Journal on Interactive Systems, 2022, 13:1 doi: 10.5753/jis.2022.2330

Fundamental elements and characteristics for telling stories using data

Mariana Borges [**Pontifical Catholic University of Rio Grande do Sul** | ***mariana.borges@edu.pucrs.br***]
Claiton Marques Correa [**Farroupilha Federal Institute** | ***claiton.correa@iffar.edu.br***]
Milene Selbach Silveira [**Pontifical Catholic University of Rio Grande do Sul** | ***milene.silveira@pucrs.br***]

Abstract

One of the biggest challenges today is how to present data to users in a clear and easy way in different domains of application. One way to do this is to tell stories with data, as this makes it possible to improve user engagement and understanding of a given context. By using narrative elements, it is possible to provide readers with better explanations of what is being presented. To understand the research scenario about this topic, a systematic review of the literature was performed, through a snowballing procedure. As results, the main characteristics of these visualizations were identified as well as the genres and tools used in their development, and the main visualization techniques used. These results demonstrate that the most used genres are slideshows, data comics, annotated charts, magazine style, and flow chart. One of the main open questions is about the genres used for each - different - type of data, demonstrating that the topic is emerging and necessary.

Keywords: *narrative visualization, storytelling, data visualization*

1 Introduction

There is a growing need to present data in a clear and easy-to-understand way to users in different application domains once a good visualization enhances users' comprehension (Heer et al., 2010; Ward et al., 2010).

Visualization tools provide resources to build dashboards whose primary goal is to explain a point of view or a discovery. There are a variety of data visualization techniques: tables, bar charts, line charts, bubble charts, and histograms are examples of the most traditional ones.

Improving data visualizations is an ongoing debate topic within the Information Visualization community. Hence, over the years, researchers investigated ways to do it. One way to do this is the use of narrative visualization techniques (Segel and Heer, 2010).

Narrative data visualization (*aka* storytelling) arose because, historically, human beings narrate facts to convey some information (Kosara and Mackinlay, 2013). Given this, many authors started to investigate the use of storytelling as a way to improve users' engagement, interaction, and comprehension from data visualization. Employing narrative elements in data visualization can offer ways to explain specific subjects not reached by the traditional techniques (Kosara and Mackinlay, 2013; Figueiras, 2014a).

In a seminal study, Segel and Heer (2010) characterized the components involved in narrative visualization into three dimensions: genres, visual narrative tactics, and narrative structure tactics. Each dimension comprises resources that together aid the designers of narrative visualizations.

The researchers have explored different tools, techniques, and patterns for narrative data visualization. For example, Riche et al. (2018) presented an introduction to data-driven storytelling. By gathering data visualization researchers and journalists, their book defined the topic and provided a set of examples for storytelling.

In other study, Ghidini et al. (2017) identified that Narrative Visualization could be considered a very recent research area, despite being emerging, growing, and receiving increasing attention from academia. Likewise, Santos and Silveira (2019) also noted that this is an emerging topic and identified the need to discuss and explore the fundamental concepts for designing narrative visualizations in theory and practice.

Considering the above and the ongoing goal of enhancing data visualization, we decided to investigate recent works on the topic. Thus, to delimitate our research, we defined five research questions to investigate which tools are used in the construction of narrative visualizations, the main application' domains in the use of these techniques, and the fundamental elements and characteristics in the development of narrative visualization able to engage the reader according to the author's intention. To accomplish this research, we employed a Snowballing procedure (Wohlin, 2014).

To present the research done, the next sections will detail consecutively its background (Section 2), the research methodology (Section 3), the results (Section 4), the threats to validity (Section 5) and the conclusions (Section 6).

2 Background

Advancing in data visualization research, Segel and Heer (2010), Hullman and Diakopoulos (2011), and Kosara and Mackinlay (2013), considered the use of narrative visualizations as a means to enhance user engagement and comprehension in data visualization. Narrative elements can offer explanations in a way hardly reached by traditional means. According to Segel and Heer (2010), some techniques that can be used along with narrative genres are the messages and interactivity. Messages can provide explanations about the

visualization, which are titles, subtitles, labels, annotations and audio. Interactivity allows the reader to have control over the visualization. Some common forms of interactivity in the narrative visualization are buttons navigation, filter, search, zoom, timeline, etc. It is important to use these elements very carefully, because just as they can help the reader, they can also hinder the intention of the author of the visualization, generating some confusion in the reader.

Narrative visualization differs from the concept of telling stories (Segel and Heer, 2010). Text or video stories usually show a sequence of events and progression. In contrast, data visualization is similar because it can be organized in a linear sequence but offers interactivity which can instigate ones to seek alternatives explanations.

In the seminal study of the topic, Segel and Heer (2010) organized the design space for narrative visualizations into three dimensions: genre, visual narrative tactics, and narrative structure tactics. The genre relates to ways to present history, the number of scenes, elements, ordering, and the number of frames.

The visual narrative dimension encompasses devices for assisting and facilitating a narrative, such as highlighting capability. The third dimension, narrative structure, involves mechanisms to communicate the narrative structure, like a progress bar and a time slider. They recognized seven genres of narrative visualization that can be more or less suitable according to the data under exploration: Magazine style, Annotated chart, Partitioned poster, Flow chart, Comic strip, Slide show and Film/Video/Animation. They noted that genres may work best depending on the type of narrative being told. The choice of these genres depends on a set of factors, such as the complexity of the data, the narrative and the target audience. According to them, the techniques and genres should be employed to balance author-driven and reader-driven narratives. This perspective led to narratives in which the readers are provided with more interactive devices, i.e., reader-driven, and narratives with less interactivity, author-driven. The latter is more linear and tells the story.

Following the studies, Hullman and Diakopoulos (2011) explored the rhetoric in narrative visualization and how it frames the data to convey a particular message. They identified a set of techniques and forms for rhetorical and how they can be used to directly or indirectly prioritize certain interpretations. They also demonstrated how they interact with users' knowledge and their social context.

In *Storytelling: the next step for visualization*, Kosara and Mackinlay (2013) observed that narrative visualization research involves other fields beyond computer science, such as psychology and social sciences. They presented researches directions regarding storytelling approaches and affordances, evaluation, how to make the visualization memorable, annotations and highlights, interactivity, and what to learn from other disciplines.

The study presented here seeks to identify novel narrative data visualizations researches and deepen the understanding of elements and fundamental characteristics to build them.

3 Methodology

Narrative visualization is an emerging topic. Therefore, to deepen the understanding of this subject's fundamental elements and characteristics, we performed a systematic review on it. To accomplish this, we used the Snowballing procedure for systematic reviews (Wohlin, 2014) .

Wohlin (2014) proposed this method that consists of defining an initial set of studies and then performing iterations of backward and forward steps seeking new papers. At each iteration, the papers under evaluation are analyzed one at a time. In the backward step, the reference list is scrutinized to seek new studies. In the forward step, studies that cite the one being evaluated are analyzed. The papers found in each iteration are stored in a pile for the next one. This process ends when no papers are found.

The Snowballing approach is considered an alternative to other systematic literature review methods. It is a way to advance in time, from a relevant study to identify another studies published since then. According to Wohlin (2014), the method should not necessarily be used as an alternative for database research, other approaches can be used for greater coverage of the literature. However, since the topic of this research is recent, we chose to use only Snowballing.

Before starting the literature review, we defined our research questions, and the inclusion and exclusion criteria. This information is addressed in Subsection 3.1, followed by the details of the snowballing procedure presented in Subsection 3.2.

3.1 Research Questions, Inclusion and Exclusion Criteria

The following **research questions** were defined to guide the review study:

- **RQ1**: What narrative visualization genres are used and their characteristics?
- **RQ2**: What data visualization techniques are being used?
- **RQ3**: What tools are being used to build visualizations?
- **RQ4**: Which application domains are using narrative visualizations?
- **RQ5**: What are the challenges when using narrative visualizations?

In **RQ1** we seek to find which genres are cited and used and what are their main characteristics in order to obtain an overview of this dimension. About **RQ2**, we seek to understand which data visualization techniques are most used in narrative visualizations. In **RQ3** we want to understand which are the most used tools and cited in the articles to develop a visualization. In **RQ4**, we search an overview of which application domains most use narrative visualizations, for example: education, health, entertainment, etc. Finally, in **RQ5** we researched the main challenges found by designers when using to develop narrative visualizations.

To conduct the Snowballing iterations and assist the researchers in deciding whether a paper should be included in the review or not, we defined the following inclusion and exclusion criteria.

The inclusion criteria are:

- The paper should have been published between 2010 [1] and 2022;
- The paper should have been published in English;
- The title, keywords and the abstract should be related to the research topic.

The exclusion criteria are:

- The publication should not be a tutorial, workshop, technical report, or a thesis;
- Duplicate papers.

3.2 Snowballing Procedure

The first step was to **identify the initial set of papers**. According to Wohlin (2014), there is no recipe to define this initial set. However, he suggested that it be identified by searching on Google Scholar to avoid bias favoring specific publishers. In addition, the studies must come from different communities and different publishers, years, and authors. One possibility to approach this in snowballing is to identify a seminal or highly cited paper in the area of the systematic literature study.

Therefore, considering the suggestions mentioned before, we searched on Google Scholar using the keywords *"narrative visualization" or "storytelling"*. We found 911 papers in the first query. Analyzing the results provided, we choose three studies to compound the initial set for starting the snowballing iterations.

The three initial studies were chosen due to their relevance to the topic. The first, published by Segel and Heer (2010) was chosen as it is the seminal article in the area, with great importance for the continuation of the research on narrative visualization. The other two articles were also chosen for their relevance to the topic, ability to answer the questions of this research and also due to the high number of citations in Google Scholar. The three studies are presented bellow.

> **P1.** Segel, Edward, and Jeffrey Heer. "Narrative visualization: Telling stories with data." IEEE transactions on visualization and computer graphics 16.6 (2010): 1139-1148.
> **P2.** Hullman, Jessica, and Nick Diakopoulos. "Visualization rhetoric: Framing effects in narrative visualization." IEEE transactions on visualization and computer graphics 17.12 (2011): 2231-2240.
> **P3.** Kosara, Robert, and Jock Mackinlay. "Storytelling: The next step for visualization." Computer 46.5 (2013): 44-50.

After defining the initial set, we proceeded to the **iterations** of backward and forward snowballing.

In the first iteration, each paper from the initial set is taken at once. Then, the **backward snowballing** step was performed. We analyzed the reference list of each paper. From this list, we initially chose candidate studies based on the inclusion and exclusion criteria.. In the following, we analyzed each candidate's title, publication venue, authors, and the most relevant parts, such as the introduction, methodology, and results. We aimed to check if they addressed narrative genres, data visualization techniques, narrative structures, or elements that could help us answer our research questions. Finally, we pushed each paper selected into a pile for the next iteration.

The **forward snowballing** step involves identifying new papers based on those papers citing each paper in the set being analyzed. Each paper's citations were found using Google Scholar. We checked the information of citing papers similarly as approached in the backward step, i.e., we initially chose candidates based on inclusion and exclusion criteria and then checked title, publication venue, authors, and the most relevant parts. Each selected paper was included in a pile for the next iteration of backward and forward steps.

In the **first iteration**, after backward and forward snowballing, from the initial set, 2 papers was selected in the backward step and 46 papers were selected in forward step. In the second iteration, after removing articles examined in previous iterations, and applying the inclusion and exclusion criteria, no new papers were selected. Thus, considering the three papers from initial set and the 48 selected during the process, a total of 51 papers were selected. Figure 1 outlines the snowballing procedure performed in this study and Table 1 shows the selected papers that are in accordance with the research topic, bringing relevant results to our questions. The analysis of the papers to be selected was done by pairs.

4 Results

This section presents the results from the analysis done in the 51 studies selected. To answer our research questions, we organized the results considering six dimensions: genres, visualization techniques,tools, application domains, approach, and structure adopted. Following, we present how the dimensions answer the research questions defined in our study. Figure 2 shows highlights from each dimension.

- The dimension **genres** answers **RQ1**: What narrative visualization genres are used and their characteristics?
- The dimension **visualization techniques** answers **RQ2**: What data visualization techniques are being used?
- The dimension **tools** answers **RQ3**: What tools are being used to build visualizations?
- The dimension **application domains** answers **RQ4**: Which application domains are using narrative visualizations?
- All dimensions answer **RQ5**: What are the challenges when using narrative visualizations?

4.1 Genres

According to Segel and Heer (2010), seven **Genres** can be used to build a narrative visualization, and they can be more or less suitable depending on the data available to construct the narrative. To identify the genres used in the papers selected, we fully read them. Our analysis revealed that the genres often used are **Slideshow**, **Data comics**, **Annotated**

[1] We defined the start period considering the main seminal paper in this field, published in 2010 by Segel and Heer (2010).

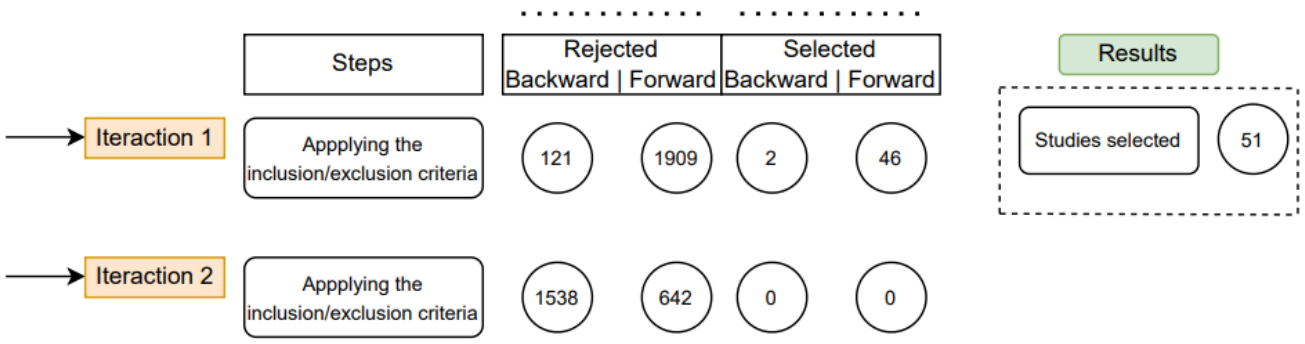

Figure 1. Snowballing design.

Table 1. Selected papers per year.

Year	Papers
2010	Segel and Heer (2010)
2011	Hullman and Diakopoulos (2011)
2012	Ma et al. (2011)
2013	Hullman et al. (2013), Kosara and Mackinlay (2013) Lee et al. (2013), Liu et al. (2013)
2014	Figueiras (2014a), Figueiras (2014b), Robinson et al. (2014), Yousuf and Conlan (2014)
2015	Lee et al. (2015)
2016	Arya et al. (2016), Boyd Davis et al. (2016), Brehmer et al. (2016), Bryan et al. (2016) Chan and Qu (2016), Gutiérrez and Pérez (2016), Stolper et al. (2016), Wang et al. (2016)
2017	Arévalo and Dewan (2017), Brolcháin et al. (2017), Carpendale et al. (2017), Echeverria et al. (2017) McKenna et al. (2017), Qiang et al. (2017)
2018	Bach et al. (2018), Chen et al. (2018), Comet (2018), Echeverria et al. (2018), Moretti et al. (2018) Nowak et al. (2018), Seyser and Zeiller (2018), Thöny et al. (2018), Tong et al. (2018)
2019	Chen et al. (2019), Rodrigues et al. (2019), Zhao et al. (2019), Zhi et al. (2019)
2020	Liem et al. (2020), Weber (2020)
2021	Joshi (2021), Lan et al. (2021), Latif et al. (2021), Norambuena et al. (2021), Roth (2021) Wang et al. (2021), Yang et al. (2021), Zhao et al. (2021)
2022	Zhao and Elmqvist (2022), Xu et al. (2022)

charts, **Magazine style**, and **Flow chart**. Figures 3 to 6 illustrate examples of the genre identified in the studies. The data used to create these sketches were obtained in the Our World in Data [2].

Figure 3 shows an example of **Slideshow** genre. This genre follows the typical slideshow format but incorporate iteration mechanisms that seek to engage and facilitate readers' understanding. From the set of selected papers, three explicitly mentioned the use of this genre. Wang et al. (2016) presented an approach to help researchers and graduate students in reading academic papers. They employed interactive slides to present literature reviews quickly. In another study, Zhi et al. (2019) proposed comparing slideshows and vertical data arrangement to investigate how they affect the reader's understanding of narrative visualizations. To evaluate this, the authors conducted a study with 180 participants. They measured the effect of the layout on user's engagement, understanding of the story content, and their ability to recall the story information. Arya et al. (2016) explored the slideshow genre to enhance the users' visit to a museum. The application developed in their study allows a user to plan a visit, put this plan into practice, and then construct a presentation summarizing moments of this visit to be shared with other visitors.

Figure 4 illustrates an example of a Data comic. When using **Data comics**, the author has the benefit of a taxonomy of transition types consisting of the moment to moment, action to action, subject to subject, scene by scene, aspect by aspect, and non-sequential transitions (Segel and Heer, 2010).

Zhao et al. (2021) listed five useful dimensions for the use of resources and genres identified by Segel and Heer:

- Target audience for which the story is intended;
- Time and space: what is the delivery mechanism of time and space for the story;
- Media components: What are the visual and sound building blocks employed?;
- Data components: how is data transmitted to the reader?;
- Viewing sequence: how is the media viewed by the reader?

[2] https://ourworldindata.org/

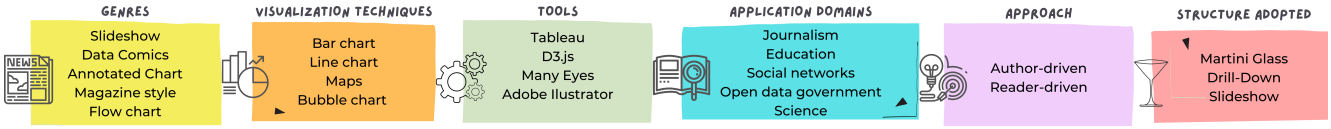

Figure 2. Highlights from the dimensions.

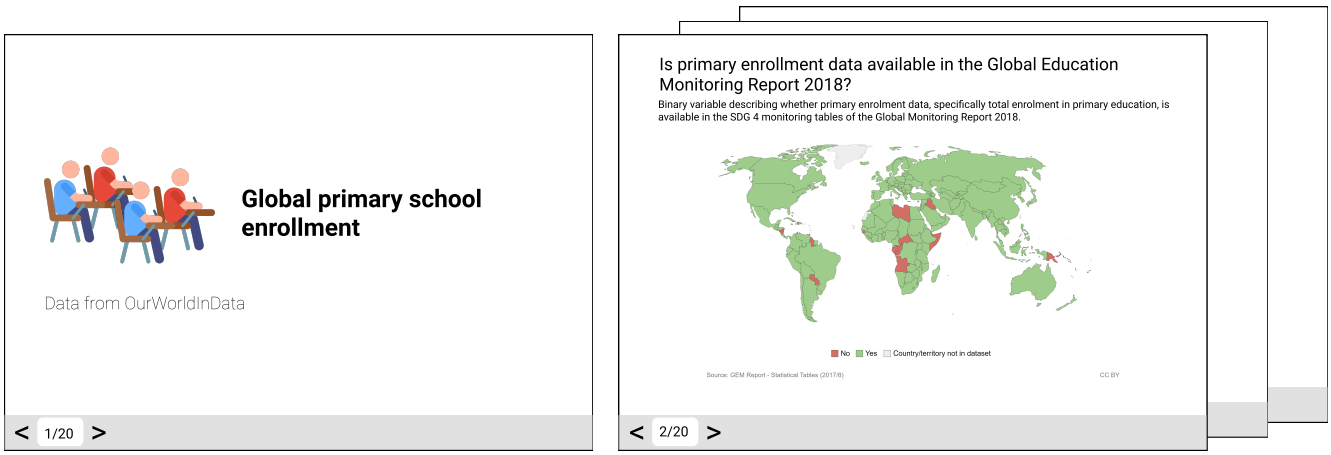

Figure 3. Slideshow example.

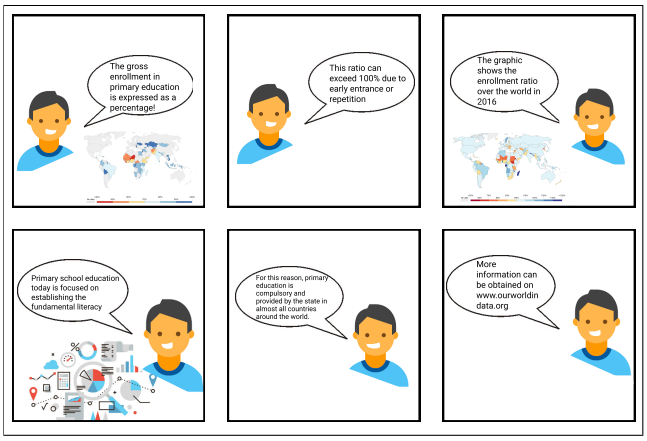

Figure 4. Datacomics example.

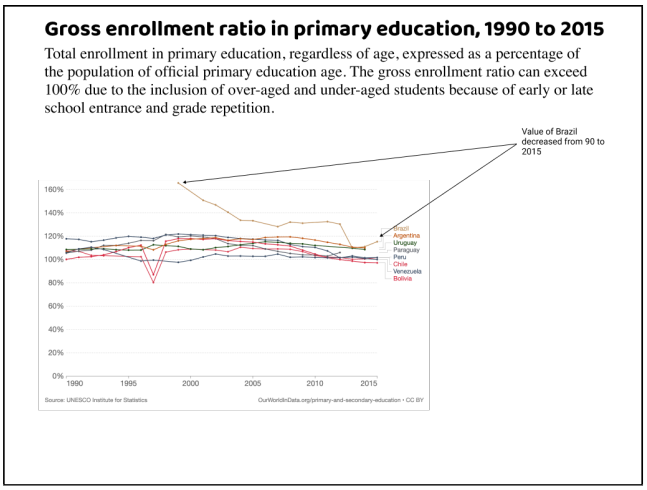

Figure 5. Annotated chart example.

They concluded that understanding these dimensions turns much easier to classify and use narrative genres and other resources.

From the selected papers, Zhao et al. (2019) and Wang et al. (2021) explored data comics. Zhao et al. (2019) investigated aspects of partitioning and sequence as mechanisms for comic strip narration. They implemented a data comics authoring tool to compare basic infographics with infographics enhanced with captions, texts, and dotted boxes organized as partitioned panels, a characteristic of comic strips. Wang et al. (2021) show ways to make comics interactive. They describe that adding interactivity to data comics can allow for non-linear storytelling, customization, levels of detail, explanations, and potentially enriched user experiences.

Figure 5 shows an example of an annotated chart. **Annotated charts** are mainly used with education data. Segel and Heer (2010) said that this genre favors the understanding of some information for the readers, and it should not be used in excess, only in essential cases, according to the author's intention.

Two studies from Echeverria et al. (2018, 2017) employed annotated charts in the education domain. The goal was to improve teachers' and students' learning by adding narrative elements in learning dashboards to enhance their communicative power. In the first study, the authors explored what data storytelling elements could be combined with the visualizations. In the second study, they integrated the findings of the first in a model to improve the visualizations of student's progress. Gutiérrez and Pérez (2016) aimed to fill a gap they perceived in research exploring narrative elements in the Business Intelligence (BI) process. The authors developed a methodology intended to enhance the interpretability of the information presented in BI dashboards by adding in these dashboards narrative elements.

Narrative visualizations are often related to benefits, such as improving data interpretation, user engagement, and increasing information memorability. In this sense, Liem et al. (2020) investigated these benefits by employing narrative visualizations (using annotated charts) to European immigration data. As a result, the authors claimed that researchers should be more careful in expecting such benefits from nar-

rative visualizations once they found smaller effects on users than expected.

The study from Chan and Qu (2016) employed a **Flow chart**, according to the example in the Figure 6, style to convey financial news. This genre presents information freely explored by the users, i.e., the author of the narrative did not define an ordering to convey the information.

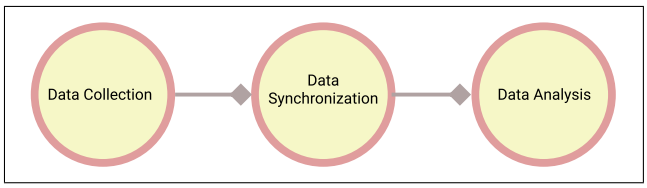

Figure 6. Flowchart example.

Seyser and Zeiller (2018) adopted the **Magazine style** (Figure 7 illustrates an example) to analyze visual storytelling in online journalism. This genre is usually adopted in journalism and it is recognized by its similarity with news in newspapers.

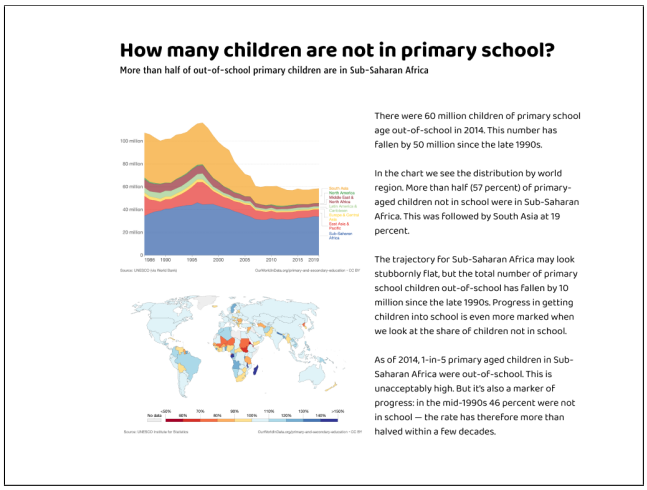

Figure 7. Magazine style example.

4.2 Visualization techniques

The **Visualization techniques** used are important elements of the narratives and can be combined to tell the narrative. Segel and Heer (2010) exemplify it bringing that **bar chart** can use annotations to convey a new narrative, providing information that the viewer would not identify on its own and **animations** can update graphics. **Annotations** can even be combined with **highlights**, **animated transitions**, and **interactivity**. Common graphics can be built in steps, with annotations and explanatory animations.

Several visualization techniques appeared in the analyzed studies. The techniques that appeared the most were bar charts (Chan and Qu, 2016; Zhi et al., 2019; Wang et al., 2016; Zhao et al., 2019; Lee et al., 2015), line charts (Echeverria et al., 2017; Chan and Qu, 2016; Wang et al., 2016; Echeverria et al., 2018; Kosara and Mackinlay, 2013), maps (Figueiras, 2014b; Thöny et al., 2018; Liem et al., 2020; Lee et al., 2015; Brolcháin et al., 2017; Moretti et al., 2018) and bubble chart (Chan and Qu, 2016; Zhi et al., 2019; Wang et al., 2016; Figueiras, 2014b). Other techniques such as animation, annotation, illustration, highlight, timeline and pie chart were also used, although rarely. Figure 8 shows a word cloud with the most used techniques.

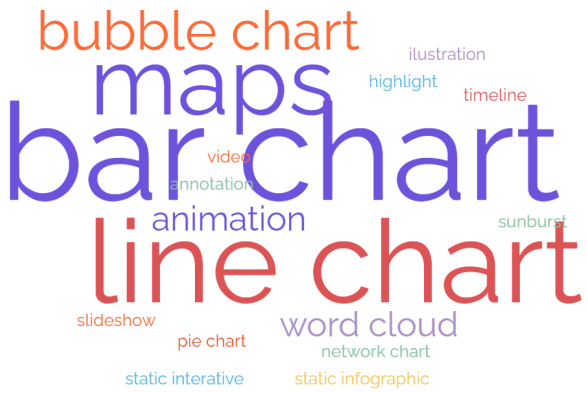

Figure 8. Visualization techniques used.

4.3 Tools

Another dimension examined in this review is related to **Tools** employed by the authors to build their narrative visualizations. To construct a narrative visualization, they can employ a handful of tools. These tools can vary from programming language libraries dedicated to building visualizations, such as **D3.js** or **Vue.js** libraries, to specific tools intended to support data visualization, such as **Tableau** or **Adobe Illustrator**. As long as narrative visualization has raised researchers' interest, data visualization tools started to provide facilities to permit authors to create stories with data. Table 2 shows the tools mentioned in the studies analyzed. Here, it is essential to say that not all studies explicitly report the tool adopted to construct a narrative visualization.

Table 2. Tools used

Tools	Papers
Many Eyes	Kosara and Mackinlay (2013) Segel and Heer (2010)
Tableau	Kosara and Mackinlay (2013) Lee et al. (2015) Segel and Heer (2010)
Adobe Ilustrator	Brolcháin et al. (2017)
D3.js	Lee et al. (2015)Nowak et al. (2018)
HTML	Zhao et al. (2019)
CSS	Zhao et al. (2019)
SVS	Ma et al. (2011)
Grapminder	Kosara and Mackinlay (2013)
Geotime stories	Segel and Heer (2010)

4.4 Application domains

The fourth dimension considered is the **Application domain** in which the studies adopted narrative visualization techniques to convey a message. Table 3 shows the genres used

in the studies belonging to each domain. It is important to say that this table summarizes studies in which both domain and genre were reported in the original studies. Therefore, studies that did not specify what domain they target were not considered. Five domains were identified in selected papers, **Culture**, **Education**, **Journalism**, **Open Government Data (OGD)**, and **Science**.

Arya et al. (2016) used narrative visualization in the **Culture** to let visitors share their experiences visiting the Canada Aviation and Space Museum's. They developed an application where people can plan their visit and share their findings with other visitors. By sharing their experience, the visitors are also creating a guide of the museum for others visitors.

By adding narrative elements to learning dashboards, Echeverria et al. (2018, 2017) target the improvement of teacher's and student's learning processes. The narrative elements added to the dashboards were labels, shaded area, key data points, title, and line. These elements are related to the dimensions of visual narrative and narrative structure reported by (Segel and Heer, 2010).

Three studies are related to the **Journalism** domain. Zhi et al. (2019) explored a New Your Times article to investigate how different information and text linking to image arrangements impact the users' comprehension of the data. The data conveyed the story of immigrants from banned nations. By providing different information arrangements, they perceived that users performed better in the slideshow layout and visualization by linking texts to other visualizations. By recognizing the inherent difficulty for people in reading and understanding financial news, Chan and Qu (2016) developed the FinaVistory tool to analyze European unemployment after 2008 and financial news about the oil price in 2014. The authors' goal was to provide a way to explain to the audience the relation between the two topics. The tool developed contains specific narrative structures, such as random access, messaging, and interactivity, to relate and convey the narrative. Seyser and Zeiller (2018) investigated the use of infographics, a type of layout that mixes characteristics from partitioned posters and annotated charts, to present complex data in online journalism and how they are integrated into articles.

Two studies combined narrative visualization for **Open Government Data (OGD)**. Gutiérrez and Pérez (2016) used data provided by the National Water Commission of Mexico (CONAGUA) to foster the use of narrative visualizations in BI dashboards. Liem et al. (2020) used European immigration data from 2000 to 2016 when researching the impact of narrative visualizations on users' attitudes and data comprehension.

One paper addressed the **scientific** domain when exploring narrative visualizations. Wang et al. (2016) used data based on references taken from academics citations on reviews to provide a way to read literature reviews. The results of employing narratives reached by this study showed that papers read become more effective and enjoyable.

4.5 Approach

The **Approach** can be author or reader driven (Segel and Heer, 2010). A fully **author-driven** approach is linear, it does not include reader interactivity. For example, non-interactive films and slideshows. This approach works best when it is needed direct communication, telling a story without alterations. A purely **reader-driven** approach has a lot of interactivity and there is no logical order.

Most examples of narrative visualizations use both approaches. An important feature of narrative visualizations is to find this balance, enabling limited interactivity in the context of a more structured narrative.

Only 15 papers explained the approach used: 4 bring a author-driven approach (Arya et al., 2016; Echeverria et al., 2017; Zhao et al., 2019; Echeverria et al., 2018); 6, a reader-driven approach (Figueiras, 2014b; Thöny et al., 2018; Wang et al., 2016; Zhi et al., 2019; Chan and Qu, 2016; Liem et al., 2020); and 5 identify both approaches (Segel and Heer, 2010; Seyser and Zeiller, 2018; Chen et al., 2019; Weber, 2020; Ma et al., 2011).

4.6 Structure adopted

The **Structure** dimension is strongly connected with the Approaches one. Segel and Heer (2010) described the structure as mechanisms that communicate the overall structure of the narrative to the viewer and allow she/he to locate itself within every view.

As described by Segel and Heer (2010), the **Martini Glass** visualization structure starts with an author-driven approach, uses questions, observations, or annotations to present a visualization. When the narrative proposed by the author is complete, it moves to a reader-driven stage, where the user is free to explore the data. The structure resembles a martini glass, with the stem representing the narrative in a single path and then opens representing several possibilities that the reader can enjoy.

Other example presented by Segel and Heer (2010) is the **Interactive slideshows** structure that incorporates interaction in the middle of the narrative within the boundaries of each slide. This structure allows the user to explore even more particular points of the presentation before advancing to the next stage of the story. It also allows interaction in the middle of the narrative and has a more reader-driven approach.

The other structure, **Drill-down story** presents a general theme and then allows the user to choose a more specific theme and get details and other related stories. This structure has more emphasis in the reader-driven approach, as the reader can choose when and what stories they want to view (Segel and Heer, 2010).

In the articles we reviewed, there is not a variety of structures identified, only those described by Segel and Heer (2010): **Martini Glass** (Weber, 2020; Lee et al., 2015), **Drill-down** (Figueiras, 2014b; Chan and Qu, 2016; Lee et al., 2015), and **Slideshows** (Echeverria et al., 2018; Wang et al., 2016; Zhi et al., 2019; Arya et al., 2016; Chen et al., 2019).

4.7 Challenges

To answer **RQ5**, we observe that there is still a gap to be able to identify which types of genres match the dataset so

Table 3. Genres employed per domain

Domain	Genres			
	Annotated Chart	Flow Chart	Magazine style	Slideshow
Culture				Arya et al. (2016)
Education	Echeverria et al. (2018)Echeverria et al. (2017)			
Journalism		Chan and Qu (2016)	Seyser and Zeiller (2018)	Zhi et al. (2019)
OGD	Gutiérrez and Pérez (2016)Liem et al. (2020)			
Science				Wang et al. (2016)

that the designer can develop a more concise narrative visualization. Some tools, such as Tableau, have some resources for the development of narrative visualizations, but they are still quite limited on that topic. Most still do not consider the elements of narrative visualizations as part of its functionality. Most articles cite the use of structures and approaches already identified by Segel and Heer (2010), but they do not bring different contributions on these dimensions.

5 Threats to validity

The following points summarize the main threats to the validity of our work and the strategies we followed to try overcome them:

- **Initial set of papers definition:** Before starting the snowballing process, we were required to define the initial set of papers. Two rounds of discussion were conducted until we agreed on what papers should compound this group. Then, after applying the rules defined by Wohlin (2014) we defined our initial set. Three researchers conducted this selection and the backward and forward rounds for including or excluding a paper. We did not encounter significant disagreements during the process as a whole.
- **Results:** Some papers of our set explicitly declared domain, narrative structures, or other characteristics that helped us place them into the six dimensions. However, the authors did not clearly define these characteristics in some cases. Then, the results represent our interpretation from the information obtained from the papers full read analysis. To mitigate the interpretation bias, we heid meetings among the authors to discuss the analysis produced.

6 Conclusion

In this paper we identified the fundamental elements and characteristics for telling a story using data. Among the discoveries, we were able to understand the genres, techniques, tools, application domains, approaches, and structures adopted in a narrative visualization.

Genres present challenges in the topic of narrative visualization. We can identify that the most used genres are slideshow, data comics, annotated charts, magazine style, and flow chart and characterize these genres, but there is still an opportunity to understand the types of data that can be used in each type of genre and **application domain** in question.

One important issue about genre is that Segel and Heer (2010) did not explored the relationship between genres and data types. They just claimed that the choice for one genre or another relies on the data under consideration. This relation remains unclear.

Regarding the **Tools**, some were listed among the papers, but it is not clear which ones are more adequate to construct narratives, or what types of knowledge the user must have in order to use them. A study aimed at identifying what tools are most appropriate for building narratives would be interesting, taking into account the user experience, the data used and the intended narrative.

About **Visualization techniques**, it is important to highlight the need for combinations that led the user feels involved in the story, independent of the application domain used. It is very important to use annotations and highlights in any type of graph, as they generate insights for the reader that would not be possible to identify.

The **Structure** and the **Approach** dimensions also complement each other. There were no new structures identified in the papers, other than those that Segel and Heer (2010) had described in their seminal paper, and not all papers identify these dimensions clearly. Therefore, the study of these dimensions can still be deepened, in order to understand if they depend on other dimensions.

One of the main opportunities listed in this research is to study genres in depth, creating real examples and connecting them to data types. There is also a need for in-depth studies on each dimension, aiming at the narrative views and the relationship between them. It is an emerging theme that deserves even more attention from the academic community.

Acknowledgements

This study was financed in part by Uol EdTech, and in part by the Coordenação de Aperfeiçoamento de Pessoal de Nivel Superior – Brasil (CAPES) – Finance Code 001

References

Arya, A., Gerroir, J., Mike-Ifeta, E., Navarro-Newball, A. A., and Prakash, E. (2016). Meseum: Personalized experience with narrative visualization for museum visitors. In *International Conference on Human-Computer Interaction*, pages 179–190. Springer.

Arévalo, S. and Dewan, A. (2017). Unveiling storytelling and visualization of data. In *14th Student Colloquium at University of Groningen*, page 38.

Bach, B., Wang, Z., Farinella, M., Murray-Rust, D., and Henry Riche, N. (2018). Design patterns for data comics. In *Proceedings of the 2018 chi conference on human factors in computing systems*, pages 1–12.

Boyd Davis, S., Vane, O., and Kräutli, F. (2016). Using data visualisation to tell stories about collections. In *Proceedings of EVA London*, pages 221–228. BCS Learning and Development Ltd.

Brehmer, M., Lee, B., Bach, B., Riche, N. H., and Munzner, T. (2016). Timelines revisited: A design space and considerations for expressive storytelling. *IEEE transactions on visualization and computer graphics*, 23(9):2151–2164.

Brolcháin, N. Ó., Porwol, L., Ojo, A., Wagner, T., Lopez, E. T., and Karstens, E. (2017). Extending open data platforms with storytelling features. In *Proceedings of the 18th Annual International Conference on Digital Government Research*, pages 48–53.

Bryan, C., Ma, K.-L., and Woodring, J. (2016). Temporal summary images: An approach to narrative visualization via interactive annotation generation and placement. *IEEE transactions on visualization and computer graphics*, 23(1):511–520.

Carpendale, S., Thudt, A., Perin, C., and Willett, W. (2017). Subjectivity in personal storytelling with visualization. *Information Design Journal*, 23(1):48–64.

Chan, Y.-Y. and Qu, H. (2016). Finavistory: Using narrative visualization to explain social and economic relationships in financial news. In *2016 International Conference on Big Data and Smart Computing (BigComp)*, pages 32–39. IEEE.

Chen, Q., Li, Z., Pong, T.-C., and Qu, H. (2019). Designing narrative slideshows for learning analytics. In *2019 IEEE Pacific Visualization Symposium (PacificVis)*, pages 237–246. IEEE.

Chen, S., Li, J., Andrienko, G., Andrienko, N., Wang, Y., Nguyen, P. H., and Turkay, C. (2018). Supporting story synthesis: Bridging the gap between visual analytics and storytelling. *IEEE transactions on visualization and computer graphics*, 26(7):2499–2516.

Comet, L. T. M. (2018). Visual narrative data-driven storytelling. *ACM Transactions on Multimedia Computing, Communications and Applications*, 1(1):1–20.

Echeverria, V., Martinez-Maldonado, R., Granda, R., Chiluiza, K., Conati, C., and Shum, S. B. (2018). Driving data storytelling from learning design. In *Proceedings of the 8th international conference on learning analytics and knowledge*, pages 131–140.

Echeverria, V., Martinez-Maldonado, R., and Shum, S. B. (2017). Towards data storytelling to support teaching and learning. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction*, pages 347–351.

Figueiras, A. (2014a). How to tell stories using visualization. In *2014 18th International conference on information visualisation*, pages 18–18. IEEE.

Figueiras, A. (2014b). Narrative visualization: A case study of how to incorporate narrative elements in existing visualizations. In *2014 18th International Conference on Information Visualisation*, pages 46–52. IEEE.

Ghidini, E., Santos, C. Q., Manssour, I., and Silveira, M. S. (2017). Analyzing design strategies for narrative visualization. In *Proceedings of the XVI Brazilian Symposium on Human Factors in Computing Systems*, pages 1–10.

Gutiérrez, A. and Pérez, C. B. (2016). Data visualization process through storytelling technique in business intelligence. *Avances en Interacción Humano-Computadora*, (1):49–52.

Heer, J., Bostock, M., and Ogievetsky, V. (2010). A tour through the visualization zoo. *Communications of the ACM*, 53(6):59–67.

Hullman, J. and Diakopoulos, N. (2011). Visualization rhetoric: Framing effects in narrative visualization. *IEEE transactions on visualization and computer graphics*, 17(12):2231–2240.

Hullman, J., Drucker, S., Riche, N. H., Lee, B., Fisher, D., and Adar, E. (2013). A deeper understanding of sequence in narrative visualization. *IEEE Transactions on visualization and computer graphics*, 19(12):2406–2415.

Joshi, A. (2021). Learning cues to improve the understanding of explanatory storytelling.

Kosara, R. and Mackinlay, J. (2013). Storytelling: The next step for visualization. *Computer*, 46(5):44–50.

Lan, X., Shi, Y., Wu, Y., Jiao, X., and Cao, N. (2021). Kineticharts: Augmenting affective expressiveness of charts in data stories with animation design. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):933–943.

Latif, S., Chen, S., and Beck, F. (2021). A deeper understanding of visualization-text interplay in geographic data-driven stories. In *Computer Graphics Forum*, volume 40, pages 311–322. Wiley Online Library.

Lee, B., Kazi, R. H., and Smith, G. (2013). Sketchstory: Telling more engaging stories with data through freeform sketching. *IEEE transactions on visualization and computer graphics*, 19(12):2416–2425.

Lee, B., Riche, N. H., Isenberg, P., and Carpendale, S. (2015). More than telling a story: Transforming data into visually shared stories. *IEEE computer graphics and applications*, 35(5):84–90.

Liem, J., Perin, C., and Wood, J. (2020). Structure and empathy in visual data storytelling: Evaluating their influence on attitude. In *Computer Graphics Forum*, volume 39, pages 277–289. Wiley Online Library.

Liu, S., Wu, Y., Wei, E., Liu, M., and Liu, Y. (2013). Storyflow: Tracking the evolution of stories. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2436–2445.

Ma, K.-L., Liao, I., Frazier, J., Hauser, H., and Kostis, H.-N. (2011). Scientific storytelling using visualization. *IEEE Computer Graphics and Applications*, 32(1):12–19.

McKenna, S., Henry Riche, N., Lee, B., Boy, J., and Meyer, M. (2017). Visual narrative flow: Exploring factors shaping data visualization story reading experiences. In *Computer Graphics Forum*, volume 36, pages 377–387. Wiley Online Library.

Moretti, M., De Chiara, F., and Napolitano, M. (2018). Beyond transparency: Making the italian public administration more accessible through data storytelling. In *IV*, pages 247–250.

Norambuena, B. F. K., Mitra, T., and North, C. (2021). De-

sign guidelines for narrative maps in sensemaking tasks. *arXiv preprint arXiv:2112.12205*.

Nowak, S., Bartram, L., and Schiphorst, T. (2018). A micro-phenomenological lens for evaluating narrative visualization. In *2018 IEEE Evaluation and Beyond-Methodological Approaches for Visualization (BELIV)*, pages 11–18. IEEE.

Qiang, L., Bingjie, C., and Haibo, Z. (2017). Storytelling by the storycake visualization. *The Visual Computer*, 33(10):1241–1252.

Riche, N., Hurter, C., Diakopoulos, N., and Carpendale, S. (2018). *Data-driven Storytelling*. A K Peters Visualization Series. CRC Press/Taylor & Francis Group.

Robinson, S. J., Williams, G., Parnami, A., Kim, J., McGregor, E., Chandler, D., and Mazalek, A. (2014). Storied numbers: Supporting media-rich data storytelling for television. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video*, pages 123–130.

Rodrigues, S., Figueiras, A., and Alexandre, I. (2019). Once upon a time in a land far away: Guidelines for spatio-temporal narrative visualization. In *2019 23rd International Conference Information Visualisation (IV)*, pages 44–49. IEEE.

Roth, R. E. (2021). Cartographic design as visual storytelling: synthesis and review of map-based narratives, genres, and tropes. *The Cartographic Journal*, 58(1):83–114.

Santos, C. Q. and Silveira, M. S. (2019). Qualitative studies on human-data interaction and data analysis collaboration in latin america. In *Proceedings of the IX Latin American Conference on Human Computer Interaction*, pages 1–4.

Segel, E. and Heer, J. (2010). Narrative visualization: Telling stories with data. *IEEE transactions on visualization and computer graphics*, 16(6):1139–1148.

Seyser, D. and Zeiller, M. (2018). Scrollytelling–an analysis of visual storytelling in online journalism. In *2018 22nd International Conference Information Visualisation (IV)*, pages 401–406. IEEE.

Stolper, C. D., Lee, B., Riche, N. H., and Stasko, J. (2016). Emerging and recurring data-driven storytelling techniques: Analysis of a curated collection of recent stories.

Thöny, M., Schnürer, R., Sieber, R., Hurni, L., and Pajarola, R. (2018). Storytelling in interactive 3d geographic visualization systems. *ISPRS International Journal of Geo-Information*, 7(3):123.

Tong, C., Roberts, R. C., Laramee, R. S., Wegba, K., Lu, A., Wang, Y., Qu, H., Luo, Q., and Ma, X. (2018). Storytelling and visualization: A survey. In *VISIGRAPP (3: IVAPP)*, pages 212–224.

Wang, Y., Liu, D., Qu, H., Luo, Q., and Ma, X. (2016). A guided tour of literature review: Facilitating academic paper reading with narrative visualization. In *Proceedings of the 9th International Symposium on Visual Information Communication and Interaction*, pages 17–24.

Wang, Z., Romat, H., Chevalier, F., Riche, N. H., Murray-Rust, D., and Bach, B. (2021). Interactive data comics. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):944–954.

Ward, M. O., Grinstein, G., and Keim, D. (2010). *Interactive data visualization: foundations, techniques, and applications*. CRC Press, 2 edition.

Weber, W. (2020). Exploring narrativity in data visualization in journalism. *Data Visualization in Society*, pages 295–311.

Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th international conference on evaluation and assessment in software engineering*, pages 1–10.

Xu, X., Yang, L., Yip, D., Fan, M., Wei, Z., and Qu, H. (2022). Fromwow'towhy': Guidelines for creating the opening of a data video with cinematic styles. *arXiv preprint arXiv:2202.02709*.

Yang, L., Xu, X., Lan, X., Liu, Z., Guo, S., Shi, Y., Qu, H., and Cao, N. (2021). A design space for applying the freytag's pyramid structure to data stories. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):922–932.

Yousuf, B. and Conlan, O. (2014). Constructing narrative visualizations as a means of increasing learner engagement. In *HT (Doctoral Consortium/Late-breaking Results/Workshops)*, pages 1–4. CEUR-WS.org.

Zhao, J., Xu, S., Chandrasegaran, S., Bryan, C., Du, F., Mishra, A., Qian, X., Li, Y., and Ma, K.-L. (2021). Chartstory: Automated partitioning, layout, and captioning of charts into comic-style narratives. *arXiv preprint arXiv:2103.03996*.

Zhao, Z. and Elmqvist, N. (2022). The stories we tell about data: Media types for data-driven storytelling. *arXiv preprint arXiv:2202.00047*.

Zhao, Z., Marr, R., Shaffer, J., and Elmqvist, N. (2019). Understanding partitioning and sequence in data-driven storytelling. In Taylor, N. G., Christian-Lamb, C., Martin, M. H., and Nardi, B. A., editors, *Information in Contemporary Society - 14th International Conference, iConference 2019, Washington, DC, USA, March 31 - April 3, 2019, Proceedings*, volume 11420 of *Lecture Notes in Computer Science*, pages 327–338. Springer.

Zhi, Q., Ottley, A., and Metoyer, R. (2019). Linking and layout: Exploring the integration of text and visualization in storytelling. In *Computer Graphics Forum*, volume 38, pages 675–685. Wiley Online Library.