NewsLines: Narrative Visualization of News Stories

Mariana Costa, Sérgio Nunes

INESC TEC, Porto, Portugal
FEUP - Faculty of Engineering, University of Porto, Portugal

Abstract

Visual representations have the potential to improve information understanding. We explore this idea in the development of NewsLine, an open-source web-based prototype that focuses on narrative visualizations of news content. Having structured data as input, the prototype produces a storyline which showcases the narrative's events and participants, allowing the user to interact with the visualization in a number of ways. We built an information hub around the storyline to allow for multiple levels of exploration, specifically the main visualization, the event information module, and the sidebar. The visualization depicts the sequence of events that make up a news story, as well as the interactions between the involved parties in each event. The event information module presents additional information on a particular event. The sidebar is the "control center" of the visualization, unlocking a number of interactions and configurations. The prototype was evaluated with a user study with journalists and also with an online survey which gathered feedback from 178 potential end users. From these, 106 participants (60.6%) provided a rating of four or above (one to five scale) when asked to quantify their interest in using the application. Moreover, participants were asked to rank the importance of the visualization elements used. The results highlight that two elements stand out as the most important, the events and the entities. Overall, the participants generally found the application to be useful, but in need of some work in order for it to be made available to a broader public.

Keywords

narrative visualization, news visualization, text visualization, prototype

1. Introduction

In today's world, we are subject to an uninterrupted stream of information on a daily basis. The relentless pace at which news articles are communicated to the public often leads to an inability to process but a small fraction of them. Furthermore, not all news articles are unique nor relevant. This combination of excess, repetition, and questionable relevance places further stress on the reader, who must allocate a considerable amount of time to not only interpret the narratives within each news articles but to also carefully choose which articles are of interest.

One way to tackle this generalized disinterest revolves around minimizing the difficulties faced during the consumption of news. These difficulties can stem from two distinct stages of this process: the selection of news articles and their interpretation. In regards to the latter, not only must the reader determine whether or not a particular article is relevant, they must also make sense of the underlying narrative and contextualize it. We believe that there is value in

the study of journalistic narratives through semi or fully-automated means that can be applied to news stories regardless of their perceived impact or their origin. The possibility of using an application to easily explore news stories through representations other than pure text can open up new doors in terms of how the reader consumes the news. It is with this in mind that we propose the use of visualization to present the underlying narrative of a set of related news stories in a memorable and accessible manner.

In this paper we describe the development and evaluation of NewsLines, an open-source web-based JavaScript prototype for narrative visualization. In Section 2 we briefly present two existing solutions for the visualizing of storylines. Next, in Section 3 we present the NewsLines prototype and describe its main interface elements. We then present the user study and the survey conducted to evaluate the prototype in Section 4. Overall, the positive feedback gathered corroborates the potential which this type of visualization has to help news readers during the news consumption process. The paper ends with a summary of the contributions and an outline of possible future work.

2. Related work

Narrative visualization, a term popularized by Segel and Heer [1], tackles the interplay between narrative and visualization. However, as discussed by Padia [2], introducing storytelling elements in visualization is different from visualizing the narratives themselves. Whereas narrative visualization has been the subject of a considerable amount of academic projects, efforts into the study of these visual narratives are scarcer [2]. One technique originated by Munroe [3], the storyline visualization, attempts to visually encode narrative elements in a way that succinctly encapsulates the flow of a story through the manifestation of the interactions between its participants and the major events in which they partake. Entities are represented by lines that evolve from left to right, horizontal and parallel to each other. Adjacent lines indicate an interaction between the corresponding entities and events are encoded by nodes. This technique has been used extensively, both in more general applications [4, 5], as well as in specific scenarios such as meeting summarization [6], the evolving interactions between developers in software projects [7], and the analysis of plots in literary works [8].

We found that existing works rarely adopt a narrative-driven perspective, i.e., in which narrative elements such as actors and interactions are placed at the forefront. Our goal is to explore this opportunity by implementing a visualization solution that exposes the narrative elements of news stories, thus contributing to a better understanding of the underlying narratives.

3. Implementation

As a starting point to the implementation, we identified two existing projects, namely Narrative Charts [9] and Comic Book Narrative Charts [10]. Both projects make use of D3[1], a visualization engine built in JavaScript that allows for the creation of visualizations in browser-based environments, and both are made available as open-source software. These tools aim to mimic

---

[1]https://d3js.org/

```
1   {
2     "characters": [
3       {
4         "id": "rioters",
5         "name": "Rioters",
6         "affiliation": "rgb(204, 0, 0)",
7         "synonyms": ["proud boys", "large crowd", "mob", "ashli babbit"]
8       }
9     ],
10    "scenes": [
11      {
12        "characters": ["rioters"],
13        "description": "Members of the Proud Boys, a far-right group, join protesters, who
            ↪  had been assembled on the Capitol lawn since 10 a.m.",
14        "title": "Proud Boys",
15        "date": "2021-01-06 11:37:00",
16        "location": "Capitol"
17      }
18    ]
19  }
```

Listing 1: NewsLines JSON document example.

the storyline visualizations of Munroe [3]. Given a JSON file of characters and events, the systems output the position of each event in a storyline, leaving the styling and presentation of the output to the discretion of the user.

We opted to use Narrative Charts [9] as a starting point due to the thorough documentation that accompanied the project. Based on this engine, we built a browser-based prototype which, given as input a JSON document detailing entities and events belonging to a story, presents the entities in a visualization that allows the user to explore the story through a narrative-driven lens, focusing on the relationships between entities, events, locations, and dates. The source-code of the prototype can be found on GitHub [11]. A live version is also made available[2], together with a collection of example narratives in JSON format[3]. The project is licensed under the MIT license.

Each storyline is defined in an independent JSON document. The JSON structure is based on the one originally used by the Narrative Charts project with a few modifications. Listing 1 presents a basic example containing the definition of one character and one event where the character participes. The JSON Schema definition of the proposed format can be found in the repository[4]. The JSON files used for testing were produced manually, as studying the automatic generation of these files from already existing documents falls beyond the scope of this work.

The interface, depicted in Figure 1, is composed of three main views: the visualization, the event information module, and the sidebar. The visualization (**A**) is the core element of the
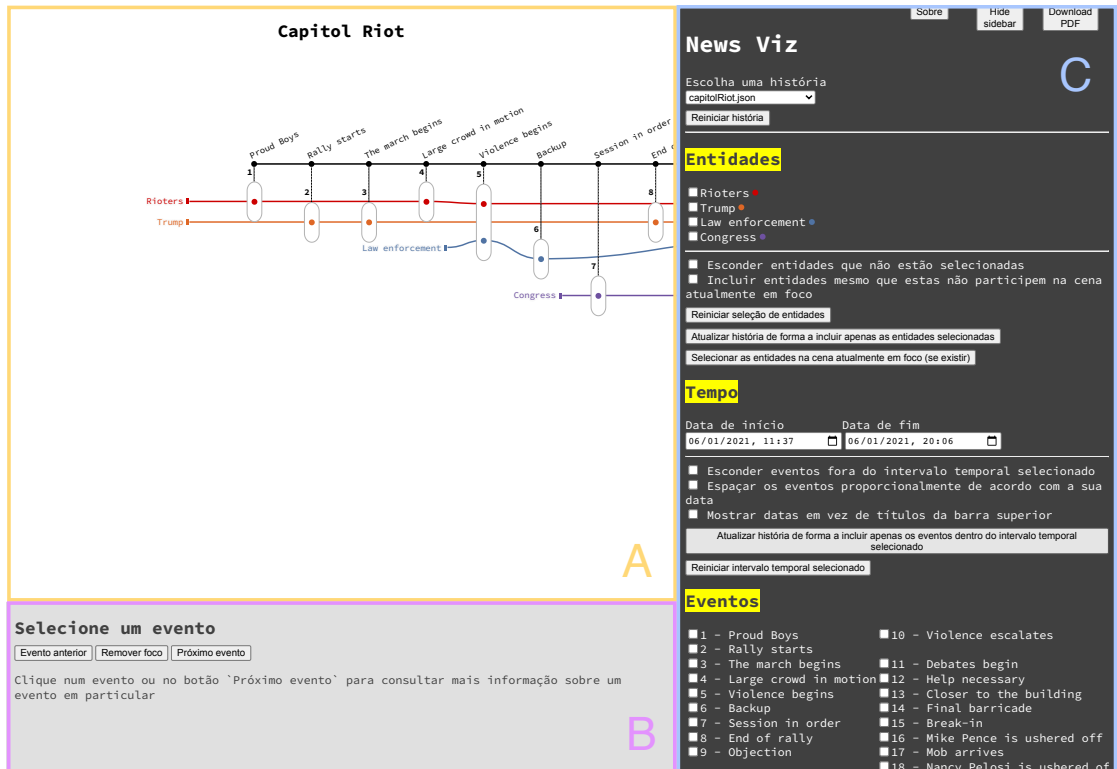

Figure 1: An overview of the interface of our application. Element **A** depicts the visualization (the core element of the interface), element **B** depicts the event information module (which provides more information on a particular event), and element **C** depicts the sidebar (the hub for the majority of the interactions available).

interface. We worked with the layouts provided by Narrative Charts (i.e., the positions of the introduction nodes, event nodes, entity lines, and appearances within an event) to produce an interactive storyline visualization, complemented by additional features such as: a timeline, the ability to drag events and introduction nodes, the ability to focus on a event to obtain more information on it, and the ability to hover a event to activate a tooltip (containing the date and location of the event, if available).

The event information module (**B**), placed at the bottom of the interface, displays a detailed description of the event in focus, along with some more contextual information such as the event's title, date, and location. The entities are highlighted using the color of the entity's line in the visualization. While the latter provides a bird's eye view of the events that make up that storyline, the event description module allows for a more detailed view of each individual event.

Finally, a collapsible sidebar (**C**) adds a number of additional interactions with the visualization, including: entity, event, and time frame selection; updates to the storyline based on those selections; and the option to filter elements according to a number of criteria (e.g., hiding entities which are not selected). In addition to these three main visual components, the user can also download the visualization as a PDF (reflecting all filters and updates applied to it) and
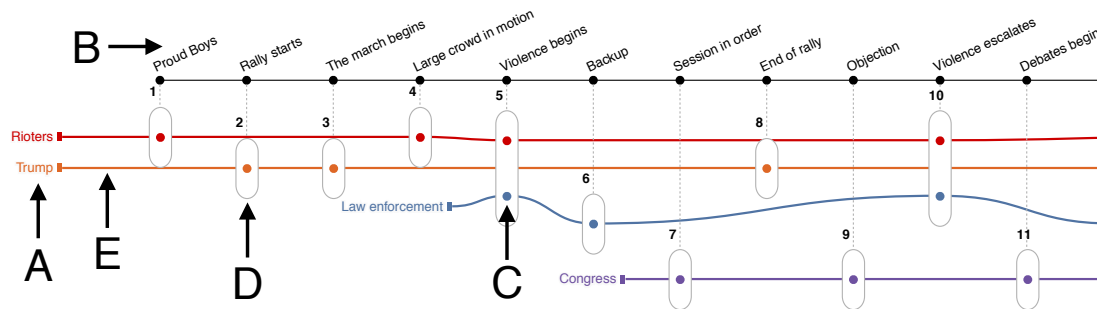
Figure 2: Crop of a PDF provided by the application, depicting the visualization.

visit an about page which contextualizes the project and provides information such as licenses and links to the source-code of the project and of the Narrative Charts implementation.

Figure 2 provides a description of the visualization. Central to this view is the notion of events, characterized by a description, a title, participants (i.e., appearances), a date (optional), and a location (optional). These events are translated visually in the form of ellipses, indicated as element **D**. Inside those ellipses are circles which depict the appearance of entities in that respective event (element **C**). The color of these circles correspond to the entity to which that appearance belongs. The lines flowing across the visualization represent the entities participating in the storyline (element **E**). Each line is attributed a random color if no user-specified color exists. The source and destination of the lines can either be an event or an introduction node (text legends that associate a line to a character's name, depicted as element **A**). Above the events and entity lines is a timeline that connects the events to a specific point in time or the title of the event in question (element **B**).

While the sidebar unlocks the majority of interactions with the visualization (e.g., filtering events by date), some do not require access to that module. The user can access a contextual tooltip by hovering over an event, which allows them to quickly access the date and location of the event. The tooltip is only visible if the event in question is in focus, or if no events are in focus. It is also possible to drag the events in the vertical direction, the introduction nodes in both vertical and horizontal directions, and the titles or dates in the timeline in the horizontal direction. Finally, clicking on an event will focus on it, updating the event information module with the details of that particular event.

In summary, we based our prototype on an existing open-source implementation of a storyline visualization, Narrative Charts. In addition to the adaptions made to the visualization itself, we built an information hub around the storyline to allow for multiple levels of exploration and control. The interface is thus composed of three main elements: the visualization, the event information module, and the sidebar. The visualization depicts the sequence of events that make up a news story, as well as the interactions between the involved parties in these events. The event information module presents additional information on a particular event. Finally, the sidebar is the "control center" of the visualization, unlocking a number of interactions.

Table 1
Tasks in the first part of the test.

	Functionality	Interaction taxonomy
T1	Focus on event	Select, explore, abstract/elaborate
T2	Move the events	Reconfigure
T3	Filter by date	Filter
T4	Select entities	Select
T5	Filter events	Filter
T6	Filter entities	Filter
T7	Update storyline	Reconfigure
T8	Reset storyline	Reconfigure
T9	Download PDF	-

4. Evaluation

To evaluate the efficacy of the developed prototype in exploring news stories from a narrative perspective, we conducted a user testing session with five expert users and a survey among potential end users. These are described in the next two sessions.

4.1. User Tests

We organized five individual test sessions, in which the participants were asked to complete a set of tasks and explore the system freely. All participants (all men) had strong ties to journalism, and all were working in the field. First, the participants were introduced to the prototype; the interviewer walked through the main components of the interface and some of the main concepts pertaining to the visual metaphor used. Then, the participants were asked to complete a number of tasks and, between each task, were asked to rate the usefulness of the underlying feature; these tasks are listed in Table 1. This exercise was followed by a period of open exploration, in which the participants were asked to interact with the system in whichever way they saw fit, voicing any comments or suggestions along the way. Finally, the test session concluded with a set of post-test questions related to the overall perception of the prototype.

The collected feedback was assessed using thematic analysis [12] and the results are briefly summarized under three main topics: positive feedback, negative feedback, and important features. For each topic highlighted below we include an example comment from the participants.

Considering the **positive feedback**, highlights include:

- Appealing visualization, *"the visualization is very simple, in a good way; it is easy to absorb"*;
- Situates readers, *"allows the reader to quickly visualize who participated, when, and where"*;
- Different way of seeing things, *"it is very interesting, it is a different way to look at things"*.

Considering the **negative feedback**, highlights include:

- Too many options, *"not sure if it is necessary to have so much control over the application"*;

- Unclear interactive elements, *"I would not understand what that button [to space out the events according to the timestamp] does if you had not told me"*;
- Redundant visibility levels, *"they are somewhat redundant, the first option [to hide the entities] does the same thing [as removing the entities entirely]"*.

Finally, considering the **most important features**, highlights include:

- Reset the visualization, *"need to have the option to reset, to go back to the original view"*;
- Select/highlight entities, *"the option to highlight entities was the most useful, it helps uncover certain narrative threads"*;
- Space out the events, *"[spacing out the events according to their timestamp] is important, it lets us know exactly which periods were the most intense"*;
- Event scroll, *"it is probably the most natural way to explore the narrative"*.

4.2. End-Users Survey

The survey was sent out to all members of the University of Porto, including students, and teaching and non-teaching staff members. Questions were divided into two main topics: news reading habits and demographic information, and analysis of the prototype. First, the respondents were shown two (appropriately annotated) screenshots of the visualization and scene information module, a brief video clip demonstrating some of the core features (e.g., scrolling through the events and dragging the event nodes around), and a brief textual description of the interface and its elements. Keeping in mind the time constraints, we chose to hide the sidebar, as its interest in a static context is quite limited. This introduction was followed by a series of questions related to the perceived importance of each element, which the respondent was asked to rate from one to five, one being "not important at all" and five being "extremely important".

The majority of participants (64.7%) were between 18 and 24 years old, followed by 28.2% between 25 and 39, and the remaining 7.3% between 40 and 59. This is in line with our expectations, as the survey was distributed using a university-wide mailing list targeting the University of Porto's members. The overwhelming majority of participants (93.2%) were students. Participants were asked to rank the importance of several visualization elements from one to five. The results can be see in Figure 3. Two elements stand out as the most important in the eyes of the participants: the events and the entities. The descriptions were ranked third in terms of relevance, which is in line with our expectations, as they provide crucial information on the events that make up the story. Surprisingly, the element which gathered the least amount of positive ratings is the title. The occurrences (i.e., the participation of entities in a particular event, encoded as a small colored circle inside of events) gathered the most neutral responses, which may indicate that their purpose is not obvious. All features had a majority of ratings above "four", which means we cannot single out any that may be unnecessary.

Participants showed more interest in using the visualization tool as opposed to editing the visualization, with 63% of participants rating their interest in using the prototype as a "four" or "five", in contrast with the 39% who provided those same ratings for the question related to editing. This is in line with what we expected, as editing the visualization is a feature that is perhaps more desired by power users. Moreover, the example we included in the survey did not lend itself to any major edits; there were no significant line overlaps and there were no obvious
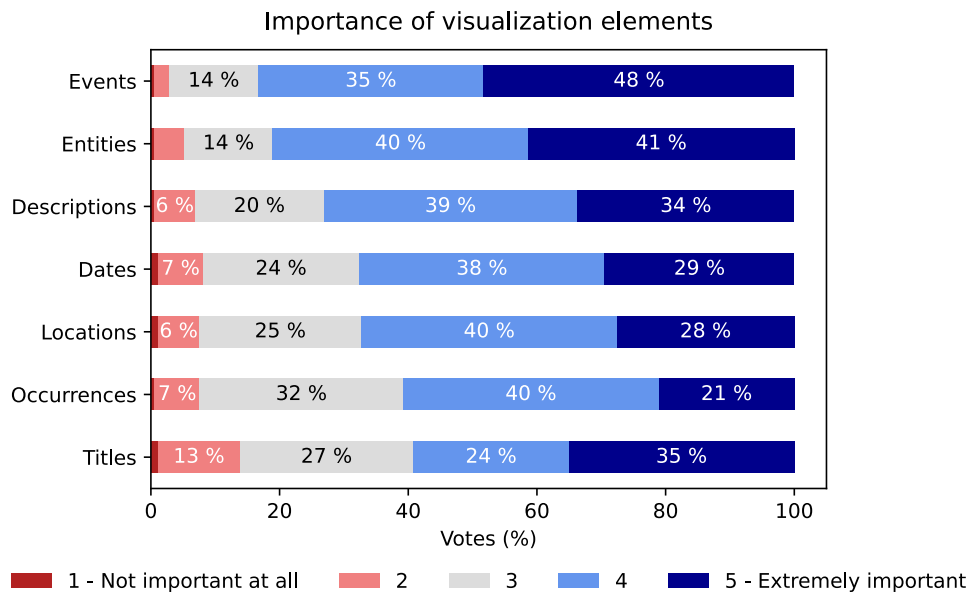
Figure 3: Importance of certain visualization elements.

problems with the visualization's readability. The need for manual adjustments may have been clearer if the visualization was more cluttered.

As for the relation between reading frequency and interest in the prototype, those who read the news more than three times a week show the highest percentage of "five" ratings, at 29%. Between those who read the news once a day and those who do so multiple times daily, the percentage of "three" and "four" ratings vary significantly, with those who read once a day providing 14% more "four" ratings and 16% less "three" ratings, with all other values remaining roughly the same between the two participant groups. This may indicate that participants who read the news multiple times per day do not benefit as much from visualizations focused on particular stories, since their consumption rhythm allows them to keep up with current news more easily. Additionally, people who read more often are already accustomed to the news sphere and may be more reluctant to changes. For people who read more than three times per week, the number of "four" ratings decreases, but the number of "five" ratings increases. For those who read between one and three times per week, the inverse happens. Finally, the participants who read the news less than once a week showcase the highest number of positive ratings; however, only 12 of the 178 responses belong to this group thus the ratings may not be representative.

The feedback collected corroborated the potential that storyline visualizations have in the study of journalistic narratives. From the 178 survey responses collected, 106 (60.6%) provided a rating of "four" or above when asked to quantify their interest in using the application. We verified that this result is stable across all age ranges. Overall, the participants generally found the application to be useful, but in need of some work in order for it to be more accessible to a broader public.

5. Conclusions

In this paper we present NewsLines, a browser-based prototype centered on a storyline visualization to highlight events and entities in a story. This visualization tool takes as input a JSON document detailing the events and entities of a story, and displays these elements in such a way that facilitates the study of the underlying narrative: how events relate to each other, how entities interact with one another, how events are spread out in time, and so on. NewsLine is made available as open-source software.

We conducted a first evaluation of the implemented solution with a small controlled user test and also with a large scale survey to the members of the University of Porto. The feedback collected both during the user tests and in the survey corroborated the potential that storyline visualizations have in the study of journalistic narratives. The journalists who participated in the user tests claimed that application provides an interesting perspective on news stories, one that is likely to be of interest both to the average reader, as well as to the journalist.

With the survey, we wanted to reach a broader audience and determine whether a visualization tool such as the one we developed would interest the general population. Given the brevity of the questionnaire, we were limited in terms of which parts of the visualization we could evaluate. Regardless, from the brief demonstration we included in the survey, the participants showed considerable interest in using the application as a complement to news articles. The comments left by some participants expressed concern about the usability of the application, which they found lacking in clarity and appeal.

After the development of the first version of the NewsLine prototype, future development should focus on user experience — addressing concerns raised during the conducted user tests and surveys, specifically the accessibility of the user interface, and the overall user experience. Another line for future exploration is to conduct users tests with journalists in real-word settings and test a broad range of news stories — exploring, for instance, which types of news stories are better candidates for a NewsLines based representation. Finally, we would like to integrate NewsLines in a text processing pipeline, where the intermediate JSON representation would be automatically extracted from news texts and thus quickly produce visualization for any set of news articles.

Acknowledgments

References

[1] E. Segel, J. Heer, Narrative visualization: Telling stories with data, IEEE Trans. Vis. Comput. Graph. 16 (2010) 1139–1148. URL: https://doi.org/10.1109/TVCG.2010.179. doi:10.1109/

TVCG.2010.179.

[2] K. Padia, Storyline Visualization Techniques for Linear, Non-linear, and Diegetic Narratives, Ph.D. thesis, North Carolina State University, Releigh, North Carolina, USA, 2019.

[3] R. Munroe, xkcd: Movie Narrative Charts, Available at https://xkcd.com/657/, 2009. [Online; accessed Jan 2023].

[4] S. Liu, Y. Wu, E. Wei, M. Liu, Y. Liu, Storyflow: Tracking the evolution of stories, IEEE Trans. Vis. Comput. Graph. 19 (2013) 2436–2445. URL: https://doi.org/10.1109/TVCG.2013.196. doi:10.1109/TVCG.2013.196.

[5] Y. Tanahashi, K. Ma, Design considerations for optimizing storyline visualizations, IEEE Trans. Vis. Comput. Graph. 18 (2012) 2679–2688. URL: https://doi.org/10.1109/TVCG.2012.212. doi:10.1109/TVCG.2012.212.

[6] Y. Shi, C. Bryan, S. Bhamidipati, Y. Zhao, Y. Zhang, K. Ma, Meetingvis: Visual narratives to assist in recalling meeting context and content, IEEE Trans. Vis. Comput. Graph. 24 (2018) 1918–1929. URL: https://doi.org/10.1109/TVCG.2018.2816203. doi:10.1109/TVCG.2018.2816203.

[7] M. Ogawa, K. Ma, Software evolution storylines, in: A. Telea, C. Görg, S. P. Reiss (Eds.), Proceedings of the ACM 2010 Symposium on Software Visualization, Salt Lake City, UT, USA, October 25-26, 2010, ACM, 2010, pp. 35–42. URL: https://doi.org/10.1145/1879211.1879219. doi:10.1145/1879211.1879219.

[8] M. John, S. Lohmann, S. Koch, M. Wörner, T. Ertl, Visual analytics for narrative text - visualizing characters and their relationships as extracted from novels, in: Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 2: IVAPP, (VISIGRAPP 2016), INSTICC, SciTePress, 2016, pp. 27–38. doi:10.5220/0005669800270038.

[9] S. Elvery, Narrative Charts (version 1.0.0) [source-code], Available at https://github.com/abcnews/d3-layout-narrative, 2016. [Online; accessed Jan 2023].

[10] N. Iskander, Comic Book Narrative Charts [source-code], Available at https://github.com/niskander/ComicBookNarrativeCharts, 2021. [Online; accessed Jan 2023].

[11] M. Costa, News Story Viz (version 1.0.0) [source-code], Available at https://github.com/marianafcosta/news-story-viz, 2021. [Online; accessed Jan 2023].

[12] V. Braun, V. Clarke, Using thematic analysis in psychology, Qualitative Research in Psychology 3 (2006) 77–101. URL: https://www.tandfonline.com/doi/abs/10.1191/1478088706qp063oa. doi:10.1191/1478088706qp063oa. arXiv:https://www.tandfonline.com/doi/pdf/10.1191/1478088706qp063oa.