Interactive Web-based Disaster Data Visualization

Aniket Panda, Anthony Pietrofeso, James Earl Johnson

Abstract— Disasters continuously impact both individuals and the world as a whole. Oftentimes, being able to visualize and interpret data associated with disasters is critical in establishing a response to disaster events, establishing patterns and understanding risks, and preparing and anticipating future events. To do this, we create a dynamic online tool for mapping, locating, and filtering various natural disasters and disaster metadata. Other tools exist, but these are often limited to specific datasets and offer limited exploration of disaster event metadata outside of the geographic context. Because of this, specific events can be geolocated and inspected, but it is difficult to draw general conclusions about the larger trends related to features such as temporal context, intensity, validity, causes, duration, and others for multiple disasters. For our project, we iterate on existing disaster data visualizations, and include multiple views to explore data in a spatial context that display multiple types of disaster and their locations. In addition, we also include scatterplot views for interpreting disaster event metadata in addition to the geographic context. We explore improving the quality of the visualizations by employing techniques we have seen in class such as the faceting of multiple views, working with layered data, and utilizing design principles and awareness of perception when designing our charts. This addition of multiple views and trend comparison was not utilized in existing tools, and we believe that combining these features with geospatial data increases the discoverability and visibility of trends in disaster data.

1 Introduction

Natural disasters have an immediate and unexpected impact on the lives of all who experience them. In this paper, we create a dynamic visualization to support sorting, locating, filtering, and trend analysis of geospatial disaster data to aid scientific advances in the field. Directly, scientists, researchers, and analysts can use this tool to identify larger trends in disaster data, preparing for and responding to natural disasters, as well as using the visualization elements and improvements as inspiration for future works and designs that can bring a new of understanding and interpreting the data.

Visualizations of disaster data often only explore a single dataset or a curated set of datasets. By supporting and allowing the ingestion of multiple types of datasets we enable processing and analysis of new, dynamic data. While our data is constrained by some fields that we key specific attributes on, other fields can be fully dynamic and analyzed if they contain numerical data.

Additionally, visualization websites for disaster data largely focus on geolocation and inspection of single events or groups of events. By leveraging additional views in addition to the map view, we allow users to answer more compelling questions. These questions would not just be motivated by what and where, but also address the how and when. Through our visualization we provide the visual richness to encourage the discovery of new trends and relationships that have previously gone unexplored due to a focus on specific events. Our visualization organizes views in both a geographic and geospatial context, as well as supplemental charts to decouple the data from the geospatial lens. In this context, a user can compare multiple attributes in a scatter plot to aid in trend analysis.

The aims of this research are:

- Utilizing existing disaster event visualization techniques and developing and surfacing new techniques for visualizing disaster event metadata backed by visualization research

- Combining these techniques on a deployable online web application that supports user interaction and multiple views.

- Allowing for ingestion of new datasets in semi-dynamic formats and displaying the data in the tool.

• *Aniket Panda, Anthony Pietrofeso, James Earl Johnson are graduate students at the University of Arizona. aniketpanda100@arizona.edu, appietro516@arizona.edu, johnsjam@arizona.edu*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

2 Background

Throughout history volcanoes, earthquakes, and hurricanes are just a few of the natural disasters that have affect a human population. These powerful disasters have caused many people to die, become displaced, or caused material damage. Visualization of these natural disasters helps in mitigating these effects, by educating and informing the people, as well as enabling further analysis of the data. It helps in guiding policies to help with living through a disaster or avoiding it all together.

There are many visualizations of natural disasters, with many of them focusing on one type of natural disaster or another. Some natural disasters have relationships with other natural disasters. For example some high magnitude earthquakes can cause Tsunamis. Or a hurricane could cause flooding. The target audience for this paper would be scientists who study natural disasters and governments need to make informed decisions to react the threat of multiple natural disasters. This paper focuses on a visualization that displays multiple maps with natural disasters on it. This allows for the focus of natural disasters to be done in a more global sense, because most natural disasters are location based. Our visualization also contains the ability to show the combined relationships of two maps with natural disasters onto a single map (located in quadrant 4). We were inspired from using visualization tools such as the natural hazards viewer has helped guide the layout and function of our visualization [3].

2.1 Related Work

Many tools exist for browsing and filtering data related to natural disasters. One such tool is the Natural Hazards Viewer maintained by the National Centers for Environmental Information [3]. The tool features a single view showcasing a world map. The user can zoom and pan this map to view different areas of the world. Different types of disasters are indicated by differently shaped symbols on the map, and events can be filtered on the left sidebar. Clicking a symbol creates a popup displaying the metadata for the disaster event that the map item represents.

The view we create is similar to the National Hazards Viewer for our map view, pictured below in Figure 1.

Other examples of recent, officially produced visualizations abound online. Many such visualizations are more limited than our project is going to be. They tend to satisfy the design principle of graphical integrity, but under perform when it comes to graphical excellence. Consequently, they serve as benchmarks we iterate on and improve.

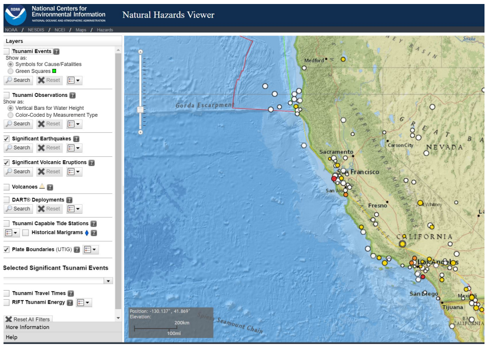

Fig. 1. The National Hazards Viewer

One source of such visualizations is FEMA's online data visualization resource [1]. It includes various colored maps that are provided for data such as the number of fire incidents per county and historical flood risk cost. However, more detailed data such as the cause of the fires and change in the number of flood events over time are expressed in separate bar charts and line graphs below the maps. This prevents the user from easily making interesting comparisons of the more detailed data from one county to another because a bar chart or line graph can only appear for one county at a time.

Another site, by Our World in Data, includes colored maps that convey information such as total number of natural disaster deaths and natural disaster deaths as a share of total deaths [8]. However, these are on separate maps, meaning the two variables cannot be evaluated together. Therefore, the interaction between the two variables is difficult to understand.

In Natural Disaster Visualizations by Alex Martinez, colored maps allow a user to switch between viewing total incidents, total deaths and total damages [7]. However, users can only view one map at a time, so evaluating the relationship between the variables is, once again, difficult. A stacked bar chart is provided that shows the breakdown of different disaster types over time, but these are global data and their spatial distribution cannot be expressed on the map itself.

We believe the above is a limitation because a single view could become susceptible to overcrowding and chart junk, while also making it difficult to compare the relationship between separate variables without encoding them in a separate channel (which can be difficult on a 2D map). We observe problems in class around this, where we start to see obfuscation of items on the map that are traditionally solved through filtering or semantic zooming, which restricts available information and makes it difficult to evaluate data in an n-ary relationship. For non-spatial data, we can solve this with a SPLOM or parallel coordinates chart, but would lose the geographic element unique to cartography, which is often a highly valuable feature.

In the 2018 paper, "Hyper-resolution monitoring of urban flooding with social media and crowdsourcing data" Wang et al. describe new methods for collecting, analalyzing, and visualizing flood data using social media [9]. While visualizing flooding is very specific compared to general disaster data, Wang's visualizations on road closure and flooding mapping as well as spatial patterns of precipitation/tweets can generalize to other types of disasters and be used as inspiration for map layers.

Additionally, "Multidimensional Visualization and Processing of Big Open Urban Geospatial Data on the Web" discusses 3d approaches for visualizing geospatial data [6]. While we don't plan on adding 3d elements, we can use this as inspiration for possible 2.5d elements (such as map markers), and can adopt elements of their interface for tagging and displaying data.

Finally, we draw on the design of our application through Wood et al.'s description of a "Mashup" in a 2007 paper. The authors believe that there is an "analogy between mashing technologies to produce an application and mashing views to explore data. Both involve flexible synthesis: one of data and functionality, the other of visual encodings of data through interactions" [10]. Through this synthesis, we believe we can create a richer visualization and a more functional application, especially when accounting for technological advancements that have emerged with the evolution of web3 technologies since the inception of the paper.

3 Building the Visualization Tool

In planning our visualization tool, we investigated other natural disaster tools such as the Natural Hazards Viewer [3] to help us with designing the visualization. During this investigation period we looked into what data we could use for our tool. We also researched papers that had any type of discussion on visualizing natural disasters. This helped with our ideas for our brainstorming session.

Our brainstorming sessions benefited from our investigation, and we created hand drawings of our website. From the hand drawings we created a mock up to help us design the prototype. We kept the scope of the prototype to visualize earthquake data for the United States. We wanted to display the size of the magnitude of the earthquake data using the size of the circle.

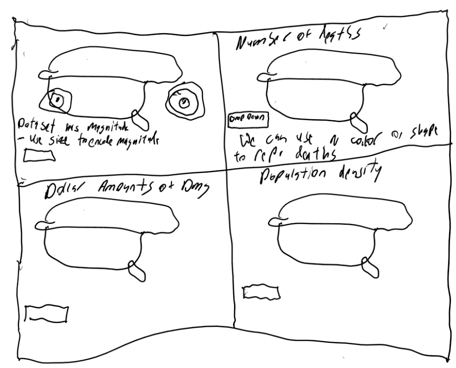

Fig. 2. Hand drawing of Website

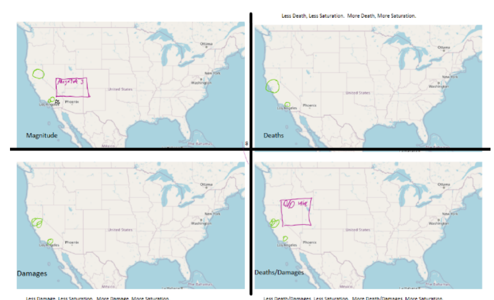

Fig. 3. Mock up of Website

When it came down to implementation we looked into APIs that we could use that would allow us to create a map and add glyphs to it. We settled on using the OpenLayers API because of its many features. It also integrates well with d3. Because the team was unfamiliar with OpenLayers we needed to use many examples to help with our implementation of the prototype.

The goal of the prototype was to give us a good foundation of what we needed to build next. In the prototype we create a 4 Maps that could be viewed at the same time. We made the prototype load a small subset of the earthquake data and we were able to demonstrate the earthquake data varied in magnitude at different points in the map.

From there we moved onto the final website. We built on top of the prototype and added more features for the user to be able to view the natural disasters data. We added a fields drop down to choose other fields besides magnitude. We added a drop down to choose which disaster to look at. We added a drop down to choose

3.1 Data

The team used the EM-DAT Query Tool [2] to retrieve our data. For our tool we using Earthquake, Coastal Flood, Landslide, Tsunami and Tropical Cyclone data. The EM-DAT Query tool exports a csv that our visualization reads to populate the data.

3.2 Technology

We used Javascript, D3, SVG and OpenLayers to implement the website [5]. The website was built and tested with Chrome and Edge browsers. D3 and SVG were good tools for making the visualizations for the website and we took advantage of the versatility of D3 when we built our prototype visualization for the website. OpenLayers is was a convenient library to create maps with because it can be integrated with D3.

3.3 Group Member Roles

All of our team members worked together to design the interactive website. During the implementation of the prototype and final website we ended up dividing the prototype and website based on features. We identified our tasks as user-facing (UX/Design/implementation of maps) and backend facing (data loading/website infrastructure/"API-esque" interactions).

Anthony Pietrofeso - Back-end
Aniket Panda - Front-end
James Johnson - Front-end

4 Impacts

Natural disasters are phenomena that have long burdened humanity and will continue to do so far into the future. Consequently, their status as a major source of damage, death and disruption will likely remain a subject of public interest moving forward. These disasters interact with societal interests in highly varied and complex ways. However, these nuances are often difficult to parse because most current visualizations of natural disaster data can only visualize a few variables simultaneously. This limits the kinds of interesting patterns and relationships that can be observed.

For example, consider the limitations of the Natural Disaster Visualizations map which were discussed in Section 2.1 [7]. Many interesting socioeconomic questions that could be asked if said limitations were lifted. For example, are total deaths and total damages correlated? If not, a poignant observation about how natural disasters tend to kill residents of poorer areas could be made. In turn, a conversation about how we should measure the effects of a disaster could be started.

Of course, these questions can still be asked, but not as strikingly. Being able to make a point visually makes it more quickly communicable and more memorable. In turn, this increases the likelihood that the point will stick with enough people that an actual change in attitude, policy or behavior is affected. The reason these visualizations have not already been made is not due to a lack of interest. The challenge is that expressing more multivariate data increases the risk of the visualization becoming too cluttered and difficult to read. This is why careful visual design is needed.

Our project reveals interesting relationships between natural disaster data to the scientific community, but more importantly, to non-scientific users. At the very least, the project will be of interest to other developers looking for ways to improve, or not to improve, their own visualizations.

5 Results

We are very happy with our results. The time and effort invested in choosing richly detailed data, flexible yet powerful tools, and insightful visualization techniques has paid off. Our application includes all the features we planned for it to and performs well on all the evaluation criteria. Much of what makes the application useful is best understood by using it, but it is still worth describing here.

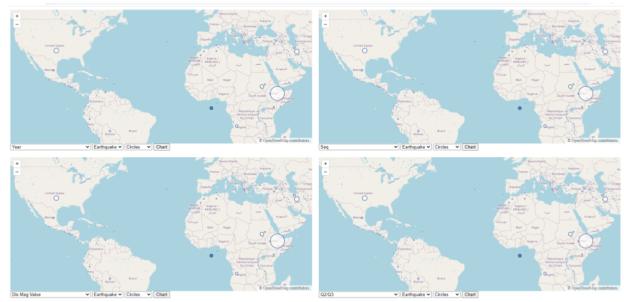

Fig. 4. A small multiples view of four different attributes.

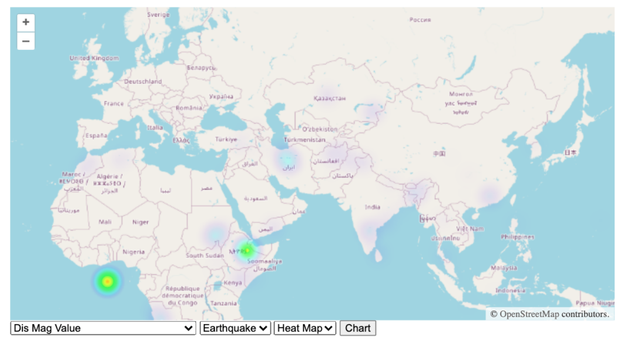

Fig. 5. A quadrant showing a heat map rather than circles.

5.1 Features

Below is a list of features that helps summarize the functionality of our interactive visualization.

1. Small Multiples The visualization simultaneously displays four maps laid out in different quadrants of the screen, as shown in Figure 4. Each map can represent a different data attribute. This allows multivariate geospatial data to be expressed more effectively by reducing the burden of the user's working memory and allowing each attribute to be expressed with the most effective channels, rather than some being relegated to less expressive ones.

2. Linked Views The maps are linked such that any zooming or dragging interactions with one map are immediately reflected in all four. Figure 4 shows this. This facilitates the user's ability to perform map comparison tasks by preserving which areas of the globe are being displayed even as the data attributes and markers are changing.

3. Attribute Selection The user can set each map to display any quantitative attribute they wish. The quadrant each attribute can be displayed in is not fixed. Additionally, a user can display the same attribute in multiple quadrants, which may be useful if they wish to view said attribute with different types of markers, with different time ranges, or with one as a map and the other as a scatterplot.

4. Direct Attribute Comparison The map in the bottom right quadrant has an option in the attribute selection dropdown menu called Q2/Q3. This causes the map to directly portray the relationship between the attributes in the top right and bottom left quadrants. This is more effective than side-by-side comparison.

5. Heat Map If they prefer, the user can make any map a heat map rather than using circles as markers. This makes disasters' areas of effect a continuous gradient rather than a collection of discrete zones. An example is shown in Figure 5.

6. Scatterplot Individual maps can be switched into scatterplots that portray the selected data attribute over time. This is useful when temporal patterns are more important to understand than geospatial ones. An example is shown in Figure 6.

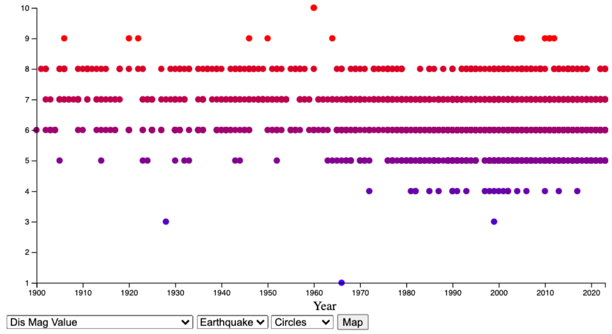

Fig. 6. A scatterplot showing earthquake magnitudes over time.

7. Filtering The user can filter the data being displayed by the type of disaster using a drop down menu, as well as change the time range over which data is displayed by using a time slider.

8. Tooltips Hovering over a disaster marker brings up a tooltip allowing the user to inspect individual data points in a more detailed manner.

5.2 Evaluation

It was important that our visualization satisfied the design principles of graphical integrity and graphical excellence discussed in class. There is a natural tension between these principles that underlies the motivation of this project. The visualizations discussed in Section 2.1 succeed in satisfying graphical integrity, but we feel that they underperform when it comes to graphical excellence. Our work improves the latter without compromising the former.

The visualizations discussed in Section 2.1 serve as baselines against which our work can be evaluated. They are good baselines because they are similar to our project. That is, they represent the same kind of data, are interactive, and prominently feature colored maps. Critically, they are also from professional or official sources and are recent, meaning they are a fair standard to evaluate our work against.

To be more specific regarding the evaluation criteria, we used a set of action target pairs to describe tasks a user might perform using our visualization. We used these pairs to compare the effectiveness of our visualization with the similar maps discussed in the Related Works section. The main advantage of our small multiples approach over the single view format used by other maps is meant to be the ease with which one can compare spatial patterns, such as trends, correlations, and shapes, among different data attributes.

For example, if a user wishes to examine the relationship between the damage, in dollar amounts, and the number of deaths, our visualization allows them to do so more easily compared to other visualizations. Figure 4. shows this. Our visualization is clearly superior to FEMA's online data visualization resource [1], Our World in Data [8], Natural Disaster Visualizations by Alex Martinez [7]. This superiority comes from the fact that small multiples views reduce the burden on the user's working memory.

At the same time, our visualization is superior to the Natural Hazards Viewer [3]. The latter, despite being detailed and holistic, is also cluttered and lacking in clarity. Since it is a single view map, it must relegate each additional attribute it tries to express to less effective channels. In contrast, each attribute displayed in our visualization is expressed with the most effective channels besides position, namely size and color.

In addition to comparing attributes, the user can discover spatial trends in the relationships between two attributes by viewing the map that plots said relationship explicitly. This is even more directly expressive than side-by-side comparison of two maps, so it is a vast improvement over the other visualizations.

The two tasks above, and our advantages in executing them, can be extended from the spatial domain to the temporal domain through the use of the feature which toggles each map with a scatterplot that shows the chosen attributes over time. Although FEMA's online data visualization resource [1], and Our World in Data [8] have features that express temporal trends, they suffer from the same kinds of single view limitations previously discussed.

Finally, our visualization helps summarize the trends in the data more effectively than its single view counterparts. Again, this advantage comes from reducing the burden on the user's working memory to make it easier for them to take in the multivariate data holistically.

Of course, there are numerous other tasks our visualization allows for. However, it is unnecessary to evaluate them because they are not relevant to the shortcomings of other visualizations that we specifically set out to improve upon.

We performed the evaluation by ourselves. Though the use of an online survey, as discussed in the project proposal, would be ideal, it was not practical given our time constraints. Furthermore, it was unlikely to provide a great deal of benefit. Our project relies on well-established design principles, not new or unusual ones. The novelty comes from the fact that this particular design has not been applied to this particular type of data. Using our own judgment is sufficient.

Performance considerations, such as how quickly the webpage loads and how responsive the interface is, have not been major issues so far. To the extent that they are, they are no worse than those of the visualizations we have been comparing our map to.

5.3 Running the Code

To run the code for the project, install nodejs/npm [4] if not already installed. Then, enter the directory and input the following into the command line:

```
git clone https://github.com/Appietro516/disaster-app-project.git
npm i
npm start
```

Open the URL provided in a browser. Chrome is recommended.

6 Conclusion

Overall, we successfully accomplished our goal of creating an interactive web based disaster visualization using publically available dynamic data.As mentioned in the evaluation section, we assess our tool on various action target pairs that the user can -perform using the tool. While our tool is largely based on the research that we conducted, we make a few critical improvements to existing geospatial data tools that allow for comparison of attributes and enabling trend discovery.

A key component of our visualization tool is the four panel "small multiples" linked views. Similar visualization tools we experimented with contained a single view. This makes it easy to gather information about specific events and clusters of events, but difficult to understand trends in overall data.For example, if we wanted to compare the trend of earthquake damage with magnitude, this would be difficult with existing tools.A second and third panel easily allow users to compare trend data for these attributes because we can reuse existing encodings on the other panels without creating visual fatigue or noise for the viewer. We believe that this simple adaptation will enable more dynamic comparisons for disaster data.

Additionally, our tool locates specific fields in the dataset for user input and data filtering, such as the "Disaster Type" for some chart controls.However, for a majority of our data, we search for dynamic numeric attributes within the dataset and display a normalized encoding of the value in the viewer. Because of this, our tool could easily extend to many types of geospatial data with minor tweaks to the code. While we demo as a "disaster data" visualizer, the tool can be generalized to many different types of datasets and still enable adequate trend discovery in these datasets. Because of this, we make some sacrifices to the visual fidelity of the tool, such as encoding all attributes using the normalized size of the circle. But we believe that the ability for processing more dynamic data would add value to future users, and we take reasonable precautions, such as normalizing the data and giving the user a legend of the magnitude.

We also introduce the ability to directly compare 2 attributes in the 4th quadrant. This is more experimental, but allows a user to understand how 2 fields relate to each other. For example, if 2 fields are relatively correlated, we will see similar sized circles. If 2 fields are inversely related, we will see very small or very large circles, or a mix of both if there is no relation. This is another technique that allows the user to "merge" multiple views to get a better understanding of if a relationship is present

We also include scatterplot and heatmap views that provide additional ways of exploring the data. We found that it was sometimes helpful to remove the data from the geospatial context, and existing tools often do not allow for an alternative chart. In addition, we provide a heatmap as another encoding of the magnitude. The heatmap allows us to more directly observe the magnitude of the effect and decouple it from individual events.

Because of this, we believe that users have a more detailed understanding of datasets using the tool and will be more richly able to explore data than with existing interfaces.

6.1 Lessons Learned

We learned several lessons during the development of our application. Initially, we found it difficult to onboard with so many new tools, as we had not worked with many of these libraries in the past, such as openlayers. There was a steep learning curve that made it difficult for us to get the initial prototype created. For future projects, we may want to narrow our focus on a few tools, or better parallelize the work overall.

Additionally, we spent large amounts of time focusing on the design of our project, we believe this was beneficial because we were able to think through and plan the key elements of the visualization. Because we spent a large time focusing on the design of the visualization, we did not spend as much time planning the implementation of the code. We believe we should have given this equal weight to design, and could have factored out multiple key abstractions or modules that would have made implementing the final visualization easier.

One final lesson we learned is that we should verify the robustness of the data prior to building a project around it. We initially tested our design with the EM-DAT earthquake data only. Later on, we introduced additional disaster types and found that the longitude and latitude fields were undefined for most of these fields. At this point, we faced a choice of spending additional effort sourcing new data for the project (and refactoring our existing implementation) or continuing to use the data we had. We ran into additional data quality issues when we realized the dataset is very sparsely populated for some attributes, which lead to a less interesting visual comparisons and made it tougher to identify relationships with a small n. It is unknown if this was lacking in the dataset, or if specific pieces of information are not collected for all disasters. Ultimately, we chose to prioritize the functionality of the visualization so we could create a proof of concept for our ideas, as the earthquake data was sufficient to demonstrate the key ideas.

6.2 Future Direction

We have several ideas for future directions. Firstly, to understand the tools real-world potential for impact, we would need to deploy the tool to a permenent online website. This would require additional systems such as authentication, dedicated hosting, and more rigorous security features for uploading data. Because of these additional requirements, we decided not to implement this in this iteration of the project which acts as a minimum viable product.

For enhancing the visualization, we would also like to explore additional encodings for the data that may make more sense for specific attributes, and ways to implement those encodings. For example, we initially considered using different shapes for the different types of disasters and proceedings with a more glyph-based visualization, however, we decided against this because we found it was easier to perceive the size of attributes relative to each other.

Because the data is semi-dynamic, we also explored other directions for sourcing the data from a real-time API and optimizing the display of a large number of points on the map. We identified this as a stretch goal in our initial proposal, but found implementation would have been out of scope when paired with the rest of the implementation of the tool. Because this did not directly enhance the data visualization, we decided not to pursue this direction and focused on the visual quality of the tool. However, we believe that live data could add a larger degree of usefulness to the tool, and would want to continue in this direction if we continued to work on the project. To be truly useful in an application setting, we also desire to be able to display datasets that are larger than our sample dataset size (1 MB) and would need to optimize the tool to run more efficiently with larger amounts of data.

References

[1] Fema data visualizations. `https://www.fema.gov/about/reports-and-data/data-visualizations`. Accessed: 2023-10-02.

[2] The international disaster database. `https://emdat.be/`. Accessed: 2023-18-02.

[3] Natural hazards viewer. `https://www.ncei.noaa.gov/maps/hazards/`. Accessed: 2023-12-02.

[4] Nodesource node.js binary distributions. `https://github.com/nodesource/distributions`.

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

[6] C. E. Kilsedar and M. A. Brovelli. Multidimensional visualization and processing of big open urban geospatial data on the web. *ISPRS International Journal of Geo-Information*, 9(7):434, 2020.

[7] A. Martinez. Visualization: The international disaster database. `https://sites.uci.edu/alexm/2019/07/14/visualization-the-international-disaster-database`. Accessed: 2023-12-02.

[8] H. Ritchie, P. Rosado, and M. Roser. Natural disasters. *Our World in Data*, 2022. https://ourworldindata.org/natural-disasters.

[9] R.-Q. Wang, H. Mao, Y. Wang, C. Rae, and W. Shaw. Hyper-resolution monitoring of urban flooding with social media and crowdsourcing data. *Computers & Geosciences*, 111:139–147, 2018.

[10] J. Wood, J. Dykes, A. Slingsby, and K. Clarke. Interactive visual exploration of a large spatio-temporal dataset: Reflections on a geovisualization mashup. *IEEE transactions on visualization and computer graphics*, 13(6):1176–1183, 2007.