Telling stories with data - A systematic review

Kay Schröder, Wiebke Eberhardt, Poornima Belavadi, Batoul Ajdadilish, Nanette van Haften, Ed Overes, Taryn Brouns, André Calero Valdez

***Abstract*—The exponential growth of data has outpaced human ability to process information, necessitating innovative approaches for effective human-data interaction. To transform raw data into meaningful insights, storytelling and visualization have emerged as powerful techniques for communicating complex information to decision-makers. This article offers a comprehensive, systematic review of the utilization of storytelling in visualizations. It organizes the existing literature into distinct categories, encompassing frameworks, data and visualization types, application domains, narrative structures, outcome measurements, and design principles. By providing a well-structured overview of this rapidly evolving field, the article serves as a valuable guide for educators, researchers, and practitioners seeking to harness the power of storytelling in data visualization.**

***Index Terms*—Visualization, Storytelling, Review**

I. Introduction

Storytelling stands as an age-old tradition, a cornerstone of human culture, ensuring the continuity of knowledge across generations. The power of stories emanates from their ability to engage our entire cognitive and emotional faculties, enabling us to empathize, connect, and understand the narrative's content. In our modern era, this storytelling tradition is being merged with data visualization, thereby becoming a crucial tool in comprehending complex data associated with global phenomena such as climate change, pandemics, and societal changes.

However, the application of traditional storytelling to the rapidly evolving domain of information visualization is not without its challenges. This integration necessitates a deeper understanding of the nuances of classical storytelling and an exploration of its potential applicability to data-driven narratives. With various interpretations of storytelling across different domains, a holistic approach is required, one that merges these diverse perspectives.

This paper addresses these challenges, aiming to provide a comprehensive review of storytelling in the context of data and information visualization. We identify a dual-fold problem: the need to consolidate various perspectives on storytelling and the need to understand the potential application of traditional storytelling techniques in the context of data-driven narratives.

K. Schröder is with HS Düsseldorf University of Applied Sciences, Düsseldorf, Germany
W. Eberhardt is with University of Twente, Twente, The Netherlands
P. Belavadi is with RWTH Aachen University, Aachen, Germany
B. Ajadilish is with Zuyd University of Applied Sciences, Heerlen, The Netherlands
N. van Haften is with Ruigrok, Amsterdam, The Netherlands
E. Overes is with Zuyd University of Applied Sciences, Heerlen, The Netherlands
A. Calero Valdez is with the Institute of Multimedia and Interactive Systems, University of Lübeck, Lübeck, Germany

Drawing upon literature studies, information visualization, and psychological and behavioral science, we explore the multifaceted nature of storytelling. Four research questions guide this exploration:

- What are the **theoretical or conceptual foundations** of data-driven storytelling, and how is the story creation process structured?
- How can data-driven storytelling be examined from a **narratological perspective**? What are the different narrative structures, and what goals do they serve?
- How can we structure individual data stories within a **taxonomy** that reflects a user's conceptual understanding of the story?
- What are the cognitive and non-cognitive **effects of storytelling**, and how were these effects measured empirically?

This study will be of interest to a wide audience, including researchers, practitioners, students, and educators from diverse fields such as narratology, visualization, design, journalism, data science, bioinformatics, communication, and computing. The insights presented here may serve as a primer for those new to visualization and storytelling and as a tool for educators seeking to structure storytelling concepts for their audiences. Moreover, the findings can guide researchers and practitioners in identifying gaps and potential future research directions in their respective fields.

The rest of this paper is structured as follows: We begin with an overview of related work, followed by a description of the method and procedure employed in this review. We then structure the selected literature based on our four research questions, concluding with a discussion on the findings and potential future research directions.

II. Related Work

Storytelling, while a long-established concept, has only recently been applied in the context of data visualization.

Segel and Heer [176] provided the first comprehensive review of data storytelling and narrative visualizations in 2010, identifying specific design strategies and relevant interaction paradigms. Their analysis also delineated different categories within genres, a classification we have adopted for our analysis of visualization types and application areas.

Building on this, Bach et al. [11] proposed grouping these genres into three categories based on their orientation: spatial, temporal, or a combination of both. Kosara and Mackinlay [102] further explored the role of both the setting and audience in three scenarios: self-running presentations, moderator-guided live presentations, and presentations for individual or small groups, which allow for more specific interactions with the audience [165].

Tong et al. [194] proposed a two-dimensional framework to categorize literature, focusing on elements such as authoring tools, user engagement, narrative structures, and transition techniques, as well as various access methods. While their analysis provided initial insights, some areas were only covered briefly. Our review seeks to build upon this by offering a more detailed examination of the impacts and implications in these underexplored areas.

More recently, Zhu et al. [227] published a survey on automatic infographic and visualization recommendations detailing how storytelling processes can be automated. Their work spans a range of related tools, models, and frameworks, from data-driven automatic visualizations and annotations to knowledge-based visualizations. We have integrated examples from their survey into our tools overview (see Fig. 6). Furthermore, Losev et al. [114] highlighted the need for leveraging diversity within the visualization community to foster new ideas and collaborations.

In the realm of journalism, Freixa Font et al. [67] provided a comprehensive overview of data-driven storytelling, underlining the importance of integrating interactivity and visualization. They discussed the potential of these techniques for engaging news readers, the tools and resources used in digital newsrooms, and the limitations of visual interactions. Similarly, Lopezosa et al. [113] argued for further analysis of data storytelling across different journalistic formats and reader perspectives.

Despite these significant contributions, a comprehensive summary of the literature across various application and research domains remains lacking. Our review addresses this gap by offering an interdisciplinary overview of data storytelling, drawing on perspectives from Behavioral science, HCI research, Marketing research, Data-Interaction research, and Communication Science. We discuss theoretical frameworks and foundations, narrative structures and their goals, and data story types across different methodological settings. Moreover, we provide a comprehensive overview of cognitive and non-cognitive effects and their evaluation methods.

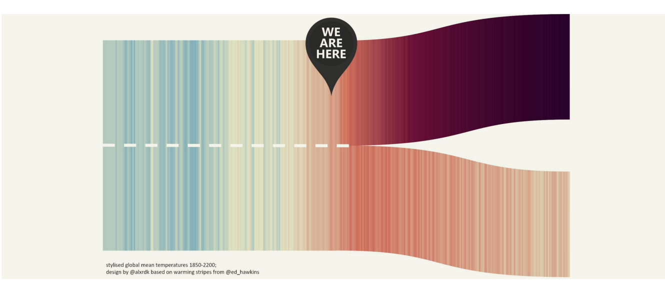

Fig. 1. The Warming Stripes alternatives by @alxrdk based on warming stripes from Ed Hawkins

Recent contributions to the field include the work of Aziz et al. [9], who explored the link between personality traits and user preferences for visual design styles in data storytelling. Lim et al. [111] discussed the educational potential of innovative data visualization in news journalism, particularly during the COVID-19 pandemic. Sanei et al. [170] used data storytelling to foster computational data literacy within the context of socioscientific issues like climate change. Shan et al. [178] proposed a design strategy for data storytelling in cultural heritage, examining data narrative from the perspectives of data science, visualization, and narratology. Lastly, Dailey et al. [51] conducted a survey on the use of narrative infographics by U.S. municipal governments, emphasizing their role in informing the public about various issues, particularly in the domain of public health and safety.

III. Method

Our research methodology follows a mixed-method review approach, aligning with established literature in the field to synthesize related findings comprehensively. We adopt the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [149] as our review protocol, which entails four main steps.

Firstly, we initiated the review process with a Scopus database search, targeting the subject areas "storytelling" and "visualization", restricting to papers published before February 2023. The query string used was `(TITLE-ABS-KEY (storytelling)) AND (visuali?ation)`. This initial phase yielded 1010 papers.

Secondly, each author received a random sample of 80 papers, with the task of mapping the core elements of their selection into a mind map. Collaboratively, we then identified individual categories and defined relevant inclusion and exclusion criteria. A custom-developed tool facilitated the process, enabling researchers to browse and tag individual papers with details such as source name, title, abstract, keywords, a "tf-idf" word cloud, and the publication year. Two authors reviewed each paper, categorizing it for inclusion or exclusion. In case of disagreement, a third author was involved in making a majority decision. Additionally, we manually searched for related work and incorporated additional papers into our sample ($n = 9$), leading to a total of 414 papers.

Thirdly, we assigned each category to two authors, who divided all papers within their category for an in-depth read. We corrected any misclassifications and cross-checked them. Publications were rated by relevance and excluded if the full text did not meet our criteria. We held weekly meetings to discuss preliminary conclusions and address papers relevant to multiple or different categories. This stage resulted in 184 remaining publications.

Lastly, in the fourth step, we organized the remaining papers into a coherent context and summarized them within the respective sections.

Three main exclusion criteria guided our review process. Firstly, we considered the relevance to the research question. Since we focus on data-driven stories, we excluded all papers that did not involve storytelling applied in the data context. This includes stories intended for entertainment or those visualizing qualitative data types in narrative visualizations. Secondly, we excluded papers that did not delve deeply into storytelling, such as those where storytelling was used descriptively or as a keyword. Lastly, we applied an incompatibility criterion, excluding papers not written in English or not published in peer-reviewed journals.

A. Definitions

Our systematic review includes only publications describing data-driven storytelling and related aspects per the following definition. The term "storytelling" has been broadly used in the visualization community without a universally accepted definition [90, 106, 116, 142, 194]. However, most definitions share a common trait of portraying a process or sequence of events. Therefore, we distill the existing definitions and **define a data-driven story as a series of related events in a (meaningful) context to facilitate understanding and decision-making concerning data**. To comprehend our perspective on storytelling and its effects, we first define some terms from narratology in relation to visualization and provide an illustrative example (see Fig. 1):

- **Story subject**: Describes what the data and story are about (e.g., in Fig. 1, this is the climate crisis).
- **Story object** (actor or carrier): The elements that depict the story (e.g., in our example, the average global temperature and the "We are here" sign).
- **Story events**: Individual arguments, data representations, or contextual information (e.g., in our example, we have two potential future story events with different outcomes).
- **Connections**: The link between the particular story events and how they are structured (e.g., in our example, the positioning of the "We are here" sign indicates the current position in time).
- **The audience**: The reader or target group of the story (e.g., in our example, people looking at the figure).
- **Effects**: Cognitive and non-cognitive effects concerning the story context (e.g., in our example, the storyteller aims to engage the audience emotionally and cognitively).

IV. Results

In this section, we summarize the results of our systematic review, following the four research questions we have outlined above.

A. Frameworks and design principles

Within this section, we take a closer look at data storytelling from a more holistic point of view and will try to answer questions like: What are the underlying mechanisms? How can the design process be conceptually/cognitively grounded? How can we structure the story creation process?

1) Fundamentals of storytelling: Storytelling is a method that creates a narrative context through a guided combination of explicit (e.g., the evil wolf in Little Red Riding Hood) and implicit knowledge (i.e., the wolf is treacherous, should not be trusted, and let into the house) [70]. As a result, story perception is an internal construction process whereby an internal model of the story is created. This construction process is influenced by the interaction of sensory impulses, their processing, internal knowledge, and associations. As a result, we need to distinguish between different dimensions that influence our understanding of a story: cognitive perception, narrative construction, and persuasive modeling.

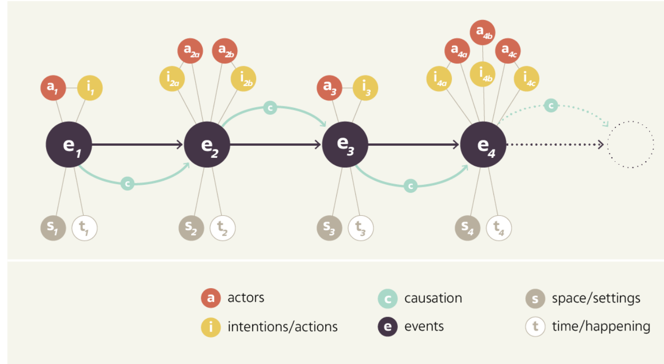

Fig. 2. Sequence of events with individual attributes adapted from Mayr and Windhager [132]

a) Cognitive perception: We know from visualization research that the cognitive processing of visual information depends on the type of information and how it is encoded visually. We look into multimodal information processing to understand how information is processed cognitively [27, 73, 132, 177] (see Fig. 3).

Multimodal cognition is based on the assumption of a dual-layer cognitive architecture. We process verbal (e.g., words or text) and visual information (e.g., figures) in parallel but independent of each other. After the stimuli representing an event "e" (see Fig. 2) is processed within the iconic memory, it gets semantically processed and organized to construct a situated model inside the working memory. This construction process also interacts with prior knowledge stored in the long-term memory, e.g., through supplementing information—whereby it functions as grammar [70] or contextualization through previous experiences (see Fig. 3).

Some approaches like annotations implicate that we perceive the verbal and visual channel sequentially, simply because we cannot look at both in parallel, Kong et al. [101] propose to visually encode the data and guide the user's attention with visual cues accompanied by an audio narrative to improve comprehension. Smith and Moore [186] propose a framework with a detailed guideline to design an auditory description display for complex interactive environments. This approach especially benefits users with visual impairments, providing access to all relevant information in a highly interactive manner and supporting user engagement. Other examples of using the verbal and visual channels in parallel are classical presentation scenarios, like Rosling's Gapminder talks [165] for a general audience. Here, Rosling directs the audience's attention visually to the relevant changes within the visualized data and explains the context verbally. The moderated form of data presentation can also support the communication between domain experts and non-domain experts in a health context [83].

b) Narrative construction: Mayr and Windhager [132] point out that narratives are not just a presentation format but "a fundamental way of organizing human experience and a tool for constructing models of reality" as they closely correspond to how we sequentially perceive the world around us. A narrative is considered a chain of related events in cognitive science whereby the individual events consist of several central data

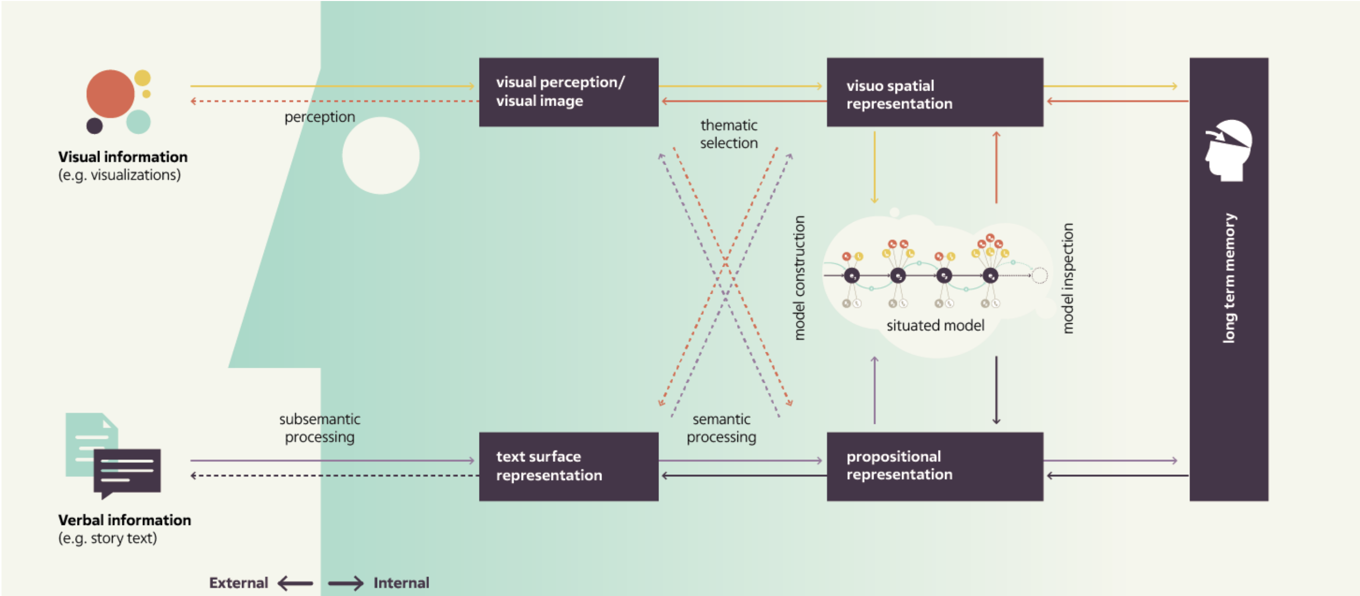

Fig. 3. Schematic showing processes involved in multimodal cognition adapted from Mayr and Windhager [132]

dimensions such as time, space, actors, etc. (see Fig. 2).

When we look at the connection between the event "e" (see Fig. 2) and how we interpret them, a critical component is the story schemata. They form an essential building block as they represent more generic concepts about the relationship of the events, related expectations, and related knowledge, such as when they function as grammar [70]. El Outa et al. [58] defines this into four layers: factual, intentional, structural, and presentational.

Schemata is an umbrella term for several knowledge structures like frames, plots, scenarios, and references. Describing every detail in an event is often unnecessary, as the audience's schemata can complete the meaning based on prior knowledge. Story schemata refer to the content of the entities and the connection between them, thereby influencing cognitive processing and understanding [132]. Led by the schemata, the recipient sequentially perceives events and the related dimensions and builds an internal representation of the story called the "situated model" [132]. The situated model evolves through continuously updated story events that are cognitively connected within the global internal story construct. While perceiving the narrative, the audience might engage with the story described as narrative immersion or transportation [21, 88]. Although immersion is a very elusive term, Isenberg et al. [88] pointed out that narrative immersion has an emotional as well as a cognitive dimension and distinguishes between(1) immersion in an absorption context, whereby the user engages with the story (emotionally), and (2) immersion in narrative/transportation context, whereby the user is getting connected with the story and the intended message. Narrative immersion is not a binary state; there are multiple levels of involvement and several influencing factors. Therefore, it remains an open question of how we can measure or control this effect to facilitate narrative visualizations.

c) Persuasive messaging: The basic idea of distinguishing between different elements in storytelling is not new. Some authors [21, 78] refer to Aristotle, who distinguishes between three relevant elements in storytelling: First, **Ethos** refers to the author's credibility and demands a clear and accurate presentation of the findings and transparency regarding the underlying data and analytical process that led to the conclusion. The relevance of a clear message contributing to the author's credibility was pointed to in different contexts [29, 73, 130]. Second, **Pathos** refers to the (emotional) relationship between the reader and the story. Due to the assumed asymmetry in literacy, prior knowledge, and perspective, it is proposed to consider the emotional connection with the topic to influence the accessibility and remove emotional barriers [21, 66, 89]. Pathos can also refer to the concept of narrative immersion [88] or the general principle of effective storytelling ("lure people in") mentioned by Groshans et al. [73]. Lan et al. [104] propose an animated design scheme and measured that it can convey positive affective emotions like amusement, surprise, tenderness, and excitement. Furthermore, the domain and visualization literacy of the audience needs to be considered within the story creation process. And third, **Logos** is about constructing the arguments, which demands considering the reader's perspective and carefully constructing the logical arguments. Other authors distinguish between the multiple aspects of storytelling: Figueiras [64] explored three dimensions that can facilitate storytelling: (1) context of the data, (2) emotional or empathetic dimension (the personal relationship with the topic, which can facilitate motivation and memorability), and (3) temporal structure of the event (which can be linear and non-linear).

Zhang et al. [221] present a framework for information unit-based data storytelling that combines multiple disciplines. The framework utilizes game development and machine learning methods to assist in the composition of data and story elements.

Several visualizations were analyzed to investigate subjectivity and identify specific techniques [34]. The authors point out that subjectivity is a controversial term within the visualization community, with objectivity as a core value. However, communication without subjectivity is hardly possible. A better understanding of subjective aspects can help us develop visual encodings that more closely reflect individual experiences and facilitate a broader view of personal perspectives from individuals. Lyu et al. [121] combine the Data-Insight-Knowledge-Wisdom method (a common method to explain human understanding in the perceptual and cognitive space) with six constitutive factors in Jakobson's communication model to create a framework for turning data into a story. Their research focuses on a theoretically grounded story creation process by splitting it into two parts: the creation of wisdom (analysis, translation, implementation) and the creation of a story (investigation, representation, development, implementation).

2) Structuring the storytelling process: The early work on defining storytelling within information visualization focused on categorizing formal aspects and definitions (e.g., genres) to create a comprehensive overview of the design space and related strategies [92, 176]. Based on this inclusive perspective, almost all visualizations would be considered narrative: we could even describe a line chart with a headline, sub-line, and annotations as a narrative visualization (annotated chart). Over time, the definition was narrowed down by emphasizing the sequential dimension (such as a set of story pieces). This delineated exploration from intention and the message of a visualization [21, 38, 102, 116, 162]. As a result, the storytelling process and its underlying mechanisms moved more into the foreground, calling for a reflection of the process, related tools, and algorithms, as can be seen in the work of Chotisarn et al. [41] and Shi et al. [179].

a) Focus on the Data Exploration: In many cases, the structure of the working process was initially derived from the context itself (mostly in data journalism or visual analytics), whereby exploration was an integral part of the process [80, 119, 189, 199]. For example, one of the **critical elements of exploration in storytelling is identifying and selecting key data features for presentation** (e.g., through snapshots) and emphasizing their contextual meaning (e.g., through captions or annotations) [80, 91, 92, 119, 199] or re-ordering [197]. As a result of this, the core statements are easier to understand in the resulting presentation scenario. The viewer does not have to go through all the necessary steps to identify the key aspects. To facilitate the reproducibility of the discovery process, Gratzl et al. [72] propose the "CLUE" approach—an interactive provenance graph that indexes and displays the relevant steps of the exploration process. The resulting transparency of the process directly contributes to the author's credibility. However, a comprehensive understanding of the exploration process usually requires specific domain knowledge, and therefore, this approach might not be applicable within scenarios with a general audience. The authors found this particularly the case in business analytics, where users have a task to bridge the gap between raw data and business insights to become a data-driven organization. To create a data-driven organization, Boldosova and Luoto [18] propose a conceptual framework with propositions about the relationship between business analytics, data-driven storytelling, and the intention to use business analytics. They state that the intention to use business analytics regularly depends on users' positive or negative experiences while interacting with it and that storytelling contributes to a positive experience. The difference between traditional data interpretation and data interpretation supported by data-driven storytelling is thus reflected in the user's experience while interpreting data and making decisions. Marjanovic [128], on the other hand, suggests a more pragmatic approach to increase visual analytics skills within the business. A good narrative for the data interpretation journey can help to accomplish that task. While Boldosova and Luoto [18] recommend that the key to implementing storytelling within business analytics is behavioral change.

Marjanovic [128] focuses more on empowering business users to develop visual data exploration skills. Using visual stories as boundary objects among primary (developers) and secondary designers (users of visual analytics). Boundary objects need to be co-created rather than exchanged.

In line with this, Minelli et al. [136] proposed visualizing development sessions to support understanding developer behavior. All authors require users to be involved in the process of creating stories and require them to be inspired by those stories. Ya'acob et al. [214] take another perspective and propose a framework to facilitate business decisions based on analytical reasoning features from three parts of visual analytics representation: higher-level structure, interconnection, and lower-level structure.

Considering multiple provenance paths in parallel can also enhance understanding in various contexts. Diamond et al. [56] developed a dashboard for data storytelling for the cultural sector where provenance was used to identify data sources. The framework used for this dashboard was the four editorial layers defined by Hullman and Diakopoulos [84]. Yousuf and Conlan [215] proposed a framework to increase student engagement where a visualization system uses storytelling to present complex data. The main focus of the framework is to construct narratives with multiple exploration paths that are personalized for individual end-users, primarily to support weak students. The authors found that personalized visual narratives facilitate understanding and engagement. Similarly, Park et al. [150] propose a system called "Storyfacets". The system provides different views of the same analysis session to support provenance exploration to a different audience—experts, managers, and laypersons.

b) Focus on mass communication: Lopezosa et al. [113] reviewed the current developments, evolving technologies, and challenges in the field of data journalism. Building on the foundations laid by the previous models within the field, Lee et al. [106] proposed a high-level overview: the visual data storytelling process, which integrated all relevant steps (see Fig. 4) to communicate data using storytelling effectively. In this context, the individual stages of the process, especially exploration and story creation, were considered iterative and not necessarily linear processes where external factors like audience, setting, and medium influence all steps. The authors raise an essential concern that the transformation of data to an understandable format might also result in a (un)intended misuse, which leads to relevant ethical questions, particularly in connection with transparency concerning the exploration process, related choices, and the underlying data.

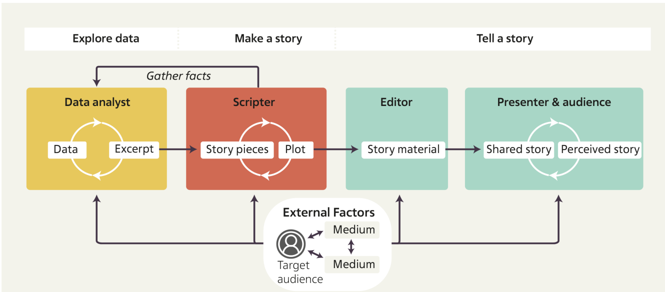

Fig. 4. Sequence of events with individual attributes

Later, this perspective was extended by identifying seven key features of data-driven storytelling within data journalism, thereby offering a more nuanced perspective on the multifaceted role of data in both analytical and transparency/ethical context [207] as well as the discussion of narrative techniques and interaction [65].

Examining the use of narrative info-graphics by the United States municipal government for public information dissemination, Dailey et al. [51] summarize the design elements of the narrative info-graphics instrumental in fostering transparent communication. These elements included: *use of clear and concise language, the use of visuals to support the narrative, the use of color to highlight key information, the use of annotations to emphasize key observations, and the use of subheadings to organize information.* Shi et al. [181] overviews six design patterns related to user input and interaction to facilitate engagement, information recall, and subjective connection to the data story. Another illustrative example of innovative public data presentation is described by de Martino et al. [54]. Grounded in the principles of accessibility and clarity in the context of European statistics, the authors used information and communication technologies to provide new opportunities for disseminating statistical data. The information is described using interactive graphics, sector-specific glossaries, references to publications, and other auxiliary resources. Another upcoming domain is the field of legal design, aiming to utilize data storytelling to facilitate interaction and understanding with legal information [155].

Current developments in open data lead to new challenges. Janowski et al. [90] proposed several data story patterns that can enhance the discovery process of open data. Brolcháin et al. [26] investigated the data communication process in an open data context. The authors stated that **open data platforms should provide storytelling features** to enable users to find and present insights within the data. The authors propose a framework (YourDataStories-YDS platform) that consists of five stages to create a data story that should facilitate users to understand and get to Open Data. First, **discovery** during which the data set is explored and traversed, including information on metadata, completeness, and highlighting outliers and anomalies. Second, **assistance**, during which the data is cleaned and processed. Third, **insight**, which is considered the most important phase, consisting of two main features: explanatory features (such as availability of visualizations and automatic narrative generation) and social features to gain more information to read the dataset (such as engaging with data-owners and discussing with other dataset users). Fourth, **leverage**, during which conclusions from the analysis are shared and discussed also to estimate validity. Fifth, **trust** including transparency that privacy and other rights are respected.

c) Focus on Scientific communication.: Typically, knowledge dissemination in science is a different role of the data exploration process and involves a specific perspective of the (target) audience [21, 29] as well as design strategies [159]. Analytics and exploration are part of the research that led to the knowledge that needs to be disseminated. Therefore, Botsis et al. [21] highlighted the importance of separating exploration and story creation to avoid the risk of biases. These can be influenced by the needs of the narrative or the other way around, resulting in inefficient stories. The major challenge in this application domain is a general asymmetry of knowledge and (domain) literacy between author and reader. Inspired by the general concept of Aristotle's ethos, pathos, and logos, the authors proposed to start with the target audience and understand the user by analyzing (domain) literacy and interviews, defining goals and sub-goals to support the clarity of the created story. To facilitate the implications of the audience towards design choices, they adapted Cairo's visualization wheel and combined it with the Newest Vital Sign (NVS) score to help story creators with design choices (see Fig. 5).

To provide an illustrative example for this concept, Fernandez Nieto et al. [62] point out that traditional learning analytics dashboards pose a significant challenge in interpretation due to the often limited data literacy of their primary users, which includes teachers and students. To address this issue, they propose using alternative ways to communicate data insights using visual narrative interfaces. Rickhaus [159] emphasized how visualizations enhance scientific storytelling and lateral thinking, helping the readers understand complex information. He also recognizes the effective use of visual elements, such as color, shape, and layout, as strategic tools for emphasizing critical data points and directing the reader's focus toward key information. Consequently improving the reader's engagement. Burkhard et al. [29] introduced a visual storytelling design board to align design choices with literacy considerations. They integrated the external factors as a dedicated stage within their process. Data and conclusions already existed in their scenario as the pre-existing research findings. They created a user scenario narrative to understand better the individual steps, narrative structure, audience's perspective, and related expectations. Subsequently, the resulting findings were transformed into mock-ups for further iterations together with users through visual and interactive prototyping [29, 168].

d) Focus on Collaboration: The story creation process can also directly involve users and their individual views on the data. Collaboration can involve human-machine interaction supporting the collaborative editing process [191]. Several authors proposed to design the exploration and story creation phase as a collaborative process of multiple users [8, 94, 95], while modeling personal family migration narratives with data. In user studies, they analyzed the behavior of several family

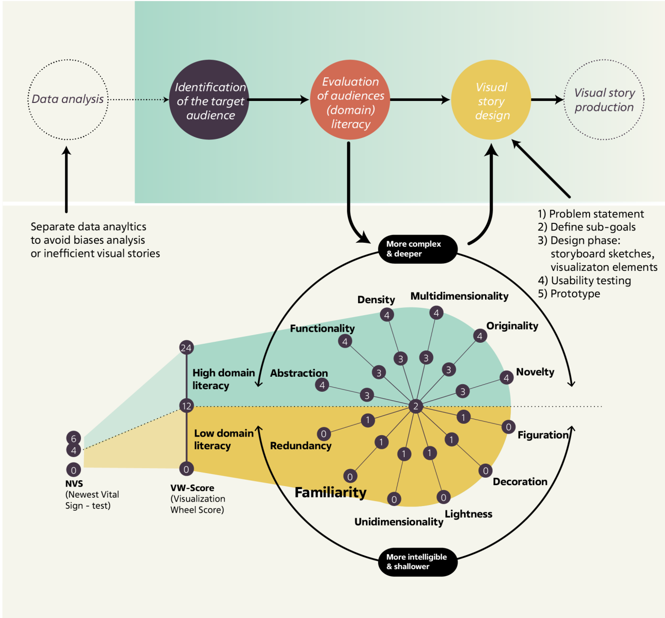

Fig. 5. Story creation process within knowledge dissemination according to Botsis et al. [21]

groups and identified ten different socio-technical practices of data wrangling. They developed a conceptual model of the data wrangling process, differentiating the individual and global/data perspective.

e) Focus on the user's conceptual model of time and space.: Geospatial visualization has a long tradition of developing visualizations in a spatiotemporal context, focusing on the idea of incorporating the space and time dimensions into narrative visualizations. The commonly used phrase in storytelling—"Once upon a time in a land far away"—elucidates the relationship between the dimensions of time and space and storytelling [161]. Rodrigues et al. [161] emphasize the aspect of employing spatio-temporal dimensions when creating interactive visualizations by providing a set of guidelines. Agreeing with the concept of the connection of narratives and spatial dimension resulting in spatial identities that, in turn, contribute to the shaping of the story, Caquard and Fiset [33] designed an application to explore the geographic structure of the story and to perceive the impact of stories in the creation of places. They achieved this by developing a map of contemporary Canadian films by characterizing the spatio-temporal dimensions of narratives. The dimension of spatio-temporal narratives leverages strongly in geographic visualization systems. In describing geo-data, the spatial dimension can represent the spatial relationship between the data entities, and the temporal dimension would illustrate the changes in the data over time. The additional information about the data could then be represented in a thematic dimension [193]. Thöny et al. [193] explore storytelling techniques in 3D geographic visualizations to help conceptualize geospatial data over time. Visualizing geodata in a three-dimensional manner provides the opportunity to make it interactive and present the information in a captivating and intuitive way. This way of presenting geo data motivates the otherwise overwhelmed user to explore the relevant information more effectively. They discuss using various components in the 3D maps that help tell data stories. Story maps are built by considering whether the data belongs to spatial, temporal, or thematic dimensions. For data storytelling, storylines follow a theme or a person from a data-related perspective. Geographical 3D maps have exploited the use of storylines to provide a better immersive experience of the story to the users. Storyboards and scene components are the remaining essential elements that Thöny et al. [193] discuss, which can be used to create interactive stories. Another example where location mapping has been used for creating effective visualizations is the work done by Chaudhary and Arora [36], where they use geographical visualizations to present the citizen complaints in India to the responsible authorities to help in decision making. Along the same lines, Lan et al. [103] introduce a new web mapping platform where users can tailor their own story maps and effectively identify geographic patterns of social and economic phenomena through narrative mapping. This would prove to be a great tool for social scientists and policymakers

3) Automated systems to structure the storytelling process:

a) Automated visualization recommendations and reasoning.: The evolution of data-driven storytelling has generated a demand for automated tools to assist story creation. Upcoming genres have led to the development of different storytelling forms, including DataToon [99] and Timeline Storyteller [25] and others designed to effectively communicate data or visualization and videos [5, 98].

In the field of automated systems specifically for visualizing data in quantitative research, we found one paper in our survey by Choudhry et al. [42] who presented CAUSEWORKS, which is a textual narrative system that uses visualization and text to describe causal data. The system provided narratives summarizing data changes and projecting trends. Although we found many works focusing on specific types of data based on the domain it was collected from, for example, Chotisarn et al. [40] developed a prototype to support bubble chart animations based on Twitter data. The tool automatically creates animations and facilitates authoring through captions and filtering functionalities. Another example of this type of data presentation is Rosling's presentation of economic data [165]. This inspired Shin et al. [182] to design a semi-automatic storytelling system for data presentation called Roslingifier. The system provides three views supporting data presentation with auto-detected events based on storytelling techniques such as natural language narratives, visual effects highlighting events, and temporal branching. To summarize time-varying data in a comprehensive narrative, Bryan et al. [28] propose "Temporal Summary Images" (TSI), that automatically identify points of interest through computing noteworthy changes (e.g., strong increase/decrease) to recommend potentially relevant data features.

We also reviewed several works on data videos. For example, Shi et al. [180] propose Autoclips for automatically generating data videos when a sequence of data facts is given as input. They construct an algorithm that creates videos that have comparable quality to human-made videos. The work of Shu et al. [184] adds Word Clouds as an authoring tool that interactively crafts word clouds and animations to generate storytelling videos.

b) Automated sequencing.: Along with efficient visual encoding and identifying key aspects, ordering events is crucial in developing data-driven stories. The section narrative structures (see section IV-B) elaborates extensively on theoretical aspects of sequences and narrative structures in general. In comparison, this part gives an overview of sequencing support tools. Hullman et al. [87] have significantly contributed to the field by researching automatic sequencing in narrative visualization. The underlying idea is to look at a story as a series of views connected by transitions. The authors incorporated the transition costs and a global weighting to calculate the most efficient transition type. As a result, the GraphScape method was created to illustrate the theory [100]. Later, this approach was combined with visual encoding support. Obie et al. [141, 145] proposed a framework for logically ordered data stories and the related tool "gravity"—built on Vega lite for visualization support and recommendation for effective narrative ordering based on Graphscape supporting collaboration and presentation [143]. The DataToon tool [99] also suggests automatic transitions and panel recommendations for dynamic network data comics, which support narrative ideation and storyboarding. Another approach by Walsh et al. [201] used an algorithmic method to find stories in large data lines. This technique is dynamic by allowing for displaying and collapsing the timeline. Focusing on the provenance of the exploration process, some authors [12, 32] created a machine-learning model to develop storylines for data visualization users. The model learns from past user exploration of the data set to build a preferable storyline. For example, the algorithms create a drill-down sequence based on the user's exploration habits [32]. An archival library study took a similar approach by adapting the narrative sequence to the user input [13]. Looking at the individual story nodes, Zanda et al. [217] used the Story Network Principle to generate a tree of story paths for data visualization. Overall, these techniques have been shown useful as they allow for customization to individual user preferences.

4) Tools to create data stories: Tools may help the user by choosing efficient visual encodings, by choosing accurate contextualization, or by simplifying the process of presentation. In the following, we present the three categories of tools that we found in the articles we reviewed (see Fig. 6).

The first set of tools generally tries to help users **explore and present** their data. As one example, Roels et al. [163] proposed a set of general requirements and introduced MindXpres, a presentation platform that combines interactive data exploration with storytelling. To help the general audience gain insights and have a better understanding of time-series data, Lu et al. [115] proposed an interactive authoring tool that converts 2-D time series to data videos. To detect changes in time-series data and illustrate them in video sequences, the authors employ several algorithms to complete tasks that otherwise would have required much effort if done manually. Another category includes tools that help users annotate existing data visualizations and present them as visual data stories on the web. Typical tools from this category allow importing existing visualizations and help the user to design scenes and annotations

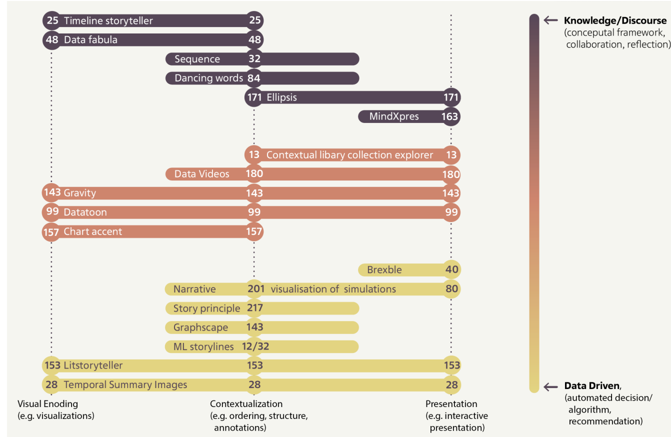

Fig. 6. Story Generation Tools: the number refers to the reference number

for parts of the story. Toolkits come in a variety of feature sets. Tools like ChartAccent provide simple tools to annotate figures from data [157]. Tools like Ellipsis [171] additionally provide multiple tools to generate different narrative structures, and tools like Flourish provide fully-fledged data exploration and visualization software with a focus on data-driven storytelling. The last category contains tools designed to automate the full process of data-driven storytelling, typically with the help of visualization. Tools in this category use various algorithms to find meaning in the data (e.g., unsupervised machine learning) and automatically pick or recommend ideal narrative structures for the given data. As an example Cruz and Machado [48] have developed a conceptual framework for telling a story from data, introducing Data Fabulas, "the set of events, agents, actions and chronology extracted from a dataset" with the idea of enhancing cognition about information. Ping and Chen [153] proposed an interactive system that helps researchers understand a field of research from the scientific literature. Their system is designed to be used as supplementary information for systematic reviews to not only grasp the overall research trends in a scientific domain but also get down to research details embedded in a collection of core papers. Their tool supports interactive visual storytelling at multiple levels. It allows for answering various research questions, covering macro-scale and micro-level questions. As entities of investigation, the authors use concepts or terminologies gathered by employing various text-mining methods and then map them to basic visual elements to form visual storylines. We have visualized tool examples in Fig. 6.

B. Narrative structures

To further understand how we can utilize storytelling to enhance our understanding of quantitative information, we first need to understand the building blocks of a story. The core components of storytelling are the collection of events and their temporal and causal relationship that we call narrative structure. Relevant questions are: How is the information organized? What different narrative structures are there, and what goals are pursued by them?

In traditional narratology, **a narrative has two components** (see Fig. 7): **the story** (relating to the content, what is told?) **and the discourse** (relating to the expression, how is it told?) [35]. While these refer to texts or films, visualizations are based on a different narrative structure, namely the space in which the story can be presented [45]. A fundamental difference lies in the additional dimension created by the data representation and has implications for layout and the temporal structure [25].

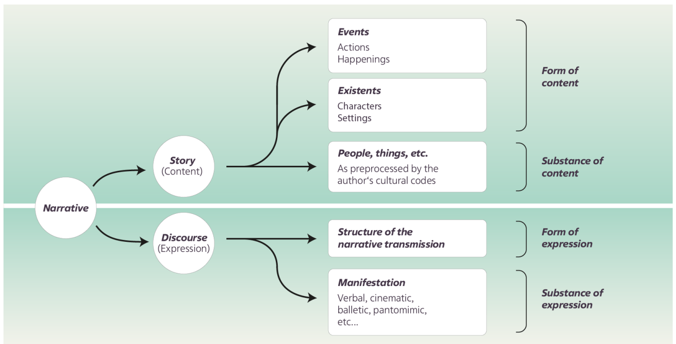

Fig. 7. Narrative structures according to Chatman [122]

1) Existents: If we look at the formal content, we can consider a story a collection of existents and events. Existents refer to actors/characters and settings, while events include multiple dynamic entities such as actions and happenings that shape the causal relationships [122].

a) Characters: Characters are a fundamental building block of classical stories, but how do they behave when data plays a central role? Multiple articles have mentioned the need for a sequence with a clear protagonist and other characters for data stories via data visualization [6, 123, 139, 166, 224] or data videos [82]. The data in data visualization can be presented as characters in a story [166]. Additionally, a character can be separated from the data by presenting an animated agent with audio [79] or without audio [17]. When the data stories include personal history data, the user him-/herself can also be seen as the character or player [96]. The user can also become an actor in the story through interactivity, allowing the user to steer the story [123]. Roth [166] indicates that data stories in geography can have a protagonist and an antagonist. According to the characters, different structures for the story can be shaped. Similarly, Bounegru et al. [22] illustrated five narrative patterns in network visualizations which were created around characters, either single actors, multiple actors, entire organizations, or the association/opposition of multiple characters. The two articles illustrate a rare example of how traditional storytelling with the use of characters is applied to data visualization. While articles mention the use of characters, there is no consistent notion of what is and what is not a character, nor how they can be used. The knowledge is due to a limited academic review of used characters in data stories by practitioners.

b) Settings: Data-driven storytelling settings can be used to place the visualized data into a particular spatial and temporal context. The context can assist in characterizing a story to create a purpose [79]. It can be set by the story itself or by the environment of the user [96]. This can be realized in multiple ways, e.g., for timeline data visualization, the scale can be used and manipulated to place the data visualization in a context [25]. The path of the timeline can also be formed and shaped to represent the story's setting. The use of maps is an example of spatial settings [166].

2) Events and sequence of events: A storyline consists of a sequence of events. The basis for this is establishing a logical connection that enables the user to interpret the individual events as a coherent overall plot. During a story, multiple storylines can emerge in parallel and work up to a conclusion, where all the storylines of the plot are drawn together. This can also relate to the message [106].

a) Structure: In data storytelling, the terms exploratory and communicative (or explanatory) are taken as the general way to structure a story [162]. However, it can be argued whether the two types are purposes instead of narrative structures. A clearer insight into narrative data story structures is given by Segel and Heer [176] who argue that there are three different structures: linear, exploratory, and a combination thereof [176]. The latter refers to hybrid linear-nonlinear patterns, for example, a Martini Glass structure where the story starts explanatory and afterwards opens into an exploratory panel. The three structures have been the basis of many data storytelling studies, such as those by Weber et al. [207] and Rodríguez et al. [162]. Other studies have shared similar findings stating that data stories can be structured as guided, guided exploratory and exploratory-viewing [202] or information-seeking, comparative, and iterative for scientific storytelling [123]. Parallelism can be seen as a fourth structure [87, 194]. The structure tells the story with repetitive sequences. The parallel structure is explored for timeline visualizations by Brehmer et al. [25] who state that the stories can be unified, faceted, segmented, as well as faceted and segmented.

Although the general structures give little insight into the sequence of the story, commonly data stories use temporal or chronological sequences [79]. The plot of a story with a (partial) explanatory structure can be told in a different sequence. Besides, the three-act structure is referred to create data stories [46, 166]. The three acts include various narrative elements such as setting, hook, and plot twist [79, 166]. Lidskog et al. [108] frame the three acts in the historical account (the what, the why) and what to do about it. The narrative moves beyond a 'story with a morale' by inciting action on the audience's side. Often, these narratives can be taken apart further, as illustrated by Ma et al. [123], who states that traditional stories develop as follows: know the audience and their knowledge level, set the stage, character introduction, develop plot, show relevant story, and implication for the reader. Laurel refined Freytag's triangle, which is referred to by Thöny et al. [193]. The paper states that tension is created by the phases of exposition, (trigger) incident, critical action, and crisis to a climax. Then, the (re)solution occurs, and the plot ends in a relaxation phase.

b) Elements: The central element of the story-narrative should be the problem from which the setting, purpose, agent, acts/events, and means/helpers depart. Arevalo et al. [6] follow the theory of D. Jones and Anderson Crow [50] and Murray and

Sools [138]. Their papers adapted the central elements from ElShafie [59]: connecting with your audience, raising problem awareness, relating to a practitioner's world, acknowledging remaining challenges for practice, and giving a take-home message. Amini et al. [4], Steinert et al. [190], Yang et al. [213], expands further on the theory by Freytag [4, 68, 193] as adjusted for visual narratives by Cohn [45] by examining dramatic structures for data videos and proposing guidelines [213] Cohn [45]'s definitions for **the four story blocks are Establisher, Initial, Peak, and Release** (E, I, P, R). Patterns can be examined in data stories to create narrative categories by the following method [Element+] where Element is one of E, I, P, R and the "+" sign indicates repetition of the preceding element. According to Amini et al. [4] "E+I+PR+" pattern was the most common category, as well as "E+I+P" and its subset pattern of "EIP." So far, the study by Roth [166] is the only study that has mapped several sequences of events that reflect traditional story plots. Roth [166] states that for geographical visualization, eight different structures can be used: destruction, genesis, emergence, metamorphosis, cause & effect, convergence, divergence, and oscillation. These are closely related to the development of the different characters in the stories. Nardi [139] describes the approach of re-storying through which a story is constructed from original pieces of data involving the creation of characters, settings, and events. Here, it is not solely recounting a sequence of events, but re-storying is about composing a new story through evaluating and interpreting.

3) Representation of a story: Behera and Swain [14] briefly touch on the potential for data stories to differentiate themselves in their stories' purpose and the targeted audiences. Other articles agree that narrowing down the audience for the story is important [6]. However, they fall short of explaining different audience groups for data stories. Regarding the purpose of the data story, it is often the story's topic or main message, which can be explained in a few sentences [96]. Weber et al. [207] states that three purposes for data stories exist: to tell, explain, or argue visually. The article bases the definition on theoretical considerations of traditional narratives.

Ojo and Heravi [147] took a different approach and defined seven types of data stories after examining 44 winning projects, namely reveal information of personal interest, enable a deeper understanding of a phenomenon, reveal anomalies and deficiencies in systems, track changes in systems, refute claims, reveal information about an entity in increasing levels of details, and reveal unintended consequences. Similarly, Bounegru et al. [22] states five narrative view types: Exploring associations around single actors, detecting key players, mapping alliances and oppositions, exploring the evolution of associations over time, and revealing hidden ties.

a) Narrator: The narrator of the data storytelling influences the narrative structure. Three types of narrators have currently been reviewed in data storytelling: the designer, the user, and the visualized character. A few papers have studied the narrators [166, 223]. The user can guide the story in interactive visualizations [123]. Heyer et al. [76] took it further and allowed the user to act as the narrator by stating prior beliefs. Designers are narrators by framing the story [49, 84] such as using annotations [49]. Lastly, the data visualization itself can be the narrator by using voice narrations [17] or animated agents that narrate through voice [79]. Hullman and Diakopoulos [84] illustrate various design rhetoric and viewing codes that the designer can use to emphasize a specific message or guide the structure of the story. Moreover, the presentation structure of data videos is another influence the narrator can have [82].

b) Genres: In chapter IV-C4, the data visualization types (or genres) for storytelling have been discussed. As aforementioned, the prominent research of Segel and Heer [176] has mentioned key genres that have been saturated by other authors, pointing out new narrative visualization genres. Similarly, Roth [166] points out seven genres (static visual, long-form infographic, dynamic slideshow, narrated animations, multimedia, personalized story map, and compilations) specifically for geographic maps. However, it is non-conclusive if these time and geographic layouts can be used for other visualization types. Michel and Ladd [135] discussed a multimedia long format through which the narrative can continuously expand by adding new exhibits to the structures. At the same time, the genre uses specific codes that allow customizing the narrative to different formats and media. Similarly, Seyser and Zeiller [177] coined the term *Scrollytelling* adapted from the long-form articles used in journalism. Again, the authors highlighted the vertical aspect along with multi-modality. As opposed to Michel and Ladd [135], the *Scrollytelling* genre does have a definite endpoint. Majooni et al. [126] took a different perspective and analyzed the performance of different layout designs. The natural left-right (up-down) eye movement could improve comprehension of infographics and data visualization. Other research has applied these types of genres, such as Metoyer et al. [133] who took the drill-down genre of Segel and Heer [176]. In addition, the medium can also be seen as the structure of the story in the spatial layout. Zhao et al. [224] illustrated that data visualization stories could be partitioned and sequenced into data comics to create a meaningful order in the visual space. Common data stories such as long formats or slideshows are designed vertically or horizontally in the visual space. On the other hand, Brehmer et al. [25] constructs four genres for timeline visualizations: linear, radial, grid, spiral, and arbitrary. These visual flows can help outline the story structures for data visualization.

The genre of a narrative gives structures to the overall type of story. The term genre would refer to romance, action, comedy, or documentary in classic movies or story books. Data storytelling uses similar genres such as scientific [108, 123], geographical [166], and biography data storytelling [25]. Hook [82] found that nine genres were used for data videos: advertisement, comedy, cookery, documentary, drama, horror, music, science fiction, and unclassified.

4) Combination of the discourse and story: A few papers highlight the combination of narrative structure techniques used for the discourse and story for the following genres: timeline visualization [25], interactive data videos [82], spatio-temporal visualizations [161], and geographical visualizations [166]. Other research has looked into visualizing story plots [156, 209]. For example, Qiang and Bingjie [156] visualized the story structure of the movie Inception, which has a multitude of narrative layers. In Human-Computer Interaction, the generation of visual storylines is closely related to computer games and educational applications. This type of research has focused on creating story paths for interactive environments. Accordingly, algorithms generated relationships between entities and visualized the different storylines for one narrative in a tree-like graph [49, 160]. The underlying story path structure is often based on established theories such as Rhetorical Structure Theory [49].

C. Data story types

In this section, we propose a classification scheme for individual framework types to help future researchers within their respective research domains:

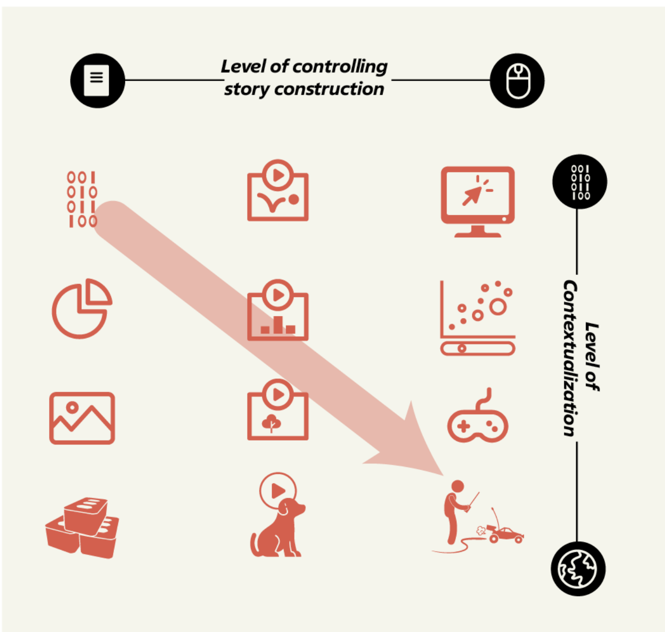

Fig. 8. Overview of the individual level of contextualization from raw data on the top left to the real or virtual world on the bottom right.

Within the conceptualization process of a story, we can distinguish between two dimensions: the level of contextualization and the level of control within the story construction. The first one refers to the semiotic distance of the visual language utilized to represent the data through perceivable means, while the latter one refers to how the temporal and causal structure of the story are constructed by the user.

Contextualization helps to understand what the data stands for (the meaning), while story control can facilitate understanding the underlying concepts and relationships (the related concept). While the author-driven and reader-driven approach introduced by Segel and Heer [176] describes data storytelling from the perspective of information visualization, we describe data storytelling from an interdisciplinary perspective of human data interaction, narratology, and psychology. Here, we look at the amount of contextual information and the level of control and how they influence the internal construction process of the final data story. We view the information presented in a data story as a combination of implicit and explicit knowledge aiming to explain the meaning of the data, where the level of contextualization plays a major role. Depending on the amount of contextual information in the data story, we categorize the data stories into four categories: stories presenting the literal information in the data, stories that use visualizations that represent the data along with textual explanations, stories that use metaphors representing the contextual information along with the information about the data and finally physically or virtually representing the data in a real/virtual world, where the contextual information is highlighted even more.

As described in the previous section IV-A1b, a story is constructed as a combination of events with several attributes that are temporally and causally connected. Controlling the order of events and other attributes can impact the story. Hence the level of control can be determined based on how many attributes are controlled directly and indirectly and how they are presented temporally. Looking at the users' perception, how story construction is controlled is another dimension that can generally be designed for any level of contextualization. While perceiving a narrative construct, we distinguish between three individual levels: **a static presentation**, where the user can experience and explore the information solely with his eyes (e.g., an infographic, a data comic, or an annotated chart) **a dynamic presentation** where the temporal structure of how the individual story blocks are perceived is controlled (e.g., a data video or animation) **an interactive presentation** There, the temporal structure and possibly also the content can be manipulated by the user (e.g., an interactive data story, see Fig. 8) This section gives an overview of the used visualization and data types within specific categorization schemes.

1) Low Context - Verbatim:

a) Annotated charts: Annotated charts combine visualizations by additional forms of communication to emphasize and explain specific parts of the dataset. Textual and graphical annotations are often used in visualizations to highlight points of interest and draw readers' attention to them. Examples of these additional forms have been given by Kosara and Mackinlay [102], such as written text, audio, video, or links to more information. They also point out that guidance of the reader can be done via visual cues like arrows or other methods. Satyanarayan and Heer [171] indicate this as being base-level annotations. Although annotations are often positioned statically, annotations can be made dynamic by binding them to data points [105, 171]. Annotated charts seem to be well connected by creating empathy, derived from Peng [151] and Figueiras [64]. We also found that annotated charts are used above average for enhancing time series data.

Interactivity, in general, can enhance the contextual dimension of data [63, 91]. By allowing the users to interact with the data and visualizations, interactions provide a sense of control to the users to explore the data and thereby understand the context better. At a basic level, interactions like slide transitions, click and zoom, or context menus are used to provide additional information on the data.

2) Medium Context - Narrative Visualization:

a) Infographics: Infographics, or information graphics, combine graphics, images, and text in one visualization, are seen as an efficient way to communicate complex data, information, or knowledge [177]. Since Charles Minard's work in 1858 and 1869, infographics have been widely used. Jacob [89] report an increase in the use of infographics in journalism since the 1980s, as they often have replaced photographs. In the analyzed literature, there have been several categorizations of infographics. Otten et al. [148] have defined three main categories (data graphics, maps, and diagrams). In comparison, Albers [2] recognizes four categories (simple infographic, snapshot with graphic needs, complex infographic, and information flow/process). Seyser and Zeiller [177] argue that because of the strength in visualizing complex data, infographics also form the basis of Scrollytelling, which we discuss in this article as interactive visualization.

b) Posters.: In the articles we reviewed, posters were hardly used. Bryan et al. [28] investigated creating posters using time-varying data to identify and visualize points of interest and create data stories from them. He combined a time-based visualization extended with comic strip-style data snapshots of relevant steps and annotations in his result.

c) Slideshows: Kosara and Mackinlay [102] describe slideshows as an efficient way of storytelling to visualize and explain data, not to analyze, whereby the data is presented in a linear, controlled way. Their primary focus is on slideshows for large audiences. They argue that this form of storytelling is missing interaction, considered one of visualization's most important aspects. They also stress that interaction facilitates faster and more practical data analysis because of the reader's ability to change the view and potentially add different data quickly. This would define slideshows (and data-videos) as being author-driven rather than reader-driven [3]. Schroeder et al. proposed automated slideshows to explain financial data [174] utilizing abstract [172] and metaphorical [173] representations. However, Segel and Heer [176] emphasize that slideshows work well to make complex datasets and narratives accessible.

d) Interactive visualizations: Interactive visualizations recently became increasingly popular, especially in journalism [69]. The concept of interaction (as applied in computer games, for example) differs significantly from what we associate with a classical narrative, which is traditionally told without interaction, which helps to keep focus in a storyline [102]. A non-linear interactive data story, on the other hand, is characterized by a stronger focus on free exploration and has been enabled by open and accessible large-scale datasets and interactive data visualization tools. Interactive non-linear narratives seem to be supported by the increasingly growing role of data and the growing preference for using data visualizations in the public news media as these tools are powerful for enriching narratives about data-related topics [176]. Terms often used in interactive non-linear contexts are data wrangling and collaborative storytelling. Jiang and Kahn [95] examined the learning opportunities related to data wrangling practices, such as when social data scientists wrangled data and created models to explain social changes. Wong and Lee [210] make a connection with collaborative storytelling to emphasize that storytelling is being used for fostering creativity, using common techniques like brainstorming, elaboration, and associative thinking. In alignment with these principles, Hasan et al. [74] employ interactive data comics as a means to enhance the co-design process. Then, dynamic interactive narratives change the perspective of the user from being a passive (reader) into an active (meaning-maker and actor). By adding user experiences into the process of transforming author-driven narratives to reader-driven, a story can also be structured by dynamic factors such as the data or individual characteristics of the user. Obie et al. [142] conducted a confirmatory user study to compare author-driven narratives based on presentation videos in relation to interactive visualizations in terms of comprehension and memorability. They concluded that users prefer interactive visualizations based on accuracy but found the data in the presentation videos is easier to understand. In this context, Shi et al. [181] identify six Breaking The Fourth Wall (BTFW) design patterns, a technique of integrating interaction with narrative to increase reader engagement. Upon conducting a user study to assess their benefits and concerns, the authors concluded that BTFW interaction improved self-story connection, user engagement and information recall, while raising concerns about balance, privacy, and learning.

3) Medium/High Context - Metaphorical Visualization:

a) Data comics.: Data comics are an upcoming genre within data-driven storytelling. They combine the freedom of 2D spatial layout presented in infographics and annotated charts with the linearity of narration found in videos and live presentations, enabling readers to consume the story at their own pace [11] and providing high interactivity and interoperability [43, 205]. Although the term comic strip seems to exclude scientific applications, this genre has been used for uncovering dynamic networks [10], trying to model business processes by key stakeholders [185], statistics [10], physics [15], and data clips [10]. Simões et al. [185] found that comic strips are handy for capturing tacit process knowledge but less effective in getting a complete process overview. On the other hand,Bach et al. [10] found that graph comics can improve understanding of complex temporal changes. Zhao et al. [224] declare data comics to be more engaging, space-efficient, faster, and easier compared to infographics. They concluded that the comic layout helps readers view the set of visualizations as a whole story without explicit instruction. Although we found that data comics are also used for exploratory purposes, the comic genre was largely used for explanatory purposes. Wang et al. [203] presented a successful and reproducible case study of how to teach data comics workshops to interdisciplinary students.

b) Storyline visualizations: Watson et al. [206] mention that visualization techniques on storyline visualization platforms are often used to attract new ways of human interaction. Our starting point to look into this interaction was the exploratory empirical study about users' reception and usage behavior with interactive information graphics done by Burmester et al. [31]. They found that graphic usage duration differed between users and that story-based approaches, although motivating readers, might lead to less intensive reception of information as interactive information graphics tend to expose readers to too much information [31]. Endert et al. [60] tried to merge human intuitive capabilities of interactive visualization with the big data processing capabilities of analytics. They argue that the involvement of human analysts in the task of creating storylines from large amounts of data is too explorative, questions are ill-defined or unknown a priori, and training data is not available. They conclude that purely visual methods are insufficient, but using visualization as a medium for human-data interaction is recommended. They suggest design principles where both user input and visual feedback are placed in the context of the process of an analyst. Danner [52] studied the dynamics between writers of stories and organizations that influence the shape and effectiveness of stories. Story writers must ensure that storylines are statistically true, actionable, and humanizing [52]. The work of Watson et al. [206] and Ping and Chen [153] give a good overview of various storyline visualization tools and applications.

c) Video.: Videos have long served as tools to tell stories and to inform people, as in documentaries. Videos when combined with data visualizations referred as data videos [192] can have a positive effect on learning [196], engagement [208], and viewers' focus [75]. This would explain the popularity of data videos in recent years as found by Amini et al. [5] [4]. However, creating effective visual data stories can be challenging due to the need for interdisciplinary skills. Xu et al. [212] propose multiple opening styles for data videos. Sun et al. [191]developed Erato as a human-machine cooperative data story editing system that uses an interpolation algorithm to help users create smoother transitions between frames. Additionally, adaptive narratives and personalized data stories are promising presentation approaches for enhancing engagement, as noted by Bonacini [19].

Interactivity, in general, can enhance the contextual dimension of data [63, 91] Interactivity is sometimes seen as a way to present visual information in interactive infographics. Interactive narratives can be linear, non-linear, or dynamic. Linear narratives are often associated with Scrollytelling, which combines storytelling and scrolling, as it allows the reader to explore the topic in depth by scrolling through the visual. The central element is a mostly full-screen animation with embedded elements like visualizations, video, textual, audible, or photographic content [137, 177]. Seyser and Zeiller [177] noted that scrollytelling articles are often text-centric and recognize image-centric articles where images/photos and videos order the elements. An elastic narrative allows following a predetermined order. In that respect, it should be considered that the story will branch off on specific points to allow a deeper understanding of the story. Riedl and Young [160] demonstrated that branching is effective as "they let users perform actions concurrently with system-controlled character actions."

4) High Context - Multimodal Experience:

a) VR Data stories.: With Virtual reality(VR) technologies acquiring greater immersive capabilities and increased levels of interaction, the possibilities of research for interactive visualizations and data research have increased. In fact, interactive visualizations made up nearly 42% of the analyzed literature in this review. Data representation in VR is no longer confined to static representation of the image. Data can now be visualized in 3D along with narration, sounds, and sensations giving the user a full immersive experience while interacting with the data. Although the field is relatively new, a significant amount of research has been done in understanding the data by visualizing it in a 3D environment [167]. When analyzing the work on visualizing data in an immersive environment, we looked into what are the different contexts in which representing data in VR would be helpful. How would VR technologies be helpful for data visualization? What are the different interactions studied by different research groups?

Lugmayr et al. [117] found several advantages for using 3D, VR, and immersive technologies for visualizing financial information. Among these advantages are presenting large amounts of data in a limited space and overcoming limitations due to a restricted physical space. By combining quantitative and qualitative data, information can be presented holistically and viewed from different perspectives. It was also concluded that using VR to visualize data supports data exploration, can increase user engagement, and allows the separation of different sets of data intuitively. A rare example in this field is the work by Ren et al. [158], who developed a prototype system to create immersive data-driven stories, supporting collaborative authoring and the use of different VR and AR devices. This tool was based on the work of Satyanarayan and Heer [171].

According to Marques et al. [129], augmented reality (AR) can enhance narrative visualization by improving view, focus, and sequence through simulation. Although research in this field is still in its early stages, future studies could investigate no-code AR authoring tools for designers. For instance, Ren et al. [158] created a prototype system that allows for collaborative authoring and supports using various VR and AR devices to produce immersive data-driven stories. Their work builds upon the research of Satyanarayan and Heer [171], who also explored authoring tools for AR. In another example, Soler-Domínguez et al. [188] proposed a set of scripted tools to accommodate storytelling in a 3D environment by collecting and visualizing the user's navigational data. Chen et al. [39] present a prototyping tool for creating augmented table tennis videos along with providing a design space for characterizing augmented sports videos based on their constituents and how to organize them. Madni* [124] employ storytelling techniques in VR to promote collaborative decision-making among different stakeholders in the process of systems engineering. Data storytelling is also getting increased attention in the VR research community. The review article by Rubio Tamayo et al. [167] provides a detailed summary of the state of the art of research and discusses ways in which collaboration in VR could benefit from data storytelling. Along the same line, Liestøl [110] explore the possibility of how to keep the balance and divide the control between the explanatory and the exploratory aspects of storytelling in VR. Xu and Ragan [211] provide an overview of the most used methods to preserve narrative control in VR. To address the problem of attention guidance, they propose using a virtual character to direct users' attention towards the target in a VR environment. Another example of using virtual characters for storytelling is to deliver complex geospatial information in the form of narratives. They compare the narratives by virtual characters with audio/text-only modalities [16].

D. Storytelling effects and evaluation

While the narrative building blocks shape how the story unfolds, the influence of the story, however, builds upon the user's interaction with the resulting artifact. When we listen to stories, our brains immediately react and both cognitive and non-cognitive areas of the brain are stimulated [1, 200]. To intentionally design data-driven stories that support as intended, it is crucial to understand the effects of storytelling and the mechanisms behind them. For this purpose, we summarize how different effects have been studied in storytelling visualization research. To provide an overview of our findings, we analyze the papers with regard to five categories—affective, cognitive, interactivity, indirect, and behavioral effects. For all studies, we identify the independent and dependent variables and related effects. We identify what methods were used to study the effects and whether studies were conducted in the lab or in the field. Lastly, we look at the individual sample sizes to conclude the expected replicability of the effects (see Fig. 9).

1) Variables of interest: Using storytelling in visualization has been claimed to improve several different aspects of communication [176]. It is critical to first look at the difference between information and knowledge. Information, in many cases, refers to numerical facts directly retrievable from the visualization, while knowledge refers to the change in cognition in a user that enables them to act or think differently about the subject. Knowledge requires interpretation [37]. Therefore, it is critical to separate these concepts into possible effects and variables. Before we identify how dependent and independent variables are interconnected, we identify which variables have been studied. The papers we selected for our review that used storytelling in data visualization utilized the following dependent variables:

- **Accuracy** refers to readers having a more correct understanding after viewing the visualization [204]
- **Aesthetics** refers to stronger perceptions regarding aesthetics [101, 192]
- **Attention** refers to the brain's process of selecting (visual) information for processing, measured using eye-tracking [53]. Yet, attention in the psychological sense would require a more extensive experimental paradigm.
- **Attitude** refers to beliefs and convictions a person has in the real world. Attitudes can be expressed in cognitive, affective, and behavioral manners. Studies measure changes in the willingness to judge the real world referent [109, 187].
- **Awareness** refers to the ability to improve situational awareness (e.g., of changes in a project) [30, 61, 187]
- **Cognitive Load** refers to the perceived mental exertion required to extract knowledge from a visualization. [126]
- **Communicativeness** refers the quality of the communication induced by the visualization [120]
- **Data wrangling strategies** refers to activating users to perform a more wide variety of data analysis strategies[94]
- **Depth of exploration** refers to activating users to perform deeper analyses[55]
- **Ease of Use** refers to how easy the visualization is to use [219, 225]
- **Effectiveness** refers to how effective the visualization is to convey data [202, 216]
- **Enjoyment** defines the subject's joy from interacting with a data visualization [204].
- **Engagement** refers to the depth of the emotional connection between the topic and the users' emotions. [30, 204]
- **Focus** refers to an improvement in filtering out irrelevant information which is not part of the visualization [204, 222].
- **Information retrieval** refers to the quality of how factual information is retrieved from the visualization [175, 222]. visualization. [63, 112]
- **Insights** refers to the number of knowledge units new to the reader extracted from a visualization [30]
- **Interaction** refers to the number of interactions users do with an interactive visualization [225]
- **Insights** refers to the number of individual novel insights gathered from the visualization [30]
- **Interpretation** refers to how data or information is converted to knowledge. Visualizations can aid interpretation by helping the viewer to attain the "correct" or "intended" knowledge from the provided information [10, 61, 112, 127].
- **Likeability** refers to the overall positive affective reaction of the user towards the
- **Memorability** refers to the quality of being memorable, without necessarily measuring recall [144, 194].
- **Understanding** refers to the amount of correctly extracted knowledge. [30, 49, 63, 64, 101, 126, 142, 142, 144, 193, 204, 227]
- **Message credibility** refers to the credibility of a message [112]
- **Navigation** refers to the ease of explorability of a visualization. [63]
- **Recall** refers to the amount of correctly recalled information from a visualization. [30, 142, 204, 218]
- **Reading Experience** refers to the reported quality of experience while reading data [112, 226]
- **Usability** refers the overall improvement in usability when using storytelling [120, 192]
- **Value** refers to the perceived value of the reported data in journalistic contexts [53]

2) Evaluation methods used: Among all the eligible studies we found a variety of different approaches being used to understand storytelling in empirical research. Mixed-method designs are frequently used to better understand the impact of storytelling [53, 204]. To investigate the impact of storytelling on the user, most studies use lab or field experiments [64, 95, 204]. From all eligible papers, we found that only a few papers used crowd-sourced samples as a data-gathering technique. We expect to find only a few papers using crowd-sourcing because methods such as eye-tracking were often used to understand storytelling but are impossible to apply in crowd-sourcing [126]. Moreover, this is possibly due to the novel interest in storytelling in visualization research. When a field is new, internal validity may be valued more highly than external validity. As a consequence, field experiments are rare [30, 204].

3) Sample sizes: To understand the potential for replicability, we also **look at the sample sizes** in the studies that we review. We assume that larger samples should yield effects that are more likely to replicate. By using the definition of effect sizes from Cohen [44] ($\delta = 0.8$ as large, $\delta = 0.5$ as medium, $\delta = 0.2$ as small), assuming a between-subject difference in means design, and using power analysis (error rates: $\alpha < .05$, $\beta < 0.2$) we derive approximate sample sizes for experiments (i.e., small effects need large samples). For survey data—which are more prone to noise—we use the stricter error rate of $\alpha < .01$ and round our sample size to the nearest hundred. For qualitative data, we use sample sizes commonly used in research practice [131, 140]. We picked the lowest recommendations we could find for the categories. We categorize sample sizes as follows. For experimental data, we chose small (less than 25), medium (25–50), and large (over 50). For survey data, we chose small (less than 100), medium (100–500), and large (over 500). For qualitative data (Interviews, Focus Groups, etc.), we chose: small (less than 5 [140]), medium (5–15 [131]), and large (over 15).

4) The importance of engagement: In this section, we want to summarize the findings distilled from empirical studies on storytelling in data visualization. While in the first place, narrative visualizations are designed to get the individual's attention and facilitate understanding, if a visualization is designed effectively, the individual will stop scanning and will engage with the content [176]. This note aligns with the two-systems approach to *judgment and choice* as presented by Kahneman [97]: the fast and instinctive thinking—system 1—generates first impressions, intentions, and feelings. These are used as suggestions for the slow, deliberate system 2, which may then turn these into conscious beliefs and actions. Effective communication needs to connect with system 1. Otherwise, system 2 is less likely to engage, and people might subsequently not integrate the information [183]. Naturally, most effects of data-driven storytelling that are being researched, therefore, fall under the system 1 thinking mode. Both cognitive and affective involvement with the story is crucial for engagement: appealing features not only call for attention but also increase memorability and, thereby comprehension [6]. Whether communication was effective in achieving engagement can be determined by evaluating content, process, and outcome [164]. Content-wise, a story should feel relevant and be perceived as useful. Ideally, the process of designing the story includes defining key actors. Moreover, outcome measures should include measures of engagement [6]. **Engagement can be defined as an "affective, cognitive, and behavioral connection"** [146]. We investigate the state of research and discuss these three different measures of engagement.

a) Affective: The first reaction to a story is often an emotional one, and authors have explored different elements that contribute to this emotional response, such as attractiveness, interest [6], and likeability [63, 64]. These effects have been studied particularly well in journalistic contexts. For example, the work of De Haan et al. [53], shows that news consumers notice, read, and appreciate news visualizations more when they are seamlessly integrated with a news story. However, the impact of data storytelling extends beyond journalism. Figueiras [63] found that storytelling visualizations improve the likability, comprehensibility, and navigation of data. In an effort to gain insight into the spontaneous behavior of groups of people, Van Den Bosch et al. [198] presented a personalized weather forecast system that used interactive data videos to trigger social reflection that, in turn, induced emotional and narrative engagement among the family members. Additionally, data comics have been shown to improve enjoyment and engagement with visualized content [204]. They also appear to strengthen the focus of the reader by reducing the text-picture distance, thereby, facilitating understanding and reducing cognitive load. It is important to note, however, that some studies have found no impact on engagement when combining storytelling visualizations with exploratory data analysis [23].

An additional example of tube-map-like visualizations includes participants reporting higher levels of inspiration, activation, and energy, and showed more "eureka-like experiences" when using the visualization [30]. According to Lidskog et al. [108], a compelling story should include both an emotional appeal and a normative orientation. As happens with simple storytelling, data stories can also make people feel like they are losing track of time and space [118], also described as *narrative immersion* or *flow*. When individuals are confronted with their prior beliefs, emotions such as surprise or doubt can develop and prejudices are exposed [77]. As visualizations also carry a risk of oversimplification, annoyance or irritation is an emotion that can arise [46, 108], especially when people have more knowledge than what is being displayed in the story. Researchers also look at the change in attitude—or the persuasiveness of data-driven narratives—which is related to the cognitive as well as the behavioral layer of engagement [187]. Liem et al. [109] cautioned that while it is a common assumption that visuals change attitude, there is little empirical evidence for it. They also do not find support for this assumption in their studies. Heyer et al. [77] found that eliciting individuals' prior beliefs together with the display of the visualization does not increase attitude change—which is in line with the limited impact of visuals on attitude that Liem et al. [109] mentioned—but does result in people feeling surprised. Viewers of the visualizations also expressed doubt and prejudice. When given a text, people gave more prejudicial comments compared to a visualization. Summarizing their results, Heyer et al. [77] also call for investigating different design strategies for informing people on the one hand and changing their beliefs or attitudes on the other. In addition to attitude, Pérez-Montoro [152] have investigated attributions that individuals make towards a brand that has been featured in a data story, along with identification with that brand and trust. Stories can enhance perceptions and appeal [192] and raise awareness for complex topics [46, 61], yet the goal is to also connect with people on a cognitive and behavioral level.

b) Cognitive: Besides affective reactions to storytelling, there are several cognitive effects as well. Compared to reading a text with the same information, people seem to learn faster and more when looking at visualizations [77]. Visual cues (such as transparency or introducing additional elements) that guide the individuals' attention towards certain story parts have also

Dep. variable	Ind. variable	Effect	Method	Lab	Wild	Sample sizes	Papers
Accuracy	Datacomic	0	mixed-method	●		medium–large	[204]
Aesthetics	Visual cue types	+/-	quantitative	●		medium–large	[101]
	datavisualization	+/-	quantitative		●	small	[192]
Attitude	Added empathy conditions	+	quantitative		●	medium	[109]
	Interactiontechnique	0	quantitative		●	medium	[77]
	Data representation	+	quantitative	●		medium	[77]
Attention	News visualizations	+	mixed-method	●		large	[53]
Awareness	Static map visualization	+	quantitative	●		small	[30]
	Storytelling visualization	+	quantitative		●	large	[187]
Cognitive load	Layout ofi nfographics	+	eye-tracking/experimental	●		small	[126]
Communicativeness	Interactive visualization	+	quantitative	●		medium	[120]
Data wrangling strategies	Learning environment	+	qualitative	●		small	[94]
Depth of exploration	Trivia game elements	+	quantitative		●	NA,onlineviews	[55]
Ease of Use	Video and annotated visualization	+	mixed-method	●		small	[225]
	VIStory, interactivestoryboard	+	qualitative/quantitative	●		medium	[219]
Effectiveness	Feature-driven animation	+	quantitative	●		small	[216]
	Narration & interactiveslides	+	quantitative	●		small	[202]
Enjoyment	Video and annotated visualization	+	mixed-method	●		small	[225]
	Datacomic	+	mixed-method	●		medium–large	[204]
Engagement	Static map visualization	+	quantitative	●		small	[30]
	Video and annotated visualization	+	mixed-method	●		small	[225]
	Datacomic	+	mixed-method		●	medium–large	[204]
	Linking	+	quantitative	●		large	[226]
	Storytelling in visualizations	+	focusgroup	●		large	[63]
	content and design of storylines	+	mixedmethod	●		medium	[6]
	narrative visualization techniques	0	quantitative		●	large	[23]
	personalised story	0	mixed-method		●	medium	[46]
	Interactive story	+	qualitative		●	large	[134]
	home and public setting		mixedmethod	●	●	small	[198]
Focus	Graphs with(out) background stories	+	quantitative	●		large	[222]
Information retrieval	Storygraph visualisation approach	+	experimental	●		medium	[175]
	Graphs with(out) background stories	+	quantitative	●		large	[222]
Insights	Static map visualization	+	quantitative	●		small	[30]
Interaction	Linking	+	quantitative	●		large	[226]
Interpretation	Visual expressiveness, annotations	+	qualitative	●		medium	[10]
	Familiarity, storycontext	+	qualitative	●		medium	[10]
	Visualizations with storyelements	+	qualitative	●		large	[61]
	Visualizations with context information	+	experimental		●	large	[127]
Likeability	Various storytelling elements	+	qualitative/experimental	●		small	[64]
Memorability	Interpretation and recall	-	quantitative	●		medium	[142]
	Datacomic	-	mixed-method	●		medium–large	[204]
	Interactive story	+	qualitative		●	large	[134]
Message credibility	Statistical information and data visualization	+	experimental		●	large	[112]
Navigation	Variousstorytellingelements	+	qualitative/experimental	●		small	[64]
Recall	Static map visualization	+	quantitative	●		small	[30]
	Interactive visualization	0	quantitative	●		medium	[142]
	Visual cue types	0	quantitative	●		medium–large	[101]
	Slideshow layout and linking	+	quantitative	●		large	[226]
	author driven narratives in visualizations	+	qualitativeandquantitative	●		medium	[144]
	data storytelling visualizations	+	experimental	●		large	[218]
Reading experience	Interactivity	+	experimental		●	large	[112]
	Linking	+	quantitative	●		large	[226]
Understanding	Annotated charts	+	qualitative	●		small	[64]
	Static map visualization	+	quantitative	●		small	[30]
	Various storytelling elements	+	qualitative/experimental	●		small	[64]
	Layout ofi nfographics	+	eye-tracking/experimental	●		small	[126]
	Interactivevisualization	+	quantitative	●		medium	[142]
	Interactive visualization	+	qualitative	●		medium	[95]
	familiarity, story context	+	qualitative	●		medium	[10]
	Visual expressiveness, annotations	+	qualitative	●		medium	[10]
	Datacomic	+	mixed-method		●	medium–large	[204]
	Interpretation and recall	+	quantitative	●		medium	[142]
	Visual cue types	0	quantitative	●		medium–large	[101]
	Slideshow layout	+	quantitative	●		large	[226]
	authordriven narratives in visualizations	+	qualitativeandquantitative	●		medium	[144]
Usability	Interactive visualization	+	quantitative	●		medium	[120]
	Interface	+	qualitative	●		medium	[192]
	Video and annotated visualization	+	mixed-method	●		small	[225]
	News visualizations	+	mixed-method	●		large	[53]
Value	News visualizations	+	mixed-method	●		large	[53]

Fig. 9. Overview of all dependent and independent variables in papers that contained empirical studies on the effects of storytelling

been shown to help people comprehend the story better [101]. This and the overall user experience can be evaluated by testing factors such as salience and relevance using simple Likert scales [86]. Comprehension of stories evaluated with the QUEST model (e.g., [71]) has shown that visualization strategies matter for both understanding and perceived coherence of the story [93]. Telling stories with data can reduce the cognitive load of processing the information [107, 134, 192], enhance decision-making quality [18], improve understanding of data [46, 64, 222], ease understanding and perceived usefulness [6], and improve comprehensibility [63]. The effect of visual cues in the presence of verbal (audio) narrations, for example, has been studied by Kong et al. [101]. Using data-driven storytelling, their results show that integrated and separate visual cues such as glow, desaturation, depth of field, etc., help to direct attention to relevant visualizations faster and maintain the reader's focus. The authors also provide suggestions based on the role of cues and their effectiveness. They found that brightness-based cueing was perceived as both most effective and aesthetically pleasing. Yet, their study did not show an actual influence on the participants' comprehension and recall. These findings tie into the results by [30], who also found that using a tube-map-like visualization helps to gain attention and improves focus. Similarly, Zhao et al. [222] found that background story influences participants' focus areas during interactive graph explorations. Borkin et al. [20] offer a clear overview using a visualization taxonomy and experimental evidence on which techniques work best for recognition and recall. Using a tube-map-like visualization to show a project plan helped participants understand project goals better. The visualization improved the quality of discussions about the plan, improved individuals' attention to details, improved memorability of information, as well as recall [30]. While the evidence for improvement of understanding and recall seems to increase, it is still unclear whether this also leads to changes in opinion or attitudes. Shedding light on cognitive mechanisms in information synthesis, Mantri et al. [127] conducted four experiments where they presented a pair of line charts to the users sequentially. They found that when the two charts depicted relationships in opposite directions, participants tended to weigh the positive slope more. In a medium-sized study, Liem et al. [109] investigated the effect of using visual data storytelling on changes in attitude. While changes were detectable, the results showed smaller effects on participants' attitudes than expected [109]. After having attained the reader's attention it is necessary to ensure understanding or comprehension by the viewer of a visualization. The effect of storytelling on understanding has been investigated in multiple studies. Most studies that address understanding investigate the effect of different visualization types or narrative structures on the amount of correctly extracted information from the visualization. By simply connecting data visually in a way that fosters imagining a story behind the data—as proposed by Schumann et al. [175]—users are asynchronously supported in developing an awareness of the underlying data. Schumann et al. [175] used graphs to visually connect audio recordings in a repository to help collaborators understand the state of the project using a story graph. Their evaluation finds that this approach improved understanding of important information about the repository. For data comics, findings even provide instructive insights. Bach et al. [10] identified eight design factors for creating data comics that helped participants interpret the visualizations without training and only minimal annotations. They used data comics to communicate changes in dynamic networks. Similarly, Wang et al. [204] found that data storytelling fosters a more enjoyable and engaging experience if combined with a comic-style presentation, and the use of comic-style improved understanding and recall. Using a slide show layout, Zhi et al. [226] found that comprehension tasks were significantly performed better; moreover, recall by participants was improved using a slide show layout where additional interaction was provided. Using an interactive slide show was preferred by the participants. Correa and Silveira [47] conducted a focus group with data visualization professionals to understand how designers can use narrative concepts to allow readers to personalize their story experience. The results revealed that using narrative genres while designing helps maintain the narrative intent. Similarly, participants preferred an interactive storyboard visualization—VIStory—over a timeline-based visualization. Zeng et al. [219] demonstrated the efficiency of the technique with a qualitative and quantitative study, which showed better understanding and faster time on task. Their tool was designed to help understand the body of literature in visualization research and participants were able to answer more questions more quickly.

When participants perceived information through featured-driven animation, they were quicker than common interactive visualizations—albeit with equally accurate results. Therefore, Yu et al. [216] recommend their use in scientific simulations. Adding a video of a human narrator and using an author-driven narrative structure significantly improved memorability. Yet, it had no significant effect on the long-term recall of information when compared against a reader-driven narrative structure [142]. Overall, we see evidence that the use of storytelling both improves the amount of correct information derived from the visualization and the time on task, improving the effectiveness and efficiency of the visualizations. However, most sample sizes of the studies shown here were either small or medium. Larger studies and even replications may be necessary to put the found effects on a more solid basis. By adding additional features like interactivity, the effectiveness of storytelling visualizations can be further improved.

c) Effects of Interactivity: When switching from an author-driven approach (what the author wants to tell) to a reader-driven approach, storytelling visualization becomes a tool of exploration. This is achieved by letting the user interact with the visualization. This changes the process from a perception-focused to a more interrogative procedure. The reader understands some information that helps form new questions, which the visualization may then answer. The use of interactivity impacts both the overall message of the visualization, understanding, and recall by changing the engagement of the user. Interactivity allows the reader to focus on aspects of the visualization that they connect with [176].

Investigating the effects of interaction on message credibility and reading experience Link et al. [112] found that interactions do not affect them. However, the perceived interactivity was seen to improve the reading experience. In an exploratory empirical study about user reception and usage behavior, [31] investigated engagement and motivation using interactive information graphics. The results showed that the usage durations were heterogeneously distributed for different users and different interactive information graphics. Users spent more time for initial orientation without interacting when introductions were shown. They also found that story-based approaches motivated users but led to less intensive reception of information. Thus, interactivity is not without risk, as users need to be able to identify when and where to interact with the visualization. Otherwise, large parts of the visualization could remain unexplored. This effect can be addressed using elements of games. A user study comparing two information graphics—either using or not using trivia questions—Diakopoulos [55] found that users showed increased exploration of the data space when the trivia game elements were included. These findings were collected in a field setting, further strengthening the evidence of these results. Many forms of interactivity have been used in digital journalism, as found by Alexandre [3]. In their review of interactive storytelling visualizations from journalism, they find different patterns of how interactivity is achieved. Most frequently, the details-on-demand technique was used in the field. More importantly, the authors conclude that more research is needed to understand the effects on and the experience of readers when using interactivity in storytelling visualizations in journalism.

d) Indirect effects: Next to the several positive effects of data-driven storytelling, there are certainly also indirect effects, some of which may be undesirable. The current practice of using storytelling visualizations in journalism, for example, has become focused on the application of technology rather than providing the public with knowledge and insights (for a better understanding of current events) [220]. This bears the risk of letting readers deduct "facts" from these visualizations that could be inaccurate from the journalist's point of view. While reporting the data visually and seemingly neutral, visualizations may provide more transparency in journalism; they also bear the risk of further blurring the line between factual reporting, opinionated reporting, and manipulation and leave more leeway for misinterpretation. Using storytelling visualization requires a deep understanding of the domain, a strong scientific skill-set, and a thorough ethical reflection, as their use can have real-life consequences. According to a study by Holder and Xiong [81], data visualization design can shape readers' perceptions of the people represented in the data and perpetuate harmful stereotypes. This suggests that design choices play a critical role in interpreting data visualizations related to social inequity and can contribute to stereotyping.

On the other hand, an improvement in data literacy was observed as a side effect of storytelling visualizations by Jiang and Kahn [95] when students were asked to create visualizations themselves. By "getting personal with big data", an entry point to visualization is created, making information personally relatable, the interaction with big data for the youth can be enhanced [94, 95]. In alignment with this approach, Ploehn et al. [154] proposes 'using data visualizations as tools of control' by showing the actual 'bodies' that the data represents within the visualizations. This strategy can create a sense of control and provide a transparent platform for discussions surrounding the data.

Hullman et al. [85] have called for more research on the impact of psychological biases (such as anchoring and social proof) on the effectiveness of visualizations. Additionally, for designing a story sequence most effectively, it is important to consider user experience measures [100]. Drawing insights from other scientific disciplines such as behavioral economics—where biases and framing effects have been researched extensively [7, 195]—and include those in the design process of stories can help prevent undesired effects [84].

e) Behavioral: According to Echeverria et al. [57], a data story should always include a clear call to action. Behavioral effects can be divided into intentional and actual behavioral measures. Pérez-Montoro [152] looked at purchase and recommendation intention. In a study examining the intention of the viewers to alter their health-related behavior after watching a data video, Sallam et al. [169] found the influence of personality traits in the behavior change. They also found that providing solutions to the health problem presented in the video increased the video's actionability and induced higher behavioral change intentions. Boy et al. [24] investigated low-level interaction with the data story itself. They started by analyzing Google Analytics clicking data and then took a closer look at the users' behavioral measures. Their findings revealed no significant differences between a visualization-savvy population and those less acquainted with visual stories.

While Segel and Heer [176] among others have called for more research on the engagement of the reader with the story and the data. Of the 35 papers we have included in this section, only 19 have looked at effects somehow. While many do not study but recognize the importance of researching effects, the remaining papers do not mention it at all. Definitions and degree scales of engagement such as the one put forward by Mahyar et al. [125] are certainly very helpful in this process—in line with our argument, they put decision-making based on the visualization as the highest level of engagement.

In a study examining viewers' willingness to change their health-related behavior after watching a data video, Sallam et al. [169] discovered that personality traits played a role in behavior change. They also found that providing solutions to the health problem presented in the video increased its actionability and induced higher behavioral change intentions.

V. Discussion

In this article, we have examined the literature on storytelling in data visualization, focusing on the use of frameworks, data story types, application areas, narrative structures, and their effects and measurement. This analysis was conducted to help researchers identify pertinent research gaps in storytelling in data visualization. We present an overview of the domain development and future research directions for each topic and address the questions stated in the introduction.

What are the theoretical or conceptual foundations of the story creation process, and how is it structured?

Several frameworks for developing narrative visualizations exist, each tailored to specific storytelling contexts, resulting in heterogeneity among frameworks. These frameworks aim to provide narrative recommendations based on the data and context at hand. Researchers have summarized methods for deriving recommendations from existing narrative visualizations and their creation processes. However, the relationship between underlying cognitive processes and data storytelling is still under investigation. To achieve consistency in framework design, it is essential to comprehend the cognitive aspects of information processing, narrative structures, and their influence on our perception of information.

Our review highlights recent studies that specifically focus on underlying cognitive and behavioral processes. The internal construction of stories and the attributes involved can be explained using multimodal processing theory, psychology, and basic narratological concepts. While recent work has increasingly concentrated on storytelling automation, primarily using low-level perception metrics, future work will incorporate ideas from recommendations based on more individualized metrics and narrative structures. An improved understanding of underlying cognitive mechanisms will enable visualization researchers to develop frameworks for creating compelling and effective visual data stories more automatically or semi-automatically.

How can individual data stories be structured within a taxonomy that reflects story construction? Based on our survey, we propose a classification scheme centered on the level of contextualization and the level of control within story construction. We focus on the presented data combined with the audience's explicit and implicit knowledge. Contextualization assists users in understanding the data, with different levels ranging from presenting raw data to constructing physical metaphors that represent the data in context. The level of control, on the other hand, enables users to explore and interact with the data through a data story, with static, dynamic, and interactive levels depending on the user's control over the story's event order. This classification scheme offers an overview of data storytelling's evolution and provides insights into the research direction in this field.

What data types were used in the analyzed stories, and in which areas were they applied? Recent publications on storytelling in data visualization have shifted their focus from explanatory to interactive and exploratory visualization. This is understandable, as designers of storytelling visualizations often reflect on concrete messages when designing. Storytelling is gaining a more prominent role in general communication scenarios, particularly in the public sector, including science, open data, education, journalism, and healthcare. Since the categorization by Segel and Heer, computer-driven visualization has become more dominant in fields such as interactivity and virtual reality. This shift has moved storytelling from primarily author-driven to more reader-driven, enabling increased exploration. No strong correlation between the purpose and application areas was found, although interactive storytelling seems more frequently used for exploratory purposes in education and science. We advocate for more research to elevate storytelling in the private sector.

How is the information organized? What are the different narrative structures, and which goals are pursued? While Segel and Heer focused mainly on visualization "genre," other aspects of storytelling have played an increasingly important role in understanding how narrative structures can enhance visualizations. Recent research has incorporated narratological aspects into the features investigated. Understanding the role of events, existents, content, and different dimensions of expression in a narrative visualization context

can facilitate story creation and understanding of effects. Narratology, with its long history of reflecting elements and sequencing within narratives, is gaining a more prominent role in recent and future research directions.

Surprisingly, elements central to classical literature or film studies, such as characters, the role of the protagonist, or the antagonist, have been minimally studied. However, these elements significantly contribute to the user forming an emotional connection to the story, motivating them to follow it and contextualize the content.

What are the cognitive and non-cognitive effects of storytelling, and how were they measured empirically? Studies on storytelling in data visualization have considered various outcome variables using different methods and approaches across various frameworks and application areas. Many papers utilize qualitative methods to comprehend the effects of storytelling visualization on the reader, offering deep insights into the reception process during storytelling visualization design. These qualitative approaches reveal unexpected effects and provide insights for developing theories about storytelling effects in visualizations. However, a quantitative perspective could strengthen the evidence for many of these findings.

Improvement in understanding and recall is the most frequently studied quantitative outcome, as evidenced in several studies. Nevertheless, sample sizes were not always large, and measurements were inconsistent across studies. Notably, we found few negative results.

The use of interactivity in storytelling visualizations presents conflicting evidence. Although it fosters a deeper connection between the reader and data, it risks making the visualization less transparent and navigable, potentially even reducing the breadth and depth of information retrieval. Therefore, interactivity must be planned with care.

Many papers that report techniques and tools resort to ad-hoc evaluation methods. While the community should value the engineering approach of designing a technique or tool on its own, adding subpar post-hoc evaluation into a paper—just to meet scientific progress expectations—negates much of the effort in developing a technique and might introduce problematic findings. There is growing evidence in psychological research that inconsistent and ad-hoc operationalization may induce false positives in the research body, especially if researchers can manipulate the measurements. Combined with a publication bias towards papers showing significant results, there is a risk of more papers with instrumentation biased towards false positives.

This does not mean that these findings are non-existent. However, research on storytelling and outcomes should consider consolidating research methods and instrumentation to ensure the replicability of findings. This can be accomplished by providing web repositories of measurement methods and scales, conducting specific studies on measurement validity, and conducting more field studies. While laboratory and survey experiments show high internal validity and might replicate well in other lab studies, it is crucial to study how well these findings translate to real-world settings. To establish a theory of storytelling effectiveness in visualization, we recommend engaging more real end-users in real-life settings. Using causal modeling approaches could ensure the applicability of theory to real-world problems.

VI. Conclusion and Future Work

This paper has provided an extensive overview of the role of storytelling in visualization. While storytelling is a historical phenomenon, predating even the concept of visualization, it is becoming increasingly critical in the realm of data dissemination and communication. Given the interdisciplinary nature of storytelling research, it is unsurprising that various fields focus on different aspects. The challenge for future research is to bridge these gaps and illuminate the field from a comprehensive perspective. This opens up several new lines of inquiry:

- **Integration of existents:** Traditional narratology can enhance data-driven storytelling. For instance, how can we integrate characters, protagonists, or the relationship between users and data into our narratives?
- **Understanding and utilizing the setting:** With emerging contexts like virtual reality or physicalization, we need to comprehend how these unique settings impact storytelling.
- **Evaluation approach:** Future evaluations need to consider not just cognitive effects but also emotional and contextual impacts.
- **Participatory story creation:** Storytelling thrives on the construction of an internal model. Therefore, we must explore how to create narratives across multiple disciplines.
- **Cross-disciplinary research:** There is a strong need for multi-disciplinary research approaches to incorporate relevant perspectives.
- **Understanding temporal structures:** We need to extend our understanding of temporal structures to accommodate non-linear, interactive approaches.

We observed that storytelling has been primarily utilized in the public and journalism sectors for explanatory purposes, often based on small sample sizes. Increasingly, non-linear interactive visualizations are being employed for exploratory purposes. There seems to be a scarcity of articles analyzing storytelling in the private domain, except for journalism and sports sectors, with some exceptions in the business domain.

Data-driven stories have shown immense potential in enhancing understanding and informing decision-making. In a world that generates an ever-growing volume of data, we can think of data not only as the "new coal," but also as a raw material that requires refined skills to distill into valuable "data diamonds." Storytelling in visualization lies at the heart of these skills, and its impact on visualization research has been steadily growing. We expect this trend to continue and even accelerate in the future. Data without context is meaningless; for data to have an impact, it requires a story.

Acknowledgements

We want to thank Benjamin Bach for his feedback on an earlier draft of this manuscript. We would also like to thank the anonymous reviewers from a previous submission for their feedback. The introduction of this article was improved using ChatGPT4 by using a fully written draft and using the following prompt: "Please rewrite the following introduction in a classic style of scientific prose to improve comprehension." We ensured that no new information was added by ChatGPT and re-edited the result. All authors agree with the final version of the introduction.

References

[1] N. Y. AbdulSabur, Y. Xu, S. Liu, H. M. Chow, M. Baxter, J. Carson, and A. R. Braun. Neural correlates and network connectivity underlying narrative production and comprehension: A combined fmri and pet study. *Cortex*, 57:107–127, 2014.

[2] M. J. Albers. Infographics and communicating complex information. In *International conference of design, user experience, and usability*, pages 267–276. Springer, 2015.

[3] I. Alexandre. Promoting insight: A case study of how to incorporate interaction in existing data visualizations. In *2016 20th International Conference Information Visualisation (IV)*, pages 203–208. IEEE, IEEE, 2016.

[4] F. Amini, N. Henry Riche, B. Lee, C. Hurter, and P. Irani. Understanding data videos: Looking at narrative visualization through the cinematography lens. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 1459–1468, 2015.

[5] F. Amini, N. H. Riche, B. Lee, A. Monroy-Hernandez, and P. Irani. Authoring data-driven videos with dataclips. *IEEE transactions on visualization and computer graphics*, 23(1):501–510, 2016.

[6] V. J. C. Arevalo, L. N. Verbrugge, A. Sools, M. Brugnach, R. Wolterink, R. P. van Denderen, J. H. Candel, and S. J. Hulscher. Storylines for practice: A visual storytelling approach to strengthen the science-practice interface. *Sustainability Science*, pages 1–20, 2020.

[7] D. Ariely. *Predictably irrational - the hidden forces that shape our decisions.* Harper Collins New York, NY, 2008.

[8] D. B. Axelrod and J. Kahn. Intergenerational family storytelling and modeling with large-scale data sets. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children*, pages 352–360, 2019.

[9] M. S. A. Aziz et al. A review on the visual design styles in data storytelling based on user preferences and personality differences. In *2022 IEEE 7th International Conference on Information Technology and Digital Applications (ICITDA)*, pages 1–7. IEEE, 2022.

[10] B. Bach, N. Kerracher, K. W. Hall, S. Carpendale, J. Kennedy, and N. Henry Riche. Telling stories about dynamic networks with graph comics. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 3670–3682, 2016.

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

[12] A. Barczewski, A. Bezerianos, and N. Boukhelifa. How domain experts structure their exploratory data analysis: Towards a machine-learned storyline. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI EA '20, page 1–8, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450368193. doi: 10.1145/3334480.3382845. URL https://doi.org/10.1145/3334480.3382845.

[13] Z. Battad, A. White, and M. Si. Facilitating information exploration of archival library materials through multi-modal storytelling. In R. E. Cardona-Rivera, A. Sullivan, and R. M. Young, editors, *Interactive Storytelling*, pages 120–127, Cham, 2019. Springer International Publishing. ISBN 978-3-030-33894-7.

[14] R. K. Behera and A. K. Swain. Big data real-time storytelling with self-service visualization. In *Emerging Technologies in Data Mining and Information Security*, pages 405–415. Springer, 2019.

[15] R. P. Benjamin Bahr, Boris Lemmer. *Quirky Quarks*, pages XVIII, 319. Springer-Verlag Berlin Heidelberg, 2016. ISBN 978-3-662-49509-4. doi: 10.1007/978-3-662-49509-4. URL https://doi.org/10.1007/978-3-662-49509-4.

[16] K. Blackmore, S. Smith, K. Nesbitt, L. North, S. Wark, and M. Nowina-Krowicki. Evaluating a virtual human storyteller for improved decision support.

[17] K. Blackmore, S. Smith, K. Nesbitt, L. North, S. Wark, and M. Nowina-Krowicki. Evaluating a virtual human storyteller for improved decision support. 2019.

[18] V. Boldosova and S. Luoto. Storytelling, business analytics and big data interpretation: Literature review and theoretical propositions. *Management Research Review*, 43 No. 2:204–222, 2019. doi: 10.1108/MRR-03-2019-0106.

[19] E. Bonacini. Engaging participative communities in cultural heritage: Using digital storytelling in sicily (italy). *International Information & Library Review*, 51(1):42–50, 2019. doi: 10.1080/10572317.2019.1568786. URL https://doi.org/10.1080/10572317.2019.1568786.

[20] M. A. Borkin, Z. Bylinskii, N. W. Kim, C. M. Bainbridge, C. S. Yeh, D. Borkin, H. Pfister, and A. Oliva. Beyond memorability: Visualization recognition and recall. *IEEE transactions on visualization and computer graphics*, 22(1):519–528, 2015.

[21] T. Botsis, J. E. Fairman, M. B. Moran, and V. Anagnostou. Visual storytelling enhances knowledge dissemination in biomedical science. *Journal of Biomedical Informatics*, 107:103458, jul 2020. doi: 10.1016/j.jbi.2020.103458. URL https://doi.org/10.1016/j.jbi.2020.103458.

[22] L. Bounegru, T. Venturini, J. Gray, and M. Jacomy. Narrating networks: Exploring the affordances of networks as storytelling devices in journalism. *Digital Journalism*, 5(6):699–730, 2017.

[23] J. Boy, F. Detienne, and J.-D. Fekete. Storytelling in information visualizations: Does it engage users to explore data? In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 1449–1458, 2015.

[24] J. Boy, F. Detienne, and J.-D. Fekete. Can initial narrative visualization-techniques and storytelling help engage online-users with exploratory information visualizations. *Proceedings of the CHI, Seoul, Korea*, pages 18–23, 2015.

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

[26] N. O. Brolcháin, L. Porwol, A. Ojo, T. Wagner, E. T. Lopez, and E. Karstens. Extending open data platforms with storytelling features. In *Proceedings of the 18th Annual International Conference on Digital Government Research*, dg.o '17, page 48–53, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450353175. doi: 10.1145/3085228.3085283. URL https://doi.org/10.1145/3085228.3085283.

[27] L. Brunet, J. Tuomisaari, S. Lavorel, E. Crouzat, A. Bierry, T. Peltola, and I. Arpin. Actionable knowledge for land use planning: Making ecosystem services operational. *Land Use Policy*, 72:27–34, 2018.

[28] C. Bryan, K.-L. Ma, and J. Woodring. 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, 2016.

[29] R. Burkhard, J. Perhac, S. Asada, A. Troyanov, S. Zhong, Y. Jiang, and S. Schubiger. 4d-ux: User experience design principles for coupling multidimensional visual representations in presentations. In *2018 22nd International Conference Information Visualisation (IV)*, pages 379–385. IEEE, 2018.

[30] R. A. Burkhard and M. Meier. Tube map visualization: Evaluation of a novel knowledge visualization application for the transfer of knowledge in long-term projects. *J. UCS*, 11(4):473–494, 2005.

[31] M. Burmester, M. Mast, R. Tille, and W. Weber. How users perceive and use interactive information graphics: An exploratory study. In *2010 14th International Conference Information Visualisation*, pages 361–368. IEEE, 2010.

[32] Y.-R. Cao, J.-Y. Pan, and W.-C. Lin. User-oriented generation of contextual visualization sequences. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI EA '20, page 1–8, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450368193. doi: 10.1145/3334480.3383057. URL https://doi.org/10.1145/3334480.3383057.

[33] S. Caquard and J.-P. Fiset. How can we map stories? a cybercartographic application for narrative cartography. *Journal of Maps*, 10(1):18–25, 2014.

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

[35] S. Chatman. *Story and Discourse*, page 288. Cornell University Press, 1980. ISBN 9781501741616. doi: 10.1515/9781501741616. URL https://doi.org/10.1515/9781501741616.

[36] A. S. Chaudhary and A. Arora. Storytelling data visualization for grievances management system. In *International Conference on Futuristic Trends in Networks and Computing Technologies*, pages 395–405. Springer, 2019.

[37] M. Chen, D. Ebert, H. Hagen, R. S. Laramee, R. Van Liere, K.-L. Ma, W. Ribarsky, G. Scheuermann, and D. Silver. Data, information, and knowledge in visualization. *IEEE computer graphics and applications*, 29(1):12–19, 2008.

[38] Y.-X. Chen, M. Reiter, and A. Butz. Photomagnets: supporting flexible browsing and searching in photo collections. In *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*, pages 1–8, 2010.

[39] Z. Chen, S. Ye, X. Chu, H. Xia, H. Zhang, H. Qu, and Y. Wu. Augmenting sports videos with viscommentator. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):824–834, 2021.

[40] N. Chotisarn, J. Lu, L. Ma, J. Xu, L. Meng, B. Lin, Y. Xu, X. Luo, and W. Chen. Bubble storytelling with automated animation: a brexit hashtag activism case study. *Journal of Visualization*, 24(1):101–115, 2021.

[41] N. Chotisarn, W. Pimanmassuriya, and S. Gulyanon. Deep learning visualization for underspecification analysis in product design matching model development. *IEEE Access*, 9:108049–108061, 2021.

[42] A. Choudhry, M. Sharma, P. Chundury, T. Kapler, D. W. Gray, N. Ramakrishnan, and N. Elmqvist. Once upon a time in visualization: Understanding the use of textual narratives for causality. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1332–1342, 2020.

[43] W.-T. Chu, C.-H. Yu, and H.-H. Wang. Optimized comics-based storytelling for temporal image sequences. *IEEE Transactions on Multimedia*, 17(2):201–215, 2014.

[44] J. Cohen. *Statistical power analysis for the behavioral sciences*. Academic press, 2013.

[45] N. Cohn. Visual narrative structure. *Cognitive Science*, 37(3):413–452, 2013. doi: https://doi.org/10.1111/cogs.12016. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/cogs.12016.

[46] S. Concannon, N. Rajan, P. Shah, D. Smith, M. Ursu, and J. Hook. Brooke leave home: Designing a personalized film to support public engagement with open data. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2020.

[47] C. M. Correa and M. S. Silveira. Diving in the story: exploring tailoring in narrative data visualizations. In *Proceedings of the 21st Brazilian Symposium on Human Factors in Computing Systems*, pages 1–7, 2022.

[48] P. Cruz and P. Machado. Generative storytelling for information visualization. *IEEE computer graphics and applications*, 31(2):80–85, 2011.

[49] A. Cunningham, M. Nowina-Krowicki, J. Walsh, J. Chung, S. Wark, and B. Thomas. Provenance narratives and visualisation to support understanding and trust. In *El Sawah, S. (ed.) MODSIM2019, 23rd International Congress on Modelling and Simulation.* Modelling and Simulation Society of Australia and New Zealand, dec 2019. doi: 10.36334/modsim.2019.d1.cunningham. URL https://doi.org/10.36334/modsim.2019.d1.cunningham.

[50] M. D. Jones and D. Anderson Crow. How can we use the 'science of stories' to produce persuasive scientific stories? *Palgrave Communications*, 3(53):9, 2017. doi: 10.1057/s41599-017-0047-7. URL https://doi.org/10.1057/s41599-017-0047-7.

[51] S. Dailey, B. Gilmore, and N. Rangarajan. The visualization of public information: Describing the use of narrative infographics by us municipal governments. *Public Policy and Administration*, page 09520767221140954, 2022.

[52] P. Danner. Story/telling with data as distributed activity. *Technical Communication Quarterly*, 29(2):174–187, 2020.

[53] Y. De Haan, S. Kruikemeier, S. Lecheler, G. Smit, and R. Van der Nat. When does an infographic say more than a thousand words? audience evaluations of news visualizations. *Journalism Studies*, 19(9):1293–1312, 2018.

[54] V. de Martino, S. Rossetti, and D. Rossi. New approaches for an effective e-dissemination of statistics: The case of noi italia–100 statistics to understand the country we live in. *Statistical Journal of the IAOS*, 29 (3):159–166, 2013.

[55] N. Diakopoulos. Game-y information graphics. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*, pages 3595–3600. 2010.

[56] S. Diamond, R. Basu, S. Cao, and A. Hussain. The canadian cultural diversity dashboard: Data storytelling and visualization for the cultural sector. In *International Conference on Human-Computer Interaction*,

pages 372–384. Springer, 2021.
[57] V. Echeverria, R. Martinez-Maldonado, R. Granda, K. Chiluiza, C. Conati, and S. B. Shum. Driving data storytelling from learning design. In *Proceedings of the 8th international conference on learning analytics and knowledge*, pages 131–140, 2018.
[58] F. El Outa, M. Francia, P. Marcel, V. Peralta, and P. Vassiliadis. Towards a conceptual model for data narratives. In *International Conference on Conceptual Modeling*, pages 261–270. Springer, 2020.
[59] S. J. ElShafie. Making Science Meaningful for Broad Audiences through Stories. *Integrative and Comparative Biology*, 58(6):1213–1223, 09 2018. ISSN 1540-7063. doi: 10.1093/icb/icy103. URL https://doi.org/10.1093/icb/icy103.
[60] A. Endert, M. S. Hossain, N. Ramakrishnan, C. North, P. Fiaux, and C. Andrews. The human is the loop: new directions for visual analytics. *Journal of intelligent information systems*, 43(3):411–435, 2014.
[61] G. M. Fernandez-Nieto, V. Echeverria, S. B. Shum, K. Mangaroska, K. Kitto, E. Palominos, C. Axisa, and R. Martinez-Maldonado. Storytelling with learner data: Guiding student reflection on multimodal team data. *IEEE Transactions on Learning Technologies*, 14(5):695–708, 2021.
[62] G. M. Fernandez Nieto, K. Kitto, S. Buckingham Shum, and R. Martínez-Maldonado. Beyond the learning analytics dashboard: Alternative ways to communicate student data insights combining visualisation, narrative and storytelling. In *LAK22: 12th International Learning Analytics and Knowledge Conference*, pages 219–229, 2022.
[63] A. Figueiras. How to tell stories using visualization. In *2014 18th International conference on information visualisation*, pages 18–18. IEEE, 2014.
[64] A. Figueiras. 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, 2014. doi: 10.1109/IV.2014.79.
[65] A. Figueiras and Á. Vizoso. Information visualization: features and challenges in the production of data stories. In *Total Journalism: Models, Techniques and Challenges*, pages 83–96. Springer, 2022.
[66] C. Fish. Storytelling for making cartographic design decisions for climate change communication in the united states. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 55(2):69–84, 2020.
[67] P. Freixa Font, M. Pérez-Montoro Gutiérrez, and L. Codina. The binomial of interaction and visualization in digital news media: consolidation, standardization and future challenges. *Profesional de la información. 2021; 30 (4): e300401*, 2021.
[68] G. Freytag and E. MacEwan. *Freytag's Technique of the Drama: An Exposition of Dramatic Composition and Art*. S. C. Griggs, 1894. URL https://books.google.de/books?id=1-MPAAAAYAAJ.
[69] J. George-Palilonis and M. Spillman. Storytelling with interactive graphics: An analysis of editors' attitudes and practices. *Visual Communication Quarterly*, 20(1):20–27, 2013.
[70] N. Gershon and W. Page. What storytelling can do for information visualization. *Communications of the ACM*, 44(8):31–37, 2001.
[71] A. C. Graesser, K. L. Lang, and R. M. Roberts. Question answering in the context of stories. *Journal of Experimental Psychology: General*, 120(3):254, 1991.
[72] S. Gratzl, A. Lex, N. Gehlenborg, N. Cosgrove, and M. Streit. From visual exploration to storytelling and back again. In *Computer Graphics Forum*, volume 35, pages 491–500. Wiley Online Library, 2016.
[73] G. Groshans, E. Mikhailova, C. Post, M. Schlautman, P. Carbajales-Dale, and K. Payne. Digital story map learning for stem disciplines. *Education Sciences*, 9(2):75, 2019.
[74] M. T. Hasan, A. Wolff, A. Knutas, A. Pässilä, and L. Kantola. Playing games through interactive data comics to explore water quality in a lake: A case study exploring the use of a data-driven storytelling method in co-design. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, pages 1–7, 2022.
[75] J. Heer and G. Robertson. Animated transitions in statistical data graphics. *IEEE transactions on visualization and computer graphics*, 13(6):1240–1247, 2007.
[76] J. Heyer, N. K. Raveendranath, and K. Reda. Pushing the (visual) narrative: The effects of prior knowledge elicitation in provocative topics. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–14, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450367080. doi: 10.1145/3313831.3376887. URL https://doi.org/10.1145/3313831.3376887.
[77] J. Heyer, N. K. Raveendranath, and K. Reda. Pushing the (visual) narrative: the effects of prior knowledge elicitation in provocative topics. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2020.
[78] S. Hill and C. Grinnell. Using digital storytelling with infographics in stem professional writing pedagogy. In *2014 IEEE International Professional Communication Conference (IPCC)*, pages 1–7. IEEE, 2014.
[79] D. Hilviu and A. Rapp. Narrating the quantified self. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*, pages 1051–1056, 2015.
[80] Q. Ho, P. Lundblad, T. Åström, and M. Jern. A web-enabled visualization toolkit for geovisual analytics. In *Visualization and Data Analysis 2011*, volume 7868, page 78680R. International Society for Optics and Photonics, 2011.
[81] E. Holder and C. Xiong. Dispersion vs disparity: Hiding variability can encourage stereotyping when visualizing social outcomes. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):624–634, 2022.
[82] J. Hook. Facts, interactivity and videotape: exploring the design space of data in interactive video storytelling. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video*, pages 43–55, 2018.
[83] B. I. Hougaard and H. Knoche. Telling the story right: How therapists aid stroke patients interpret personal visualized game performance data. In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, pages 435–443, 2019.
[84] J. Hullman and N. Diakopoulos. Visualization rhetoric: Framing effects in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2231–2240, 2011. doi: 10.1109/TVCG.2011.255.
[85] J. Hullman, E. Adar, and P. Shah. The impact of social information on visual judgments. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1461–1470, 2011.
[86] J. Hullman, N. Diakopoulos, and E. Adar. Contextifier: automatic generation of annotated stock visualizations. In *Proceedings of the SIGCHI Conference on human factors in computing systems*, pages 2707–2716, 2013.
[87] J. Hullman, S. Drucker, N. H. Riche, B. Lee, D. Fisher, and E. Adar. A deeper understanding of sequence in narrative visualization. *IEEE Transactions on visualization and computer graphics*, 19(12):2406–2415, 2013.
[88] P. Isenberg, B. Lee, H. Qu, and M. Cordeil. Immersive visual data stories. In *Immersive Analytics*, pages 165–184. Springer, 2018.
[89] R. Jacob. Visualising global pandemic: A content analysis of infographics on covid–19. *Journal of Content, Community & Communication*, 11(6):116–123, 2020.
[90] M. Janowski, A. Ojo, E. Curry, and L. Porwol. Mediating open data consumption-identifying story patterns for linked open statistical data. In *Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance*, pages 156–163, 2019.
[91] M. Jern. Collaborative web-enabled geoanalytics applied to oecd regional data. In *International Conference on Cooperative Design, Visualization and Engineering*, pages 32–43. Springer, 2009.
[92] M. Jern. Explore, collaborate and publish official statistics for measuring regional progress. In *Cooperative design, visualization, and engineering*, pages 189–198. Springer, 2010.
[93] A. Jhala and R. M. Young. Comparing effects of different cinematic visualization strategies on viewer comprehension. In *Joint International Conference on Interactive Digital Storytelling*, pages 26–37. Springer, 2009.
[94] S. Jiang and J. Kahn. Data wrangling practices and process in modeling family migration narratives with big data visualization technologies. 2019.
[95] S. Jiang and J. Kahn. Data wrangling practices and collaborative interactions with aggregated data. *International Journal of Computer-Supported Collaborative Learning*, 15(3):257–281, 2020.
[96] J. E. Jung, M. Hong, and H. L. Nguyen. Serendipity-based storification: from lifelogging to storytelling. *Multimedia Tools and Applications*, 76 (8):10345–10356, 2017.
[97] D. Kahneman. *Thinking, fast and slow*. Macmillan, 2011.
[98] N. W. Kim, B. Bach, H. Im, S. Schriber, M. Gross, and H. Pfister. Visualizing nonlinear narratives with story curves. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):595–604, 2017.
[99] N. W. Kim, N. Henry Riche, B. Bach, G. Xu, M. Brehmer, K. Hinckley, M. Pahud, H. Xia, M. J. McGuffin, and H. Pfister. Datatoon: Drawing dynamic network comics with pen+ touch interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*,

pages 1–12, 2019.
[100] Y. Kim, K. Wongsuphasawat, J. Hullman, and J. Heer. Graphscape: A model for automated reasoning about visualization similarity and sequencing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 2628–2638, 2017.
[101] H.-K. Kong, W. Zhu, Z. Liu, and K. Karahalios. Understanding visual cues in visualizations accompanied by audio narrations. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2019.
[102] R. Kosara and J. Mackinlay. Storytelling: The next step for visualization. *Computer*, 46(5):44–50, 2013.
[103] T. Lan, O. O'Brien, J. Cheshire, A. Singleton, and P. Longley. From data to narratives: Scrutinising the spatial dimensions of social and cultural phenomena through lenses of interactive web mapping. *Journal of Geovisualization and Spatial Analysis*, 6(2):22, 2022.
[104] X. Lan, Y. Shi, Y. Wu, X. Jiao, and N. Cao. Kineticharts: Augmenting affective expressiveness of charts in data stories with animation design. *IEEE Transactions on Visualization and Computer Graphics*, 28(1): 933–943, 2021.
[105] S. Latif, Z. Zhou, Y. Kim, F. Beck, and N. W. Kim. Kori: Interactive synthesis of text and charts in data documents. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):184–194, 2021.
[106] B. Lee, N. H. Riche, P. Isenberg, and S. Carpendale. More than telling a story: Transforming data into visually shared stories. *IEEE computer graphics and applications*, 35(5):84–90, 2015.
[107] I. Liao, W.-H. Hsu, and K.-L. Ma. Storytelling via navigation: A novel approach to animation for scientific visualization. In *International Symposium on Smart Graphics*, pages 1–14. Springer, 2014.
[108] R. Lidskog, M. Berg, K. M. Gustafsson, and E. Löfmarck. Cold science meets hot weather: Environmental threats, emotional messages and scientific storytelling. *Media and Communication*, 8(1):118–128, 2020.
[109] J. Liem, C. Perin, and J. Wood. Structure and empathy in visual data storytelling: Evaluating their influence on attitude. In *Computer Graphics Forum*, volume 39, pages 277–289. Wiley Online Library, 2020.
[110] G. Liestøl. Story and storage–narrative theory as a tool for creativity in augmented reality storytelling. *Virtual Creativity*, 8(1):75–89, 2018.
[111] V. Y. Lim, L. M. M. Peralta, L. H. Rubel, S. Jiang, J. B. Kahn, and B. Herbel-Eisenmann. Keeping pace with innovations in data visualizations: A commentary for mathematics education in times of crisis. *ZDM–Mathematics Education*, pages 1–10, 2022.
[112] E. Link, J. Henke, and W. Möhring. Credibility and enjoyment through data? effects of statistical information and data visualizations on message credibility and reading experience. *Journalism studies*, 22(5):575–594, 2021.
[113] C. Lopezosa, M. Pérez-Montoro, and J. Guallar. Data visualization in the news media: Trends and challenges. *Technology, Business, Innovation, and Entrepreneurship in Industry 4.0*, pages 315–334, 2022.
[114] T. Losev, J. Raynor, S. Carpendale, and M. Tory. Embracing disciplinary diversity in visualization. *IEEE Computer Graphics and Applications*, 42(6):64–71, 2022.
[115] J. Lu, J. Wang, H. Ye, Y. Gu, Z. Ding, M. Xu, and W. Chen. Illustrating changes in time-series data with data video. *IEEE computer graphics and applications*, 40(2):18–31, 2020.
[116] A. Lugmayr, E. Sutinen, J. Suhonen, C. I. Sedano, H. Hlavacs, and C. S. Montero. Serious storytelling–a first definition and review. *Multimedia tools and applications*, 76(14):15707–15733, 2017.
[117] A. Lugmayr, Y. J. Lim, J. Hollick, J. Khuu, and F. Chan. Financial data visualization in 3d on immersive virtual reality displays. In *International Workshop on Enterprise Applications, Markets and Services in the Finance Industry*, pages 118–130. Springer, 2018.
[118] J.-L. Lugrin, M. Cavazza, D. Pizzi, T. Vogt, and E. André. Exploring the usability of immersive interactive storytelling. In *Proceedings of the 17th ACM symposium on virtual reality software and technology*, pages 103–110, 2010.
[119] P. Lundblad and M. Jern. Geovisual analytics and storytelling using html5. In *2013 17th International Conference on Information Visualisation*, pages 263–271. IEEE, 2013.
[120] A. Lunterova, O. Spetko, and G. Palamas. Explorative visualization of food data to raise awareness of nutritional value. In *International Conference on Human-Computer Interaction*, pages 180–191. Springer, 2019.
[121] Y. Lyu, T. F. Cheng, and R. Lin. Visual data storytelling: A case study of turning big data into chinese painting. In *International Conference on Human-Computer Interaction*, pages 526–535. Springer, 2020.
[122] L. J. M. Amerian. Key concepts and basic notes on narratology and narrative. *Scientific Journal of Review*, 4(10):182–192, 2015.
[123] K.-L. Ma, I. Liao, J. Frazier, H. Hauser, and H.-N. Kostis. Scientific storytelling using visualization. *IEEE Computer Graphics and Applications*, 32(1):12–19, 2011.
[124] A. M. Madni*. Expanding stakeholder participation in upfront system engineering through storytelling in virtual worlds. *Systems Engineering*, 18(1):16–27, 2015.
[125] N. Mahyar, S.-H. Kim, and B. C. Kwon. Towards a taxonomy for evaluating user engagement in information visualization. In *Workshop on Personal Visualization: Exploring Everyday Life*, volume 3, page 2, 2015.
[126] A. Majooni, M. Masood, and A. Akhavan. An eye-tracking study on the effect of infographic structures on viewer's comprehension and cognitive load. *Information Visualization*, 17(3):257–266, 2018.
[127] P. Mantri, H. Subramonyam, A. L. Michal, and C. Xiong. How do viewers synthesize conflicting information from data visualizations? *IEEE transactions on visualization and computer graphics*, 29(1):1005–1015, 2022.
[128] O. Marjanovic. Empowering business users to explore visual data through boundary objects and storytelling. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pages 5032–5041, 2016. doi: 10.1109/HICSS.2016.624.
[129] A. B. Marques, V. Branco, and R. Costa. Narrative visualization with augmented reality. *Multimodal Technologies and Interaction*, 6(12):105, 2022.
[130] R. Martinez-Maldonado, V. Echeverria, G. Fernandez Nieto, and S. Buckingham Shum. From data to insights: a layered storytelling approach for multimodal learning analytics. In *Proceedings of the 2020 chi conference on human factors in computing systems*, pages 1–15, 2020.
[131] M. Mason. Sample size and saturation in phd studies using qualitative interviews. In *Forum qualitative Sozialforschung/Forum: qualitative social research*, volume 11, 2010.
[132] E. Mayr and F. Windhager. Once upon a spacetime: Visual storytelling in cognitive and geotemporal information spaces. *ISPRS International Journal of Geo-Information*, 7(3):96, 2018.
[133] R. A. Metoyer, T. S. Chuanromanee, G. M. Girgis, Q. Zhi, and E. C. Kinyon. Supporting storytelling with evidence in holistic review processes: A participatory design approach. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1):1–24, 2020.
[134] M. Meuschke, L. A. Garrison, N. N. Smit, B. Bach, S. Mittenentzwei, V. Weiß, S. Bruckner, K. Lawonn, and B. Preim. Narrative medical visualization to communicate disease data. *Computers & Graphics*, 107:144–157, 2022.
[135] J. P. Michel and M. Ladd. "snow fall"-ing special collections and archives. *Journal of Web Librarianship*, 9(2-3):121–131, 2015.
[136] R. Minelli, L. Baracchi, A. Mocci, and M. Lanza. Visual storytelling of development sessions. In *2014 IEEE International Conference on Software Maintenance and Evolution*, pages 416–420. IEEE, 2014.
[137] E. Mörth, S. Bruckner, and N. N. Smit. Scrollyvis: Interactive visual authoring of guided dynamic narratives for scientific scrollytelling. *IEEE Transactions on Visualization and Computer Graphics*, 2022.
[138] M. Murray and A. Sools. *Narrative Research in Clinical and Health Psychology*, pages –. Palgrave Macmillan Ltd., United Kingdom, 2014.
[139] E. Nardi. Where form and substance meet: Using the narrative approach of re-storying to generate research findings and community rapprochement in (university) mathematics education. *Educational Studies in Mathematics*, 92(3):361–377, 2016.
[140] J. Nielsen. Why you only need to test with 5 users, 2000.
[141] H. O. Obie, C. Chua, I. Avazpour, M. Abdelrazek, J. Grundy, and T. Bednarz. A framework for authoring logically ordered visual data stories. In *2019 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, pages 257–259. IEEE, 2019.
[142] H. O. Obie, C. Chua, I. Avazpour, M. Abdelrazek, J. Grundy, and T. Bednarz. A study of the effects of narration on comprehension and memorability of visualisations. *Journal of Computer Languages*, 52: 113–124, 2019.
[143] H. O. Obie, C. Chua, I. Avazpour, M. Abdelrazek, J. Grundy, and T. Bednarz. Authoring logically sequenced visual data stories with gravity. *Journal of Computer Languages*, 58:100961, 2020.
[144] H. O. Obie, C. Chua, I. Avazpour, M. Abdelrazek, J. Grundy, and T. Bednarz. The effect of narration on user comprehension and recall of information visualisations. In *2020 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, pages 1–4. IEEE, 2020.
[145] H. O. Obie, D. T. C. Ho, I. Avazpour, J. Grundy, M. Abdelrazek, T. Bednarz, and C. Chua. Gravity++: A graph-based framework for constructing interactive visualization narratives. *Journal of Computer*

Languages, 71:101125, 2022.

[146] H. L. O'Brien and E. G. Toms. What is user engagement? a conceptual framework for defining user engagement with technology. *Journal of the American society for Information Science and Technology*, 59(6): 938–955, 2008.

[147] A. Ojo and B. Heravi. Patterns in award winning data storytelling: Story types, enabling tools and competences. *Digital journalism*, 6(6): 693–718, 2018.

[148] J. J. Otten, K. Cheng, and A. Drewnowski. Infographics and public policy: using data visualization to convey complex information. *Health Affairs*, 34(11):1901–1907, 2015.

[149] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan, et al. The prisma 2020 statement: an updated guideline for reporting systematic reviews. *Bmj*, 372, 2021.

[150] D. Park, M. Suhail, M. Zheng, C. Dunne, E. Ragan, and N. Elmqvist. Storyfacets: A design study on storytelling with visualizations for collaborative data analysis. *Information Visualization*, 21(1):3–16, 2022.

[151] Q. Peng. Storytelling tools in support of user experience design. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 316–319, 2017.

[152] M. Pérez-Montoro. *Interaction in Digital News Media: From Principles to Practice*. Springer, 2018.

[153] Q. Ping and C. Chen. Litstoryteller+: an interactive system for multi-level scientific paper visual storytelling with a supportive text mining toolbox. *Scientometrics*, 116(3):1887–1944, 2018.

[154] C. Ploehn, M. Steenson, and D. Byrne. Tsuga convictio: Visualizing for the ecological, feminine, and embodied. In *2020 IEEE VIS Arts Program (VISAP)*, pages 9–18. IEEE, 2020.

[155] A. Polanía and S. de Francisco Vela. Designing interfaces for legal artifacts. using storytelling to enhance digital legal experiences. In *Advances in Design and Digital Communication III: Proceedings of the 6th International Conference on Design and Digital Communication, Digicom 2022, November 3–5, 2022, Barcelos, Portugal*, pages 137–147. Springer, 2022.

[156] L. Qiang and C. Bingjie. Storycake: A hierarchical plot visualization method for storytelling in polar coordinates. In *2016 International Conference on Cyberworlds (CW)*, pages 211–218, 2016. doi: 10.1109/CW.2016.43.

[157] D. Ren, M. Brehmer, B. Lee, T. Höllerer, and E. K. Choe. Chartaccent: Annotation for data-driven storytelling. In *2017 IEEE Pacific Visualization Symposium (PacificVis)*, pages 230–239. Ieee, 2017.

[158] D. Ren, B. Lee, and T. Höllerer. Xrcreator: Interactive construction of immersive data-driven stories. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, VRST '18, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450360869. doi: 10.1145/3281505.3283400. URL https://doi.org/10.1145/3281505.3283400.

[159] M. Rickhaus. Visual anatomy of an article–lessons learned and taught in five figures. *Chimia*, 76(9):748–748, 2022.

[160] M. O. Riedl and R. M. Young. From linear story generation to branching story graphs. *IEEE Computer Graphics and Applications*, 26(3):23–31, 2006.

[161] S. Rodrigues, A. Figueiras, and I. Alexandre. 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, 2019.

[162] M. T. Rodríguez, S. Nunes, and T. Devezas. Telling stories with data visualization. In *Proceedings of the 2015 Workshop on Narrative & Hypertext*, NHT '15, page 7–11, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450337977. doi: 10.1145/2804565.2804567. URL https://doi.org/10.1145/2804565.2804567.

[163] R. Roels, Y. Baeten, and B. Signer. Interactive and narrative data visualisation for presentation-based knowledge transfer. In *International Conference on Computer Supported Education*, pages 237–258. Springer, 2016.

[164] B. Rohrmann. The evaluation of risk communication effectiveness. *Acta psychologica*, 81(2):169–192, 1992.

[165] H. Rosling. The best stats you've ever seen TED Talk. https://www.youtube.com/watch?v=hVimVzgtD6w, January 2007.

[166] R. E. Roth. Cartographic design as visual storytelling: Synthesis and review of map-based narratives, genres, and tropes. *The Cartographic Journal*, pages 1–32, 2020.

[167] J. L. Rubio Tamayo, M. Barro Hernández, H. Gómez Gómez, et al. Digital data visualization with interactive and virtual reality tools. review of current state of the art and proposal of a model. 2018.

[168] A. Saini, K. Mathur, A. Thukral, N. Singhal, and A. Parnami. Aesop: Authoring engaging digital storytelling experiences. In *The Adjunct Publication of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pages 56–59, 2019.

[169] S. Sallam, Y. Sakamoto, J. Leboe-McGowan, C. Latulipe, and P. Irani. Towards design guidelines for effective health-related data videos: an empirical investigation of affect, personality, and video content. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–22, 2022.

[170] H. Sanei, C. Wang, L. Zhu, and S. Jiang. Remixing as a key practice for coding and data storytelling.

[171] A. Satyanarayan and J. Heer. Authoring narrative visualizations with ellipsis. In *Computer Graphics Forum*, volume 33, pages 361–370. Wiley Online Library, 2014.

[172] K. Schröder, P. Belavadi, M. Ziefle, and A. Calero Valdez. The pension story-data-driven storytelling with pension data. In *International Conference on Human-Computer Interaction*, pages 404–415. Springer, 2022. doi: 10.1007/978-3-031-06018-2_28.

[173] K. Schröder, S. Kohl, F. de Jongh, M. Putzu, M. Ziefle, and A. Calero Valdez. Rethinking pension communication–the role of metaphors in information visualization. In *International Conference on Human-Computer Interaction*, pages 416–429. Springer, 2022. doi: 10.1007/978-3-031-06018-2_29.

[174] K. Schroeder, I. Eberhardt, W. Eberhardt, and A. Henkel. Show me my future: Data-driven storytelling and pension communication. *Netspar Design Paper*, 2023.

[175] J. Schumann, T. Buttler, and S. Lukosch. An approach for asynchronous awareness support in collaborative non-linear storytelling. *Computer Supported Cooperative Work (CSCW)*, 22(2):271–308, 2013.

[176] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6): 1139–1148, 2010.

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

[178] X. Shan, D. Wang, and J. Li. Research on data storytelling strategies for cultural heritage transmission and dissemination. In *HCI International 2022–Late Breaking Posters: 24th International Conference on Human-Computer Interaction, HCII 2022, Virtual Event, June 26–July 1, 2022, Proceedings, Part I*, pages 344–353. Springer, 2022.

[179] D. Shi, X. Xu, F. Sun, Y. Shi, and N. Cao. Calliope: Automatic visual data story generation from a spreadsheet. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):453–463, 2020.

[180] D. Shi, F. Sun, X. Xu, X. Lan, D. Gotz, and N. Cao. Autoclips: An automatic approach to video generation from data facts. In *Computer Graphics Forum*, volume 40, pages 495–505. Wiley Online Library, 2021.

[181] Y. Shi, T. Gao, X. Jiao, and N. Cao. Breaking the fourth wall of data stories through interaction. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):972–982, 2022.

[182] M. Shin, J. Kim, Y. Han, L. Xie, M. Whitelaw, B. C. Kwon, S. Ko, and N. Elmqvist. Roslingifier: Semi-automated storytelling for animated scatterplots. *IEEE Transactions on Visualization and Computer Graphics*, 2022.

[183] A. Shleifer. Psychologists at the gate: a review of daniel kahneman's thinking, fast and slow. *Journal of Economic Literature*, 50(4):1080–91, 2012.

[184] X. Shu, J. Wu, X. Wu, H. Liang, W. Cui, Y. Wu, and H. Qu. Dancingwords: exploring animated word clouds to tell stories. *Journal of Visualization*, 24(1):85–100, 2021.

[185] D. Simões, P. Antunes, and L. Carriço. Eliciting and modeling business process stories. *Business & Information Systems Engineering*, 60(2): 115–132, 2018.

[186] T. L. Smith and E. B. Moore. Storytelling to sensemaking: A systematic framework for designing auditory description display for interactives. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2020.

[187] W. So, E. P. Bogucka, S. Šćepanović, S. Joglekar, K. Zhou, and D. Quercia. Humane visual ai: Telling the stories behind a medical condition. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):678–688, 2020.

[188] J. L. Soler-Domínguez, M. Contero, and M. Alcañiz. Workflow and tools to track and visualize behavioural data from a virtual reality environment using a lightweight gis. *SoftwareX*, 10:100269, 2019.

[189] F. Stalph and B. Heravi. Exploring data visualisations: An analytical framework based on dimensional components of data artefacts in journalism. *Digital Journalism*, pages 1–23, 2021.

[190] T. Steinert, U. Große, M. Hirth, and H. Krömker. Mobility data stories

for a better understanding of mobility data. In *HCI International 2022–Late Breaking Papers: HCI for Today's Community and Economy: 24th International Conference on Human-Computer Interaction, HCII 2022, Virtual Event, June 26–July 1, 2022, Proceedings*, pages 533–546. Springer, 2022.

[191] M. Sun, L. Cai, W. Cui, Y. Wu, Y. Shi, and N. Cao. Erato: Cooperative data story editing via fact interpolation. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):983–993, 2022.

[192] T. Tang, J. Tang, J. Hong, L. Yu, P. Ren, and Y. Wu. Design guidelines for augmenting short-form videos using animated data visualizations. *Journal of Visualization*, 23(4):707–720, 2020.

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

[194] C. Tong, R. Roberts, R. Borgo, S. Walton, R. S. Laramee, K. Wegba, A. Lu, Y. Wang, H. Qu, Q. Luo, et al. Storytelling and visualization: An extended survey. *Information*, 9(3):65, 2018.

[195] A. Tversky and D. Kahneman. Judgment under uncertainty: Heuristics and biases. *science*, 185(4157):1124–1131, 1974.

[196] B. Tversky, J. B. Morrison, and M. Betrancourt. Animation: can it facilitate? *International journal of human-computer studies*, 57(4): 247–262, 2002.

[197] A. Tyagi, T. Estro, G. Kuenning, E. Zadok, and K. Mueller. Pc-expo: A metrics-based interactive axes reordering method for parallel coordinate displays. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):712–722, 2022.

[198] C. Van Den Bosch, N. Peeters, and S. Claes. More weather tomorrow. engaging families with data through a personalised weather forecast. In *ACM International Conference on Interactive Media Experiences*, pages 1–10, 2022.

[199] Q. Van Ho, P. Lundblad, T. Åström, and M. Jern. A web-enabled visualization toolkit for geovisual analytics. *Information Visualization*, 11(1):22–42, 2012.

[200] S. Virtue, T. Parrish, and M. Jung-Beeman. Inferences during Story Comprehension: Cortical Recruitment Affected by Predictability of Events and Working Memory Capacity. *Journal of Cognitive Neuroscience*, 20(12):2274–2284, 12 2008. ISSN 0898-929X. doi: 10.1162/jocn.2008.20160. URL https://doi.org/10.1162/jocn.2008.20160.

[201] J. Walsh, A. Cunningham, S. Wark, M. Nowina-Krowicki, and B. H. Thomas. Narrative visualisation of simulations: Finding the stories within the data. In *El Sawah, S. (ed.) MODSIM2019, 23rd International Congress on Modelling and Simulation.* Modelling and Simulation Society of Australia and New Zealand, dec 2019. doi: 10.36334/modsim.2019.B3.walsh. URL https://doi.org/10.36334/modsim.2019.B3.walsh.

[202] Y. Wang, D. Liu, H. Qu, Q. Luo, and X. Ma. 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, 2016.

[203] Z. Wang, H. Dingwall, and B. Bach. Teaching data visualization and storytelling with data comic workshops. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–9, 2019.

[204] Z. Wang, S. Wang, M. Farinella, D. Murray-Rust, N. Henry Riche, and B. Bach. Comparing effectiveness and engagement of data comics and infographics. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2019.

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

[206] K. Watson, S. S. Sohn, S. Schriber, M. Gross, C. M. Muniz, and M. Kapadia. Storyprint: an interactive visualization of stories. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*, pages 303–311, 2019.

[207] W. Weber, M. Engebretsen, and H. Kennedy. Data stories: Rethinking journalistic storytelling in the context of data journalism. *Studies in Communication Sciences*, 2018(1):191–206, 2018.

[208] D. Weng, R. Chen, Z. Deng, F. Wu, J. Chen, and Y. Wu. Srvis: Towards better spatial integration in ranking visualization. *IEEE transactions on visualization and computer graphics*, 25(1):459–469, 2018.

[209] S. Willemsen and M. Kiss. Keeping track of time: The role of spatial and embodied cognition in the comprehension of nonlinear storyworlds. *Style*, 54(2):172–198, 2020. ISSN 00394238, 23746629. URL https://www.jstor.org/stable/10.5325/style.54.2.0172.

[210] Y.-L. Wong and C.-S. Lee. Creative storytelling enhanced through social media and intelligent recommendation. In *Proceedings of the 8th ACM Conference on Creativity and Cognition*, C&C '11, page 399–400, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 9781450308205. doi: 10.1145/2069618.2069715. URL https://doi.org/10.1145/2069618.2069715.

[211] Q. Xu and E. D. Ragan. Effects of character guide in immersive virtual reality stories. In *International Conference on Human-Computer Interaction*, pages 375–391. Springer, 2019.

[212] X. Xu, L. Yang, D. Yip, M. Fan, Z. Wei, and H. Qu. From 'wow'to 'why': Guidelines for creating the opening of a data video with cinematic styles. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–20, 2022.

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

[214] S. Ya'acob, S. M. Yusof, D. W. H. Ten, and N. M. Zainuddin. An analytical reasoning framework for visual analytics representation. In *International Visual Informatics Conference*, pages 41–52. Springer, 2021.

[215] B. Yousuf and O. Conlan. Constructing narrative visualizations as a means of increasing learner engagement. In *HT (Doctoral Consortium/Late-breaking Results/Workshops)*, 2014.

[216] L. Yu, L. Harrison, and A. Lu. Effectiveness of feature-driven storytelling in 3d time-varying data visualization. *Journal of Imaging Science and Technology*, 60(6):60408–1, 2016.

[217] A. Zanda, J. Lutz, A. Heymann, and S. Bleisch. Technological infrastructure supporting the story network principle of the atlas of the ageing society. *Geografie*, 2019.

[218] D. Zdanovic, T. J. Lembcke, and T. Bogers. The influence of data storytelling on the ability to recall information. In *Proceedings of the 2022 Conference on Human Information Interaction and Retrieval*, pages 67–77, 2022.

[219] W. Zeng, A. Dong, X. Chen, and Z.-l. Cheng. Vistory: interactive storyboard for exploring visual information in scientific publications. *Journal of visualization*, pages 1–16, 2020.

[220] X. Zhang. Visualization, technologies, or the public? exploring the articulation of data-driven journalism in the twittersphere. *Digital Journalism*, 6(6):737–758, 2018.

[221] Y. Zhang, M. Reynolds, A. Lugmayr, K. Damjanov, and G. M. Hassan. A visual data storytelling framework. In *Informatics*, volume 9, page 73. MDPI, 2022.

[222] Y. Zhao, J. Shi, J. Liu, J. Zhao, F. Zhou, W. Zhang, K. Chen, X. Zhao, C. Zhu, and W. Chen. Evaluating effects of background stories on graph perception. *IEEE Transactions on Visualization and Computer Graphics*, 2021.

[223] Z. Zhao, R. Marr, and N. Elmqvist. Data comics: Sequential art for data-driven storytelling. *tech. report*, 2015.

[224] Z. Zhao, R. Marr, J. Shaffer, and N. Elmqvist. Understanding partitioning and sequence in data-driven storytelling. In *International Conference on Information*, pages 327–338. Springer, 2019.

[225] Q. Zhi, S. Lin, P. Talkad Sukumar, and R. Metoyer. Gameviews: Understanding and supporting data-driven sports storytelling. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2019.

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

[227] S. Zhu, G. Sun, Q. Jiang, M. Zha, and R. Liang. A survey on automatic infographics and visualization recommendations. *Visual Informatics*, 4 (3):24–40, 2020.