded screen with audio during the presentation.

*Results.* We initially recruited fifteen participants, but one of them failed to finish due to an unanticipated system failure. As a result, we analyzed data from fourteen participants. They were 25.2 years old on average ( $\sigma=1.53$ ) and either undergraduate (8) or graduate (6) students. They spent on average 1 hour and 9 minutes ( $\sigma=26$  minutes) to create presentations that were an average of 5 minutes 8 seconds long ( $\sigma=2$  minutes 41 seconds). We paid 12.68 GBP per person for their participation.

Overall, the participants positively assess Roslingifier, giving scores in all questions higher than the numerical average score (4), as shown in Table 3. Specifically, they thought that Roslingifier provides an intuitive user interface (5.36), and is easy to learn (5.5) and use (6.0) They also felt that Roslingifier helps them effectively find (6.5) and highlight (6.07) the insights. There are several features that helps participants find insights, including event detection, automatic grouping, min-max band in the line charts, and traces of countries. One participant, C14, also highly evaluated the automation because they helped save effort: “Automatic clustering and drawing traces save a lot of effort in creating a story compared to when doing it manually.” C3, who decided to emphasize wars and genocides, stated that “This system makes the process of extracting events from data easy and simple, [...] The min-max band in the line charts effectively show the general trends and outliers.” We also find that participants tend to actively use detected insights as a source for highlighting, such as clustering results; “Clustering the countries with similar movement and labeling them (Central Asia) tells a clear story and provides what to search.” (C6)

To questions for evaluating perceived usefulness of individual highlighting features, participants tended to respond positively, as shown in Table 3, where all of the country (6.21) and cluster (5.36) labeling, country tracing (6.57),

<sup>1</sup>[https://en.wikipedia.org/wiki/Timeline\\_of\\_the\\_19th\\_century](https://en.wikipedia.org/wiki/Timeline_of_the_19th_century)

<sup>2</sup>[https://en.wikipedia.org/wiki/Timeline\\_of\\_the\\_20th\\_century](https://en.wikipedia.org/wiki/Timeline_of_the_20th_century)

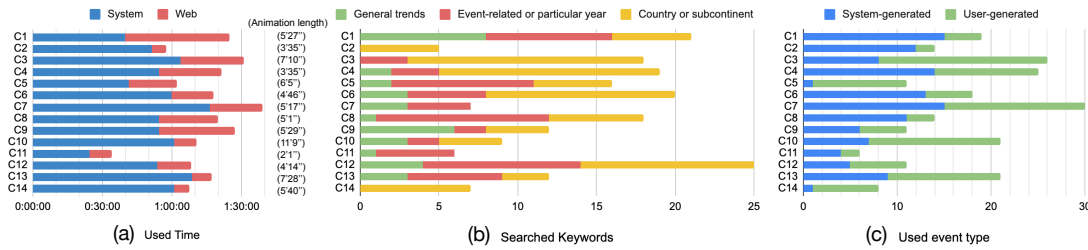


Fig. 9. (a) Time participants spent on the system and web search. (b) The number of searched keywords by their type. (c) The number of user-generated and system-generated events that participants used in their data presentation.

and slowdown (5.86) are considered useful in creating data presentations. A participant, C6, who focused on outlier countries, stated that “*Country traces are very helpful for emphasizing bubbles with different movements [...] I believe that the highlight part can also attract users in the presentation.*”

Some features were not frequently used based on participants’ intention. Rewind, for example, allows participants to help audience better understand details on the events by repeating explanation of periods. Some participants never used the operation, suggesting that they felt rewinding was not aligned with their story. For example, C5 and C12 desired to focus on describing events on world wars in a global perspective using all continents, so they thought it was not necessary to explain individual continents’ events. “*As World Wars affect across the world, I think the impact of the wars in a global perspective should be the main point of my presentation. [...] I did not find a place for using rewind in my presentation.*” (C5)

Our findings indicate that Roslingifier’s design allows participants to enjoy (6.21) their work: “*[Roslingifier] is very useful [for creating data presentations], and also fun to use as I can directly see the changes on the chart by selecting interesting periods.*” (C9) C4 expressed a similar sentiment: “*I like watching the changing history of the world in Roslingifier [...] I would recommend this to my friend who majors in history, as she would use this tool all day.*” C2 expressed that she really enjoyed using Roslingifier, and even formulated her own conclusion after using Roslingifier: “*Despite numerous wars, the entire world moves upward in the end. I felt that there is hope for humanity.*”

Figure 9 shows quantitative results from the study, where we find that participants spent significant time using Roslingifier. From our video analysis we can summarize how they performed their task in four stages. First, participants would play the animation from the beginning to understand pre-generated events. Then they would add or delete events to see the changes in value, read the provided Wikipedia pages, and search for information about outlier countries and what happened in the periods. Next they arranged events to create a storyline, highlighted countries, and searched for supporting reasons to write narratives. Finally, they played their data presentation again from the beginning, adjusted playback time of each event segment to be matched with desired speech speeds, moved cluster labels’ position, and polished the narratives. On average, participants used 74%/26% of the time on the system/web (Figure 9a).

Participants spent 18 minutes on average ( $\sigma=10$  minutes) to read search results. We can group their search topics into general trends, event-related, and specific countries or subcontinents, as Figure 9b shows. Participants made searches when they investigated patterns using Roslingifier’s features, such as line charts, detected events, country traces, and clusters. They used what they understood from searched results for enriching their presentation narratives, adding background, causation, and exact numbers. The patterns in the line charts (Figure 5b) often inspired the participants to search general trends to collect background knowledge of a certain period. For example, C9 searched for “*Europe population in the 19th century,*” when she observed from the line charts that Europe’s population increases in the early

1800s, while that of other continents remain stable. She used what she found from search result in the caption: *“From 1875 to 1906, the industrial revolution took place and capitalism developed [...]”*. Detected events also stimulated the participants, which led to event-related or time-related keyword searches. Country traces and clusters generated during the highlighting mode is one of the sources that make participants search for causation or effects of events, or exact numbers on countries. For example, C7 searched a period, “1917-1919,” to investigate a detected event and then reported the impact of the war: *“[...] World War I caused victims across all continents, reducing life expectancy around the world.”* C2 searched for “Russia Ukraine in 1933” when she found that the two countries are clustered together and showed a big drop in their life expectancy. From the search result, she provided the reason of the event in her presentation: *“From 1933 to 1934, [...] in Russia and Ukraine, life expectancy declined due to the Holodomor.”* Similarly, C12 found that Kazakhstan had a dramatic change in life expectancy while being clustered alone. He searched for “Kazakhstan in 1932 and Kazakhstan famine” to give the exact numbers of people died from the famine to audience: *“In 1932, 1.5 million people died in the Soviet Union due to the Kazakh famine, of which 1.3 million occurred in Kazakhstan.”*

Roslingifier initially generated fifteen system-generated events, and we observe that the tool effectively supports both automatic event detection as well as data exploration. We can categorize participants’ event use patterns into two groups (Figure 9c). The participants in the first group tended to keep the system-generated events. They individually checked the generated events and used search engines to better understand the events and provide detailed explanation in their story. C2, for example, used most of the system-generated events in the 1900s and additionally created two user-generated events on China’s Cultural Revolution and the oil crisis to strengthen her message—the major events in the 20th century that changed the world. She commented that *“The system is very helpful for trend analysis, I can create a few more events (i.e., China’s Cultural Revolution and the oil crisis) and pick some outliers (i.e., the middle east countries during the oil crisis) which make the story rich and interesting.”* In contrast, the participants in the second group were not much interested in the automatically generated events, but instead developed their own story by creating new user-driven events. For example, C5 deleted most of the system-generated events and created new events to focus on the changes in Europe in the first and second industrial revolutions, then address the World Wars in the 1900s from a global perspective: *“I wanted to give a presentation which are the most important events that I think in a global perspective.”* C3 also presented a distinctive story by adding 18 user-generated events. He assumed that the period where the minimum value in the line charts drops indicate something happened in a country, so created events for those periods with value drops to create a storyline. He commented that *“I wanted to show that changes in life expectancy are related to wars and genocides, and that changes in income are due to resources and changes in ideology.”*

Participants frequently used highlighting features to improve audience comprehension. They used 16.1 country labels on average ( $\sigma = 11.4$ ) to highlight the major countries of events or victorious and defeated countries of the wars, which are the main actors or results of historical events. For example, C9 labeled four middle east countries in the 1970s in the presentation, to highlight a sharp increase of income during the oil crisis. Rewind techniques are an effective technique to help audience better understand or review previous explanation with repetition. Participants tended to use not only rewinds, but also showcase temporal branching by rewinding multiple times for a period with highlighting techniques to develop engaging storylines. C1, for example, first explains World War I from a global perspective, then used the rewind four times to deliver a common pattern of the continents during the war. Then she used rewinds again to contrast movement patterns of the continents during World War II. C1 stated: *“The rewind technique emphasizes a very interesting point that in World War I, the life expectancy of the all continent dropped, but in World War II, each continent showed different movement.”* C1 gave 7 points to the rewind feature. A few participants used highlight to stress the message that they deliver, combining rewind with labeling. C4 used three rewinds in the 1800–1840 period in

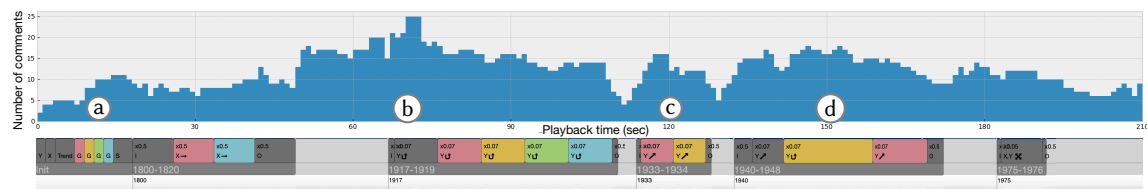


Fig. 10. The number of comments based on playback time and events in the video. See Sec. 7.2.

Europe, and labeled two countries at each rewind. She compared two countries: the UK, which started the industrial revolution, and Russia, which did not, by labeling them as a pair. She also labeled Norway and Sweden to focus on a conflict between the two countries, and then labeled Finland and Iceland to emphasize two countries with the biggest movement (due to Finnish wars and conflicts with neighboring countries).

## 7.2 User Study: Audience Viewing

We conducted another user study to demonstrate the quality of a generated data presentation and to evaluate how well the system features help audience understand data presentations.

*Method.* For this study, we created a 3.5-minute long data presentation video (Sec. 6.1) by recording the result in Roslingifier. No additional external video/image editing tools were used in the video production. We recruited participants from Prolific, a crowdsourced platform, to collect qualitative feedback on our data presentation video. After participants reached the experiment website and electronically signed our consent form, they were asked to provide demographic information. Then we requested them to read a tutorial, to watch a video, and to answer a simple question based on the contents of the video; the latter served as an attention trial to filter random clickers. Once they passed this attention trial, they were directed to the experiment page with our data presentation video, which was the main task in the study.

On the data presentation page, we asked participants to watch the video, and to leave at least three comments on (1) what made them think a segment interesting (e.g., visual effect) as well as (2) any insights that they gained from the video. We provided an annotation interface where participants could specify a start and an end time of a target video segment for every comment (Appendix G). We provided them with incentives based on the quality and number of their comments. After the study, we asked them to rate their comprehension level, the visual effects, video playback, and narratives used in the video with 7-point Likert scale (7: the strongest agreement).

*Results.* Overall, 36 participants (20 males) successfully completed the study and we paid 3.03 GBP/participant on average for their participation. They were 29.4 years old ( $\sigma=8.5$ ) on average, spent 19.3 minutes on average for the entire session, and made 6.61 comments per person (238 comments in total). The participants in general provided positive ratings for techniques as follows: comprehension (5.79/7.0), country labeling (6.24/7.0), cluster labeling (5.88/7.0), traces of countries (5.82/7.0), slowdown (6.71/7.0), rewind (6.09/7.0), and narratives (6.65/7.0).

Figure 10 shows the distribution of participants’ comments over the runtime of the video. First, we found 14 comments that recognize the explanation on the axes and trends in the initial segments, as shown in (a): “Plotting the descriptions of each corner physically on the plot really clarifies the meaning of the axes” as P1 stated. Second, the number of comments at (b) had 41 comments, where participants mentioned that the trace strongly attracts them in the segment, stressing

the big life expectancy drop after World War I: “*The trace of bubbles here clearly exemplifies how horrific the mortality rate was globally [...]*” The participants gave 5.82 points to the tracing technique.

Participants frequently mentioned labels throughout the video, but we found the moving labels on clusters in (c) to particularly capture audience attention (35 comments): “*The sudden increase in life expectancy in Kazakhstan and other Central Asian countries makes me curious to find out what was causing this sudden increase,*” P41 stated. The cluster labels received 5.88 points. We observe 30 comments at (d), where a few countries have significantly different movements from other countries. The participants also reported that the auto-generated labels on countries are useful in understanding the narratives; “*Country labeling is useful to tell a narrative about the outliers.*” They gave 6.24 points to the labels on individual entities.

Animation playback techniques received 21 comments, but participants gave higher score on them (slowdown: 6.71, rewind: 6.09). We attributed the high scores to the role of techniques—*divide and conquer stories* [44], which reduce the risks of overloading audience with information. Overall, participants thought the video is easy to understand (5.79 points) with interesting narratives (6.65 points). We also found that some participants expected more detailed stories on specific countries; e.g., “*I would like to have a little bit more info about Kazakhstan and why life expectancy rose so much.*” In [Appendix H](#), we provide clustering results of individual comments with respect to their contents.

### 7.3 Usefulness for Authors vs. Audiences

There are many features in Roslingifier which authors can use for their data presentations. However, there may be a disconnect between the features authors and audiences deem important and useful. To find out if there is any such gap, we asked the usefulness levels from both authors and audiences. [Table 3](#) (right) shows the survey results. We see that the authors put a higher value on the visual effects, such as country labeling and country tracing, as these effects play an explicit role in presenting entities of the narratives. But they did not focus on the usefulness of rewind, assuming that audiences do not need repeated but detailed explanation in narratives. C12 stated that “*I wanted to increase the year as the animation progresses, so the idea of going back and repeating the same period was a little confusing to me.*”

Audiences, on the contrary, tended to appreciate animation playback techniques (slowdown and rewind) more than authors. We think these results show that audiences prefer techniques that could allow additional time to review and digest given narratives with detailed explanations, leading to audiences’ better understanding and engagement with the story. Said P11, “*Separately showing how each continent/group of countries were affected by the Second World War was a good way to help break all the information down and keep the visual easy to digest.*”

### 7.4 Expert Review

We conducted an online, 1-hour interview with two data journalists and one narrative visualization developer (two males and a female), who have had more than 6 years of experience in the domain. One participant has been working for a broadcasting center (E1) and two others for major news media organizations (E2, E3: developer), at the time of the interview. During the interview, we showed them the same video that we created for the user study and explained how we created the video using the system. The feedback we received was overall positive. In particular, they liked the idea of auto-generated storytelling, as story presentation was one of the challenging tasks that designers face in their day-to-day operations. Said E1, “*It is an attractive tool, analyzing data stories generated automatically, [...] particularly the features such as legends, annotations, and groupings are very attractive.*”

In perspectives of data story presenters, rather than an audience, they showed their preferences in the event analytics and presentation editor views, which allow them to effectively create narratives or deliver the data presentation (“*The*

*lower timeline bar is helpful for the presenter to recognize the event in advance.*”–E3 stated). They even recommended that we include the event analytics view in the animation to strengthen the effectiveness of the data presentation video: “*The line graphs show the trend and context well, so putting them next to the scatter plot during the presentation would make a better storytelling technique.*”(E1) They agreed that it would be helpful if users could fine-tune the techniques (e.g., font size, color choice) before releasing them in the news room. When we finished the interview, E2 commented “*I would present this tool to the team. [...] We repeat data analysis and report the analysis result with a story to the management. This tool would help to shorten the repetitive processes with an easy user interface.*”

## 8 DISCUSSION AND LIMITATIONS

Our experiences designing, implementing, and evaluating Roslingifier has given rise to several insights and points for future improvement. We discuss these findings here.

*Additional Features.* While our creators tended to think the system was easy to learn and use, they requested several new functions that they would prefer to have. Examples include manual clustering, keyboard shortcuts, and undo functions. We conjecture that addition of manual clustering would improve the expressivity of data presentations, while that of shortcuts and undo would allow more efficiency in the creation process. Although we have so far not implemented these features, Roslingifier’s is open source and can be accessed here: (anonymized)<sup>3</sup>

*Event Detection Algorithms for Complex Patterns.* The goal with our work is to showcase an interactive prototype for data presentations, assuming small datasets for a general audience. However, the event detection algorithm implemented here is not sufficient to capture all complex events in the data, such as periodical events or asynchronous correlations over time. A future system, for example, could aim to capture richer patterns in large datasets, tailored to its target audience, data domains, and presentation goals.

*Comparisons to Existing Tools.* While there are many tools that can be used for data-driven storytelling, to our knowledge there exists no dedicated data presentation tool equivalent to Roslingifier. For example, Tableau Story Points and Microsoft Power BI are the closest existing tools, and provide filtering, highlighting, and captioning mechanisms. However, both of these tools are designed for general visual analysis, and lack the specialized data presentation as well as automatic event detection features of Roslingifier. Gapminder or Microsoft PowerPoint could be other alternatives, but they separate visual analysis from the presentation processes and do not provide specialized tools for story composition and editing. Because of such lack of an equivalent baseline, we chose not to provide a comparative evaluation in this work, but rather validate authoring and consuming its narratives separately.

## 9 CONCLUSION AND FUTURE WORK

We design Roslingifier based on the techniques derived from the analysis of data presentation videos. Roslingifier provides three views to support users quick prototyping of data presentation with auto-detected events and enabled storytelling techniques, such as gesture, visual effects, and animation playback. Our experimental results and expert feedback indicate that Roslingifier supports users in effectively creating data presentations that attract audiences using highlighted events and narratives. In the future, we plan on investigating automated data presentations for specialized data types beyond the scatterplots used in Roslingifier (e.g., maps, multi-dimensional data), as well as how human presenters incorporate their knowledge, insight, and experience into the data presentation produced by our tool.

<sup>3</sup>We will make the source code available upon acceptance of this paper so that anyone can extend it by adding their own features.

## ACKNOWLEDGMENTS

Quote from Hans Rosling.

## REFERENCES

- [1] Fereshteh Amini, Nathalie Henry Riche, Bongshin Lee, Christophe Hurter, and Pourang Irani. 2015. Understanding Data Videos: Looking at Narrative Visualization through the Cinematography Lens. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1459–1468. <https://doi.org/10.1145/2702123.2702431>
- [2] Fereshteh Amini, Nathalie Henry Riche, Bongshin Lee, Andres Monroy-Hernández, and Pourang Irani. 2017. Authoring Data-Driven Videos with DataClips. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 501–510. <https://doi.org/10.1109/TVCG.2016.2598647>
- [3] Benjamin Bach, Natalie Kerracher, Kyle Wm. Hall, Sheelagh Cappendale, Jessie Kennedy, and Nathalie Henry Riche. 2016. Telling Stories about Dynamic Networks with Graph Comics. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 3670–3682. <https://doi.org/10.1145/2858036.2858387>
- [4] Benjamin Bach, Zehong Wang, Matteo Farinella, Dave Murray-Rust, and Nathalie Henry Riche. 2018. Design Patterns for Data Comics. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 38:1–38:12. <https://doi.org/10.1145/3173574.3173612>
- [5] Alex Bigelow, Steven Mark Drucker, Danyel Fisher, and Miriah D. Meyer. 2017. Iterating between Tools to Create and Edit Visualizations. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 481–490. <https://doi.org/10.1109/TVCG.2016.2598609>
- [6] Chris Bryan, Kwan-Liu Ma, and Jonathan Woodring. 2017. Temporal Summary Images: An Approach to Narrative Visualization via Interactive Annotation Generation and Placement. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 511–520. <https://doi.org/10.1109/TVCG.2016.2598876>
- [7] Center for Systems Science and Engineering at Johns Hopkins University. 2019 (accessed September 8, 2020). 2019 Novel Coronavirus COVID-19 (2019-nCoV) Data Repository. <https://github.com/CSSEGISandData/COVID-19>.
- [8] Siming Chen, Jie Li, Gennady Andrienko, Natalia Andrienko, Yun Wang, Phong H Nguyen, and Cagatay Turkay. 2018. Supporting story synthesis: Bridging the gap between visual analytics and storytelling. *IEEE Transactions on Visualization and Computer Graphics* (2018). <https://doi.org/10.1109/TVCG.2018.2889054> To appear.
- [9] Zhutian Chen, Yun Wang, Qianwen Wang, Yong Wang, and Huamin Qu. 2020. Towards Automated Infographic Design: Deep Learning-based Auto-Extraction of Extensible Timeline. *IEEE Transactions on Visualization and Computer Graphics* 26, 1 (2020), 917–926. <https://doi.org/10.1109/TVCG.2019.2934810>
- [10] Noptanit Chotisarn, Junhua Lu, Libinzi Ma, Jingli Xu, Linhao Meng, Bingru Lin, Ying Xu, Xiaonan Luo, and Wei Chen. 2020. Bubble storytelling with automated animation: a Brexit hashtag activism case study. *Journal of Visualization* (2020), 1–15. <https://doi.org/10.1007/s12650-020-00690-7>
- [11] Dorin Comaniciu and Peter Meer. 2002. Mean Shift: A Robust Approach Toward Feature Space Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 5 (2002), 603–619. <https://doi.org/10.1109/34.1000236>
- [12] Weiwei Cui, Xiaoyu Zhang, Yun Wang, He Huang, Bei Chen, Lei Fang, Haidong Zhang, Jian-Guang Lou, and Dongmei Zhang. 2019. Text-to-Viz: Automatic Generation of Infographics from Proportion-Related Natural Language Statements. *IEEE Transactions on Visualization and Computer Graphics* 26, 1 (2019), 906–916. <https://doi.org/10.1109/TVCG.2019.2934785>
- [13] R. Eccles, T. Kapler, R. Harper, and W. Wright. 2007. Stories in GeoTime. In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology*. IEEE, Piscataway, NJ, USA, 19–26. <https://doi.org/10.1109/VAST.2007.4388992>
- [14] Gapminder Foundation. 2007. *Trendalyzer*. <https://www.gapminder.org/tag/trendalyzer/>.
- [15] Tong Gao, Jessica Hullman, Eytan Adar, Brent J. Hecht, and Nicholas Diakopoulos. 2014. NewsViews: an automated pipeline for creating custom geovisualizations for news. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 3005–3014. <https://doi.org/10.1145/2556288.2557228>
- [16] Samuel Gratzl, Alexander Lex, Nils Gehlenborg, N. Cosgrove, and Marc Streit. 2016. From Visual Exploration to Storytelling and Back Again. *Computer Graphics Forum* 35, 3 (2016), 491–500. <https://doi.org/10.1111/cgf.12925>
- [17] Neil Halloran. 2015. *The Fallen of World War II*. <http://www.fallen.io/ww2/>.
- [18] Harry Hochheiser and Ben Shneiderman. 2004. Dynamic query tools for time series data sets: Timebox widgets for interactive exploration. *Information Visualization* 3, 1 (2004), 1–18. <https://doi.org/10.1057/palgrave.ivs.9500061>
- [19] 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. <https://doi.org/10.1109/TVCG.2011.255>
- [20] Jessica Hullman, Nicholas Diakopoulos, and Eytan Adar. 2013. Contextifier: automatic generation of annotated stock visualizations. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 2707–2716. <https://doi.org/10.1145/2470654.2481374>
- [21] Jessica Hullman, Steven Mark 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. <https://doi.org/10.1109/TVCG.2013.119>
- [22] Nam Wook Kim, Benjamin Bach, Hyejin Im, Sasha Schriber, Markus H. Gross, and Hanspeter Pfister. 2018. Visualizing Nonlinear Narratives with Story Curves. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (2018), 595–604. <https://doi.org/10.1109/TVCG.2017.2744118>



- [23] Nam Wook Kim, Nathalie Henry Riche, Benjamin Bach, Guanpeng Xu, Matthew Brehmer, Ken Hinckley, Michel Pahud, Haijun Xia, Michael J. McGuffin, and Hanspeter Pfister. 2019. DataToon: Drawing Dynamic Network Comics With Pen + Touch Interaction. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 105:1–105:12. <https://doi.org/10.1145/3290605.3300335>
- [24] Nam Wook Kim, Eston Schweickart, Zhicheng Liu, Mira Dontcheva, Wilmot Li, Jovan Popovic, and Hanspeter Pfister. 2017. Data-Driven Guides: Supporting Expressive Design for Information Graphics. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 491–500. <https://doi.org/10.1109/TVCG.2016.2598620>
- [25] Kristijan Kolak. 2010 (accessed April 2020). *A Day In The Life Of Mister O*. <https://vimeo.com/10073883>.
- [26] Nicholas Kong and Maneesh Agrawala. 2009. Perceptual interpretation of ink annotations on line charts. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, USA, 233–236. <https://doi.org/10.1145/1622176.1622219>
- [27] Robert Kosara. 2016. Presentation-Oriented Visualization Techniques. *IEEE Computer Graphics and Applications* 36, 1 (2016), 80–85. <https://doi.org/10.1109/MCG.2016.2>
- [28] Robert Kosara and Jock D. Mackinlay. 2013. Storytelling: The Next Step for Visualization. *IEEE Computer* 46, 5 (2013), 44–50. <https://doi.org/10.1109/MC.2013.36>
- [29] Bum Chul Kwon, Florian Stoffel, Dominik Jäckle, Bongshin Lee, and Daniel Keim. 2014. VisJockey: Enriching Data Stories through Orchestrated Interactive Visualization. In *Proceedings of the Computation+Journalism Symposium*. Brown Institute for Media Innovation, New York, NY, USA, Article 2, 5 pages.
- [30] 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. <https://doi.org/10.1109/TVCG.2013.191>
- [31] Bongshin Lee, Nathalie Henry Riche, Petra Isenberg, and Sheelagh Carpendale. 2015. More Than Telling a Story: A Closer Look at the Process of Transforming Data into Visually Shared Stories. *IEEE Computer Graphics and Applications* 35, 5 (2015), 84–90. <https://doi.org/10.1109/MCG.2015.99>
- [32] Chunggi Lee, Sanghoon Kim, Dongyun Han, Hongjun Yang, Young-Woo Park, Bum Chul Kwon, and Sungahn Ko. 2020. GUIComp: A GUI Design Assistant with Real-Time, Multi-Faceted Feedback. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–13. <https://doi.org/10.1145/3313831.3376327>
- [33] Zhicheng Liu, John Thompson, Alan Wilson, Mira Dontcheva, James Delorey, Sam Grigg, Bernard Kerr, and John T. Stasko. 2018. Data Illustrator: Augmenting Vector Design Tools with Lazy Data Binding for Expressive Visualization Authoring. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 123:1–123:13. <https://doi.org/10.1145/3173574.3173697>
- [34] Junhua Lu, Jie Wang, Hui Ye, Yuhui Gu, Zhiyu Ding, Mingliang Xu, and Wei Chen. 2020. Illustrating Changes in time-series Data with Data Video. *IEEE Computer Graphics and Applications* (2020).
- [35] Sean McKenna, Nathalie Henry Riche, Bongshin Lee, Jeremy Boy, and Miriah Meyer. 2017. Visual Narrative Flow: Exploring Factors Shaping Data Visualization Story Reading Experiences. *Computer Graphics Forum* 36, 3 (2017), 377–387. <https://doi.org/10.1111/cgf.13195>
- [36] Jakob Nielsen, Rita Bush, Tom Dayton, Nancy Mond, Michael Muller, and Robert Root. 1992. Teaching Experienced Developers to Design Graphical User Interfaces. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 557–564. <https://doi.org/10.1145/142750.142986>
- [37] Charles Perin, Pierre Dragicevic, and Jean-Daniel Fekete. 2014. Revisiting Bertin matrices: New interactions for crafting tabular visualizations. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (2014), 2082–2091. <https://doi.org/10.1109/TVCG.2014.2346279>
- [38] Donghao Ren, Matthew Brehmer, Bongshin Lee, Tobias Höllerer, and Eun Kyoung Choe. 2017. ChartAccent: Annotation for data-driven storytelling. In *Proceedings of the IEEE Pacific Symposium on Visualization*. IEEE, Piscataway, NJ, USA, 230–239. <https://doi.org/10.1109/PACIFICVIS.2017.8031599>
- [39] Nathalie Henry Riche, Christophe Hurter, Nicholas Diakopoulos, and Sheelagh Carpendale. 2018. *Data-Driven Storytelling*. A K Peters/CRC Press, Boca Raton, FL, USA, 296 pages.
- [40] George G. Robertson, Roland Fernandez, Danyel Fisher, Bongshin Lee, and John T. Stasko. 2008. Effectiveness of Animation in Trend Visualization. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (2008), 1325–1332. <https://doi.org/10.1109/TVCG.2008.125>
- [41] Hans Rosling. 2006. Debunking myths about the ‘Third World’. <https://youtu.be/RUwS1uAdUcl>.
- [42] Hans Rosling, Ola Rosling, and Anna Rosling Rönnlund. 2018. *Factfulness: Ten Reasons We’re Wrong About the World—and Why Things Are Better Than You Think*. Flatiron Books, New York, NY, USA.
- [43] Arvind Satyanarayan and Jeffrey Heer. 2014. Authoring Narrative Visualizations with Ellipsis. *Computer Graphics Forum* 33, 3 (2014), 361–370. <https://doi.org/10.1111/cgf.12392>
- [44] Edward Segel and Jeffrey Heer. 2010. Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics* 16, 6 (2010), 1139–1148. <https://doi.org/10.1109/TVCG.2010.179>
- [45] Charles D Stolper, Bongshin Lee, N Henry Riche, and John Stasko. 2016. *Emerging and Recurring Data-Driven Storytelling Techniques: Analysis of a Curated Collection of Recent Stories*. Technical Report MSR-TR-2016-14. Microsoft Research.
- [46] Qianwen Wang, Zhen Li, Siwei Fu, Weiwei Cui, and Huamin Qu. 2019. Narvis: Authoring Narrative Slideshows for Introducing Data Visualization Designs. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (2019), 779–788. <https://doi.org/10.1109/TVCG.2018.2865232>
- [47] Yun Wang, Zhida Sun, Haidong Zhang, Weiwei Cui, Ke Xu, Xiaojuan Ma, and Dongmei Zhang. 2020. DataShot: Automatic Generation of Fact Sheets from Tabular Data. *IEEE Transactions on Visualization and Computer Graphics* 26, 1 (2020), 895–905.
- [48] Yun Wang, Haidong Zhang, He Huang, Xi Chen, Qiufeng Yin, Zhitao Hou, Dongmei Zhang, Qiong Luo, and Huamin Qu. 2018. InfoNice: Easy creation of information graphics. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA,

- 335:1–335:12. <https://doi.org/10.1145/3173574.3173909>
- [49] Zezhong Wang, Shunming Wang, Matteo Farinella, Dave Murray-Rust, Nathalie Henry Riche, and Benjamin Bach. 2019. Comparing Effectiveness and Engagement of Data Comics and Infographics. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 253:1–253:12. <https://doi.org/10.1145/3290605.3300483>
- [50] Martin Wattenberg. 2001. Sketching a graph to query a time-series database. In *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 381–382. <https://doi.org/10.1145/634067.634292>
- [51] Zhenpeng Zhao, Rachael Marr, Jason Shaffer, and Niklas Elmqvist. 2019. Understanding Partitioning and Sequence in Data-Driven Storytelling. In *Proceedings of the iConference (Lecture Notes in Computer Science)*, Vol. 11420. Springer, Cham, Switzerland, 327–338. [https://doi.org/10.1007/978-3-030-15742-5\\_32](https://doi.org/10.1007/978-3-030-15742-5_32)

## A REVIEWED DATA PRESENTATION VIDEOS

- [R1] BBC. Hans Rosling's 200 Countries, 200 Years, 4 Minutes. Nov 26, 2010. <https://youtu.be/jbkSRLYSojo>
- [R2] Gapminder Foundation. US in a converging world, Hans Rosling on CNN (Fareed Zakaria GPS). Mar 15, 2011. <https://youtu.be/WjVHvC9EeB4>
- [R3] Bill Gates. Hans Rosling: The River of Myths. Jan 31, 2013. <https://youtu.be/lYpX4l2UeZg>
- [R4] TED. The best stats you've ever seen. Jan 16, 2007. <https://youtu.be/hVimVzgtD6w>
- [R5] TED. Hans Rosling: The good news of the decade? Oct 7, 2010. [https://youtu.be/OT9poH\\_D2Iw](https://youtu.be/OT9poH_D2Iw)
- [R6] TED. Religions and babies | Hans Rosling. May 22, 2012. <https://youtu.be/ezVk1ahRF78>
- [R7] THINK Global School. Correlating income and life expectancy throughout history | Hans Rosling | TGS.ORG. Dec 1, 2015. <https://youtu.be/8suAGffNG6k>
- [R8] World Economic Forum. Davos 2015 - Sustainable Development: Demystifying the Facts. Jan 23, 2015. <https://youtu.be/3pVlaEbpJ7k>
- [R9] TED. Asia's rise – how and when | Hans Rosling. Nov 25, 2009. <https://youtu.be/fiK5-oAaeUs>
- [R10] TED. Hans Rosling on HIV: New facts and stunning data visuals. May 13, 2009. <https://youtu.be/3qRtDnsnSwk>
- [T1] YouTube Movies. An Inconvenient Truth. 2006. <https://youtu.be/x-VjNZBbjD4>
- [T2] RepresentUs. Unbreaking America: Solving the Corruption Crisis. Feb 27, 2019. <https://youtu.be/TfQij4aQq1k>
- [N1] ElectionsUK. The EU Referendum - FULL Results - BBC. Jun 26, 2016. <https://youtu.be/1TmUP1StPf0>
- [N2] CNN. John King: Trump enjoying a significant uptick in his political standing. Feb 6, 2020. <https://youtu.be/Is4uSbnfRWM>
- [W1] Mark 1333. Weather Events 2019 - Weather forecast - snow is coming (UK) - BBC News. Jan 31, 2019. <https://youtu.be/74sMio3c8Xo>
- [W2] UK Weather Forecast Channel. UK Weather Forecast HD: WORLD GLOBAL WEATHER FORECAST. Jan 3, 2018. <https://youtu.be/o1QMZcdN8Kw>
- [M1] CNBC Television. Jim Cramer unveils the scariest pattern in the chart book. Jan 17, 2020. <https://youtu.be/zQIZJQmnoV0>
- [M2] CNBC Television. Cornerstone Macro technician charts today's market carnage. Feb 24, 2020. <https://youtu.be/WKbnovqyIsg>

See supplemental material for additional appendices.

## B IN-DEPTH ANALYSIS ON DATA PRESENTATIONS



Fig. 11. Analysis of storytelling techniques in 7 data presentation videos, which are not presented in Figure 2.

## C CLUSTERING METHOD

When highlighting an event, we provide clusters to visually group the entities based on the position and movement. Even though only one legend is spotlighted, humans are not able to recognize every entity, especially when the entities are moving. Clustering enables users to discover important cues in highlighting periods, such as a country far away from others or a group of countries having similar movements.

We define a movement vector to represent the temporal movement of a bubble in the corresponding period and data dimensions. The movement of a bubble is represented as a matrix  $[V^{t_s}, V^{t_s+1}, \dots, V^{t_e-1}, V^{t_e}]$ , where  $V^t \in \mathbb{R}^n$  is a column vector representing values of the bubble at time  $t$ , and  $t_s$  and  $t_e$  indicate the starting and ending times of the event respectively. We use the mean shift algorithm [11] to cluster the movement vectors since it does not require the number of clusters unlike k-means or k-medoids algorithms. Parameters are chosen to have less than 10 clusters.

Note that the choice of a clustering algorithm is independent of the visualization design. Other clustering alternatives such as affinity propagation or hierarchical clustering can also be used.

## D USE CASE: CHILD MORTALITY VERSUS BABIES BORN PER WOMAN

We examine another chart using the data from Gapminder.org, child mortality versus babies born per woman (the number of 0-5-year-old dying per 1000). Size is the population. The data is from 1960 to 2018. For this data, we set the threshold to 2% for detecting events. Figure 12 presents the event exploration view of the system showing the detected events. The hull traces indicate different trends for each continent. Europe is already located in the bottom left corner in the 1960s with a small number of babies and a low mortality rate. Other continents also move toward the bottom left corner at a different speed. The system creates a group of events from 1960 to 1972 including 5 events from the world, Asia, Europe, and America. The countries are mostly spread in the beginning, in 1960 (a). European countries are located on the left while African countries are on the right. However, three clusters based on the direction of movement indicate there is not much relation between the position and movement direction. Child mortality in Asia dropped from 1960 to 1964 (b). The long trace of China shows a huge increase in the number of babies per woman after the great famine that swept across China between 1958 and 1961. The DROP(↘) events in Europe (c) and America (d) follow. Interestingly, no event in Africa was detected from 1960 to 1972. Child mortality in Africa slowly decreased in the period, then the major decrease in fertility is observed after the 2000s.

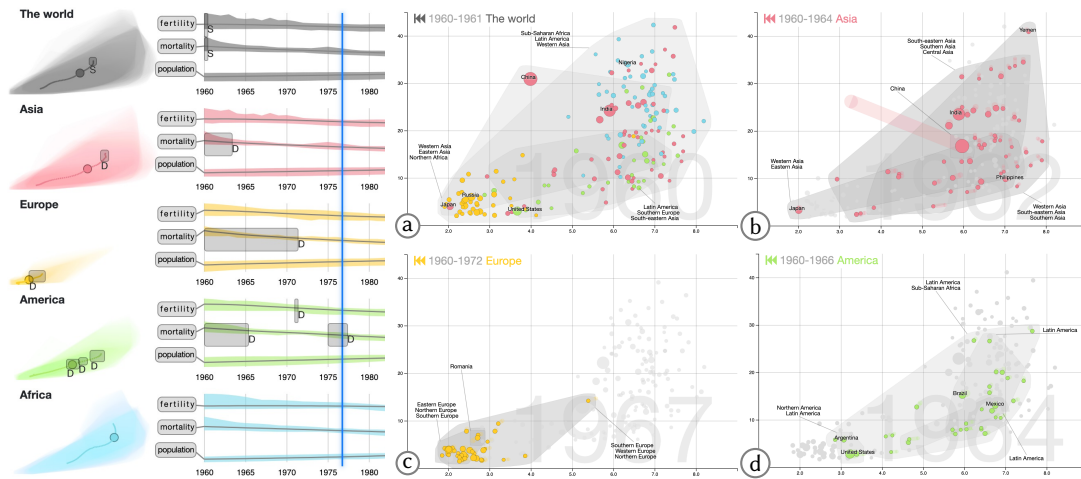


Fig. 12. The system detects 6 events from the data of child mortality and babies per woman. The charts on the right show the snapshots of the presentation output view displaying the SPREAD event in the world and DROP events in Asia, Europe, and America.

### E USE CASE: COVID-19 NUMBER OF NEW CONFIRMED CASES

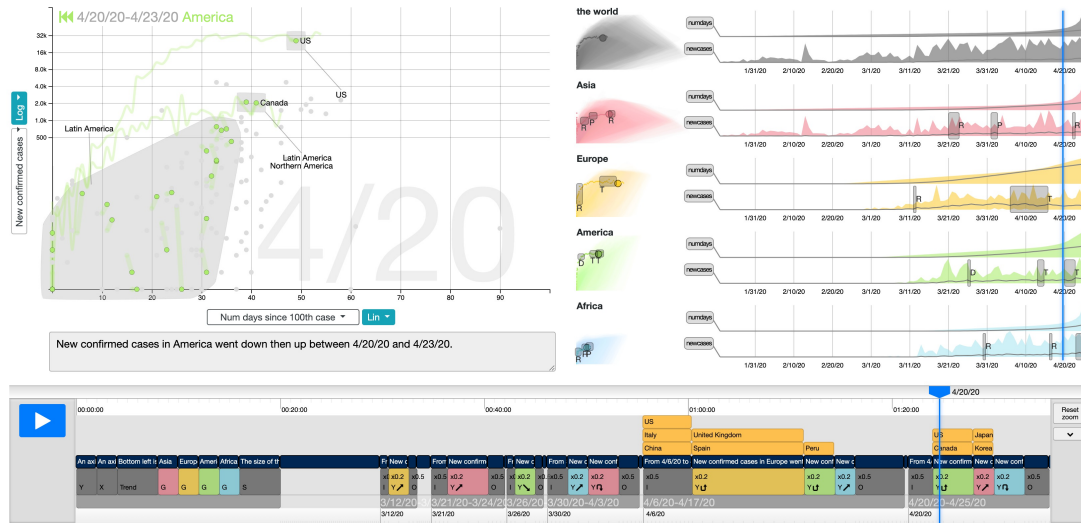


Fig. 13. The screenshot of the system displaying a story generated from COVID-19 time-series data. The X-axis is the new confirmed cases per day and Y-axis is the number of days since 100 confirmed cases. In this case study, we use the tracing technique to draw paths of countries for the entire time range.

## F AUTHORING USER STUDY: SELECTED COMMENTS BASED ON FEATURES MENTIONED

Feature	Comments
	<p>(C1) Good to see <b>complex timeseries data in one screen</b>.</p> <p>(C3) This system makes <b>the process of extracting events from data easy and simple</b>, and directly provides visual feedback in one screen.</p> <p>(C9) It is very useful and fun to use as <b>I can directly see the changes on the chart by selecting interesting periods</b>.</p>
Idea of system	<p>(C11) Great way to visualize timeline data into an animated scatterplot.</p> <p>(C2) Pre-generated animation <b>saves a lot of effort</b> in creating a presentation from the raw data.</p> <p>(C8) It is <b>easy to use</b>, and provides a good visualization of general trends.</p> <p>(C14) I like seeing a complex large amount of data in a single interface intuitively.</p> <p>(C5) An intuitive way to plot timeline data into a scatter plot.</p> <p>(C7) <b>The intuitive interface increases usability</b>.</p> <p>(C13) It intuitively visualizes the complex time-series data so that humans can easily understand.</p>
Fun	<p>(C3) <b>A fun platform</b> to create a presentation.</p> <p>(C8) The system is <b>fun to use and helpful to find events in history</b>.</p> <p>(C4) I like watching the changing history of the world.</p> <p>It would be <b>more useful and fun with domain knowledge</b>. Probably very useful for history education.</p> <p>I would recommend this to my friend who majors in history, as she would use this tool all day.</p> <p>(C1) It was very interesting that each continent has different movement in World War II where every continent drops in World War I.</p> <p>(C2) Despite numerous wars, the entire world moves upward in the end.</p> <p>I felt that there is hope for humanity.</p>
Finding insights	<p>(C12) Good visualization of the data by the time, it also efficiently finds events.</p> <p>(C11) It is <b>very effective to find new insight</b>. I do not have much knowledge of history, but I learned a lot by finding some changes in the line charts and searching the year on the web.</p> <p>(C1) Helpful to <b>find new events and outliers</b>.</p> <p>(C2) <b>Very helpful for trend analysis</b>, creater can pick some outliers which make the story rich and interesting. Detected events can help recall the historical events in the period, and <b>provide points to tell the story</b>.</p> <p>(C7) Detected events like drop and rise were very helpful.</p> <p>(C4) The line charts <b>effectively highlight drops and peaks</b>.</p> <p>(C3) <b>The min-max band in the line charts</b> effectively show the general trends and outliers.</p> <p>(C5) Color separation by continents (in the line chart) <b>highlights continent-level differences</b>.</p> <p>(C6) The drop events in the line chart help to find wars and outlier countries.</p> <p><b>Clustering the countries with similar movement and labeling them (Central Asia)</b> tells a clear story and provides what to search.</p> <p>(C8) Even without much knowledge in history, the system detects the significant events so that I was able to present the general trends in history.</p> <p>(C9) <b>The process of selecting periods from the line chart then see how countries move</b> in the period really helps me finding insights.</p>

Feature	Comments
Highlighting insights	(C14) <b>Automatic clustering and drawing traces save a lot of effort</b> in creating a story compared to when doing it manually. Animation / Clustering is the powerful point of the system.
	(C12) <b>Country labels and cluster labels help to emphasize</b> what happened in that countries/continents. It is <b>hard to see changes in population</b> .
	(C6) <b>Country traces are very helpful</b> for emphasizing bubbles with different movements.
	(C9) Want to turn on and off the gray clusters.
Animation	(C4) I like sequentially <b>introducing each color (continent) in the beginning</b> .
	(C2) <b>Intro/outro is very useful to introduce and recap the period</b> , but want to use only either intro or outro.
	(C11) <b>Intro/outro could be turn on and off</b> .
	(C13) Indication for the time point is effective in tracking the event and making animation. (C10) Good to <b>check the playtime and the next frames</b> while presenting. (C2) <b>Time sliders</b> on both the animation timeline and the event line charts can <b>easily tell the current point of time</b> .
System Improvement	(C6) Need <b>fine editing</b> of the created events, e.g., editing period.
	(C4) Want to filter events with the same kind – only show Rise events or Drop events. Better for data analysis.
	(C1) Wish there is a paused frame at the end of the animation for summarizing overall trends and story.
	(C1, C11) Want to <b>search country name</b> .
	(C13) Want to create <b>user-defined clustering</b> . (C3, C6, C7) Want to show multiple continents at the same time by <b>stacking two event frame</b> . (C4, C10, C14) Want to repeat the same events with different subtitles. (C2, C4, C10, C14) <b>Keyboard shortcuts</b> (Undo or delete) will be helpful.

## G AUDIENCE-SIDE USER STUDY: EXPERIMENT PROCESS

The experiment consists of 4 steps, as described in Figure 14. First, (1) participants are asked to fill out a demographic survey (e.g., gender, age, education) and to read tutorial pages that describe their tasks and rules. We explicitly prohibit participants from using mobile devices/tablets and low-resolution displays (1024x768), or creating new tabs or refreshing the page during the experiment. In case of a violation, the experiment does not continue and we do not collect their data. To draw rich comments, we notice that the top 20% of participants with the number of high-quality comments will receive a bonus. To check if they pay attention, (2) participants are asked to watch a video and answer two comprehension questions. Only who had passed the comprehension questions could proceed to the evaluation page. In the experiment page (3), participants are asked to write the insights or interesting features found in the video. We provide an interface to select a start and end time and leave the comment about the interval. Finally, participants are asked them to rate about the features (4) with the 7 Likert scale (7 being the strongest agreement). Here are the features:

- *Video content is easy to understand.*
- *Country labeling is helpful to understand the story of the video.*
- *Cluster labeling is helpful to understand the story of the video.*
- *Traces of countries are helpful to understand the story of the video.*
- *Slowing down for highlighting is helpful to understand the story of the video.*
- *Repeating the same time period for focusing on a legend is helpful to understand the story of the video.*
- *Narratives are helpful to understand the story of the video.*

**1. Demographic Survey**

Prolific ID   
 Age   
 Gender

**2. Comprehension Test**

Simple Question 1  
 Simple Question 2

**3. Experiment**

Start  End  Comment

**4. Post-experiment Questionnaire**

1. I think \_\_\_\_\_ is helpful to understand the story.  
 Strongly disagree ○ ○ ○ ○ ○ ○ ○ ○ Strongly agree

**3. Experiment**

WATCH THE VIDEO & ADD COMMENTS

Instructions

1. Click the play button when you are ready to watch
2. When you find some segments interesting, define the interval by pressing "Start" and "End" at the beginning and end of the segment, respectively.
3. Write about insights you gained; write about what made you feel the segment interesting (e.g., visual effects / narrative / etc)
4. Press "Add" after you finish your writing.
5. Repeat 1-4 until you cannot find more insights or interesting segments.

Note: this video does not include sound.

1945-1948 Asia

Australia and New Zealand  
 Western Asia

Western Asia  
 Southern Asia  
 South-western Asia

South-eastern Asia

South-eastern Asia  
 Southern Asia  
 Western Asia

South-eastern Asia

Life Expectancy

Income

In Asia, North and South Korea experienced a huge fall in income. Both income and life expectancy dropped in Japan. Other countries forge on to the upper right.

Start Time  End Time

Please leave your comment about each segment.

Ex) It is interesting to see most of the countries dropped in 19xx while some moved in the opposite direction.  
 Ex) The traces of the bubbles help me understand the changes in life expectancy.

Start  End  Comment Here

Add Clear

(0 / 600 characters)

Fig. 14. Four steps of the experiment used for crowd-sourced study. It consists of (1) demographic survey, (2) comprehension test, (3) experiment, and (4) post-experiment questionnaire.

## H AUDIENCE-SIDE USER STUDY: SELECTED COMMENTS BASED ON FEATURES MENTIONED

Feature	Comments
Initial Sequence	<p>(P1) <b>Plotting the descriptions of each corner</b> physically on the plot really <b>clarifies the meaning of the axes</b>.</p> <p>(P32) It was good seeing what each axis represented and <b>the blue writing really helped to summarize the meaning</b>. I liked seeing the different colours representing different continents but it was <b>hard to remember which colour was which</b>.</p> <p>(P8) I like <b>the color segments / size of bubble</b> way of displaying info on a 2 axis chart.</p> <p>(P12) The chart itself is <b>very clear and clean</b> and colours are good. the sizing of countries seems off.</p> <p>(P31) It's <b>unclear whether the size of circles represents</b> the same number of population throughout the video.</p>
Country Labeling	<p>(P12) <b>It makes me very curious when I see a random country drop so suddenly</b>. It would be <b>interesting to have that country labelled</b> so I can look it up afterwards and see why it dropped.</p> <p>(P31) I like the <b>labelling of countries of interest</b>. It might be nice to see some countries labelled throughout the entire graph cycle, such as the UK, US, China, India, etc. Though, this could be distracting to the overall message.</p> <p>(P19) The <b>labels</b> and different sized points <b>helped to differentiate between the growth/decrease of certain countries</b> and their populations.</p> <p>(P11) I really liked the <b>individual labels on some of the countries</b> –I know they were relevant to the information in that part of the video...but I wish they had been there from the start.</p> <p>(P14) <b>Country labeling</b> is useful to tell a <b>narrative about the outliers</b>.</p>
Cluster Labeling	<p>(P14) Later on in the video when <b>grey bubbles surround some countries dots</b>, I believe this could be a better method to <b>summarize interesting aspects of the video</b>. I believe the video benefits more <b>when groups of countries are labelled</b> instead of just a mass of dots.</p> <p>(P16) I like seeing when there is an <b>outlier in the data</b>, that it <b>shows which countries/regions they come from</b>.</p> <p>(P13) The <b>grey outlines around the countries</b> in each continent make it easier to <b>see continent-level movement trends</b>. The <b>labels on individual regions</b> make it <b>easier to see what is going on</b>. I also like <b>how the text stays stationary (easier to read)</b>.</p> <p>(P31) I like the <b>finer-grain separation of the boxes</b>—helps to see the <b>differences between groups of nations much better</b>.</p> <p>(P11) The <b>grey blobs</b> shifting upwards all at once really <b>emphasized what a big change it was across the whole region</b>, not just in one country.</p> <p>(P41) By using the <b>bubbles</b> you can see a <b>stark difference</b> between Australia/ New Zealand/Western/Southern Asia compared to North &amp; South Korea, Japan.</p> <p>(P23) I found the <b>cluster labeling to be more effective for my understanding</b> than the labeling of individual countries.</p> <p>(P12) <b>The labelling of the clusters didn't work for me at all</b>. I found it <b>confusing</b> and it visually made the chart more <b>difficult to understand</b>. the font on the clusters and the countries is the same. so I couldn't distinguish between them.</p>
Country Tracing	<p>(P41) Looking at the <b>traces of the bubbles</b> especially during the industrial revolution period <b>clearly highlights the differences</b> between whats occurring in the countries.</p> <p>(P19) The <b>trail</b> left from the population when it changed, helped made it <b>easier to see how the life expectancy relative to the income changed over time</b>.</p>



Feature	Comments
Country Tracing (cont.)	<p>(P34) The <b>drop really helps demonstrate how severe</b> this truly was at the time from 1917-1919.</p> <p>(P38) The <b>trace of bubbles here clearly exemplifies how horrific the mortality rate was globally</b> (with only a few small exceptions) as a result of the Spanish Flu and World War 1.</p> <p>(P12) This was the part that really <b>got my attention</b>. the <b>trace of the bubbles is so dramatic and works really well</b>. The <b>image tracing</b> thing is great. It made me really <b>curious what happened to some countries that it dropped so much</b>.</p> <p>(P37) It is almost <b>shocking to note the huge drop</b> in life expectancy as a result of the first world war, the <b>shifting bubbles make it very clear</b>.</p> <p>(P11) I liked seeing the <b>bubbles rise and fall</b> on the life expectancy axis - it helped me <b>understand how countries can recover</b> from things like war or disease, over time.</p> <p>(P32) The <b>fade lines</b> where the bubbles used to be <b>work really well</b> here because they are so long, which <b>shows how rapidly the situation changed</b>.</p> <p>(P1) That <b>different countries move in different directions is depicted well</b> but it <b>appears chaotic at times</b>.</p>
Playback Speed	<p>(P32) This whole section of not much happening helped me to <b>understand how rapid the changes were later on</b>.</p> <p>(P12) I like the way everything is <b>speeding across</b>.</p> <p>(P31) The <b>slowed-down version</b> is much <b>easier to read</b>, since otherwise it's a bunch of coloured circles changing spaces randomly. I liked the <b>slowdown and the traces, made the changes very clear and profound</b>.</p> <p>(P2) It could've been <b>slowed down at the start of the video</b> so the viewer has <b>more time to understand</b> what is going on and which colours represent which country.</p>
Rewind	<p>(P29) <b>Splitting the video by continent helped me to understand better</b> how different continents were affected by changes in life expectancy.</p> <p>(P30) The trails and <b>splitting up of the groups</b> helps me <b>understand how significant</b> the impact of war and disease were worldwide.</p> <p>(P41) By <b>individually highlighting each continent</b> you can <b>clearly see how dramatic of an impact</b> the First World War had on all continents.</p> <p>(P32) <b>Highlighting individual regions</b> really helped me see <b>how each region was separately impacted</b> over those years.</p> <p>(P37) It is helpful to <b>see different parts of various regions broken down individually</b>.</p> <p>(P11) <b>Separately showing</b> how each continent / group of countries were affected by the Second World War was a good way to <b>help break all the information down</b> and keep the visual easy to digest.</p> <p>(P1) It's staggering <b>how steep the drop in life expectancy</b> was during this period followed by the <b>steep rebound</b> and the fact this was <b>replicated around the world</b>. <b>Animation emphasizes the point well.</b></p>
Narratives	<p>(P12) What worked best for me was the <b>trivia and facts at the bottom</b>. I would have liked more of that.</p> <p>(P18) The <b>text narrative is great but a little hard to read</b> - perhaps highlight with contrasting design elements (font, colour, size)</p> <p>(P37) Make the narratives at the bottom stand out more, they <b>very much aid in understanding the graph</b> but due to the <b>size and font of the text</b> the eye is not initially drawn to it.</p>