FROM EXPLORATORY TO EXPLANATORY INTERACTIVE VISUALIZATION OF SPATIO-TEMPORAL CONFLICT DATA

BENEDIKT MAYER

DISSERTATION

zur Erlangung des akademischen Grades
DOKTORINGENIEUR (DR.-ING.)

angenommen durch die Fakultät für Informatik
der Otto-von-Guericke-Universität Magdeburg
von M.Sc. Benedikt Mayer
geb. am 18.01.1994 in Schwabmünchen

Gutachter:
Prof. Dr.-Ing. Holger Theisel
Prof. Dr.-Ing. Christian Tominski
FH-Prof. Priv.-Doz. Dipl.-Ing. Dr. Wolfgang Aigner

Magdeburg, den 24.09.2024

There comes a point where we need to stop just pulling people out of the river.
We need to go upstream and find out why they're falling in.

— Desmond Tutu

ABSTRACT

Violent conflicts cause human suffering and devastation worldwide. Conflict research studies the causes, dynamics, and impact of such conflicts, as well as ways to prevent them. Interactive data visualization can support conflict researchers in exploring corresponding data. However, so far, barely any studies have been conducted analyzing how to provide this kind of exploratory support in the field of conflict research.

In contrast, interactive data visualization is already commonly used in online stories reporting on conflicts, as it can help to explain space- and time-dependent developments more clearly. However, there is little structured research of how such data-driven visual storytelling can be performed for stories with a spatio-temporal context, like those about conflicts.

Exploratory and explanatory approaches can also be combined, allowing users to get a deeper understanding of the conveyed information. While this approach is more commonly used to target broad audiences in the context of science communication, it can also be used to communicate scientific information to researchers. However, little research has been performed in this direction as well.

To address the listed issues, we collaborated with a conflict researcher to conduct three main studies. They are positioned along a spectrum between exploration and explanation to answer the question: "How can interactive visualization be used to support the exploration and explanation of spatio-temporal conflict data?"

In the first study, we investigated how to design an application that allows conflict researchers to explore the results of a process in which multiple conflict data sets are integrated into one holistic set. We abstracted the underlying domain problems and derived a workflow and application design to solve the identified problems. An evaluation involving five conflict researchers showed that the application helped them to better understand and validate the results of the data integration. However, it also showed that even experts require a considerable amount of explanation to properly use such exploratory applications.

In the second study, we investigated how to combine exploration and explanation to communicate the workings of a scientific method to conflict researchers, based on a visual data story. We distilled the story creation process and evaluated the story with eight conflict researchers. It showed that they accepted the storytelling approach well, and their feedback allowed us to derive a set of guidelines for performing such exploratory explanation for experts.

In the third study, we collected 130 data stories with a spatio-temporal context to investigate how visual storytelling was applied in them for explaining the underlying messages. To do so, we merged and adapted three existing design spaces and used

them to classify the stories based on which storytelling techniques they employed. We analyzed the resulting data, identifying various ways to combine the storytelling techniques as well as trends that occurred over the years, e.g., towards more easily consumable stories, and towards updating and extending stories over time.

Overall, our studies showed several promising directions for how interactive visualization can support the exploration and explanation of spatio-temporal conflict data. However, it also became clear that more research is necessary. Accordingly, it should be investigated how to make exploratory applications more accessible through explanation, how to make the process of creating spatio-temporal data stories more efficient, and what the role of interaction in explanatory applications can be.

ZUSAMMENFASSUNG (SUMMARY IN GERMAN)

Gewaltsame Konflikte verursachen weltweit Leid und Zerstörung. Konfliktforschung untersucht die Ursachen, Dynamiken und Auswirkungen solcher Konflikte sowie Möglichkeiten, sie zu verhindern. Interaktive Datenvisualisierung kann die Erforschung entsprechender Daten unterstützen. Bislang gibt es jedoch kaum Studien, die sich mit der Frage beschäftigen, wie eine solche explorative Unterstützung im Bereich der Konfliktforschung aussehen kann.

In Online-Stories über Konflikte werden interaktive Datenvisualisierungen dagegen bereits häufig eingesetzt, um raum- und zeitabhängige Entwicklungen zu erklären. Allerdings gibt es wenig strukturierte Untersuchungen darüber, wie ein solches datengetriebenes Visual Storytelling für Geschichten mit einem raumzeitlichen Kontext, den Berichte über Konflikte häufig besitzen, durchgeführt werden kann.

Explorative und erklärende Ansätze können auch kombiniert werden, um Anwender*innen ein noch tieferes Verständnis von vermittelten Informationen zu ermöglichen. Dieser Ansatz wird häufiger verwendet, um im Rahmen von Wissenschaftskommunikation ein breites Publikum anzusprechen, er kann aber auch für die Vermittlung wissenschaftlicher Informationen an Forschende genutzt werden. Allerdings wurden auch in dieser Richtung bisher nur wenige Untersuchungen durchgeführt.

Um die genannten Probleme zu untersuchen, haben wir in Zusammenarbeit mit einem Konfliktforscher drei zentrale Studien durchgeführt. Sie sind entlang eines Spektrums zwischen Exploration und Erklärung angeordnet, um die Frage zu beantworten: „Wie kann interaktive Visualisierung genutzt werden, um die Exploration und Erklärung von raumzeitlichen Konfliktdaten zu unterstützen?“

In der ersten Studie haben wir untersucht, wie eine Anwendung gestaltet werden kann, die es Konfliktforschenden ermöglicht, die Ergebnisse eines Prozesses zu erforschen, bei dem mehrere Konfliktdatensätze zu einem ganzheitlichen Datensatz integriert werden. Wir haben die zugrundeliegenden, fachspezifischen Probleme abstrahiert, um einen Workflow und ein Anwendungsdesign zur Lösung der identifizierten Probleme abzuleiten. Eine Evaluierung mit fünf Konfliktforschenden hat gezeigt, dass die Anwendung ihnen dabei geholfen hat, die Ergebnisse der Datenintegration tiefergehend zu verstehen und zu validieren. Allerdings hat sich auch gezeigt, dass selbst Expert*innen einen erheblichen Erklärungsbedarf haben, um solche explorativen Anwendungen angemessen nutzen zu können.

In der zweiten Studie haben wir untersucht, wie Exploration und Erklärung kombiniert werden können, um Konfliktforschenden die Funktionsweise einer wissenschaftlichen Methode anhand einer Visual Data Story zu vermitteln. Wir haben den Erstellungsprozess der Geschichte abstrahiert und sie mit acht Konfliktforschenden evalu-

iert. Es hat sich gezeigt, dass die Forschenden den Storytelling-Ansatz positiv aufgenommen haben, und aus ihrem Feedback konnten wir eine Reihe von Leitlinien für die Erstellung solcher explorativer Erklärungen für Expert*innen ableiten.

In der dritten Studie haben wir 130 Data Stories mit einem raumzeitlichen Kontext gesammelt, um zu untersuchen, wie sie Visual Storytelling nutzen, um Zusammenhänge zu erklären. Zu diesem Zweck haben wir drei bestehende Design Spaces zusammengeführt und angepasst, um die Geschichten anhand der jeweils in ihnen verwendeten Erzähltechniken zu klassifizieren. Wir haben die daraus resultierenden Daten analysiert und verschiedene Möglichkeiten zur Kombination der Erzähltechniken identifiziert, sowie Trends, die im Laufe der Jahre aufgetreten sind. So konnten wir zum Beispiel Trends hin zu leichter konsumierbaren Geschichten beobachten, ebenso wie hin zur Aktualisierung und Erweiterung von Geschichten über den Verlauf mehrerer Ereignisse hinweg.

Insgesamt konnten wir in unseren Studien mehrere vielversprechende Wege aufzeigen, wie interaktive Visualisierung die Exploration und Erklärung von raumzeitlichen Konfliktdaten unterstützen kann. Es wurde jedoch auch deutlich, dass weitere Forschung notwendig ist. Entsprechend sollte untersucht werden, wie explorative Anwendungen durch Erklärungen zugänglicher gemacht werden können, wie die Erstellung raumzeitlicher Data Stories effizienter gestaltet werden kann und welche Rolle Interaktion in erklärenden Anwendungen spielen kann.

CONTENTS

Part I

FOUNDATIONS

We motivate the research question of the dissertation and provide background information required for the core chapters.

1 INTRODUCTION

In the recent years, the human cost of violent conflicts has drastically increased. This is not least driven by a rise in large-scale conflicts not seen since the end of the Cold War that also directly affect Western societies, including the war of Russia against Ukraine and the Israel-Hamas war. However, in addition to these conflicts, which receive a lot of coverage in Western media, there have also been many other conflicts around the globe, such as the Yemeni civil war and the Tigray war. According to the UNHCR, by September 2023, more than 100 million people worldwide were displaced because of war, violence, persecution, and human rights violations [Unh].

Peace and conflict research strives to understand the causes and dynamics of such conflicts, as well as potential ways to end them or to prevent them in the first place, for instance, via aid projects. Statistical approaches are used to analyze data reporting on conflicts and derive insights from them. Visualization research has shown that interactive visualization solutions can support such analytical processes. These solutions can provide more direct access to the data and ways to explore them in different levels of detail, facilitating the creation and validation of hypotheses. However, corresponding applications of exploratory visualization are still sparse in conflict research.

Aside from researchers, also the broad public wants to learn about conflicts, although typically not on such a technical level as researchers who have deep domain expertise. Rather, broad audiences are interested in the key facts about the most recent developments, or a summarized version of research insights and experts' opinions explaining the developments. For this purpose, data visualization is already a well-established tool. For instance, it is commonly used in data-driven online stories produced by newspapers such as *The New York Times* and *The Guardian*. In this area, visualization researchers are trying to catch up with the industry developments, analyzing patterns prevalent in the stories produced by practitioners to distill the most effective approaches. However, regarding stories based on spatio-temporal data, visualization research is still lacking, especially when it comes to conflict data stories.

Authors of data-driven online stories aim at making the communicated content easily understandable, engaging, and memorable. Techniques to do so include using story arcs, like the Freytag's Pyramid, progressively building up views, and visually abstracting information. Such explanatory storytelling techniques are typically rather used to target broad audiences. However, we argue that the communication between researchers can also benefit from these techniques. Especially when they are combined with more exploratory approaches, these storytelling techniques can improve the efficient and engaging transfer of knowledge. For instance, such stories can be used to explain the workings of a scientific method to researchers who are unfamiliar with it. Yet, little research exists on how such forms of communication between researchers can be supported in practice, and what potential pitfalls exist in this context.

To summarize, the application of interactive visualization in the context of conflict data is still underexplored. Several directions exist, ranging from exploration- to explanation-focused approaches. In this work, we follow three main directions, namely

- exploration for conflict research experts,
- the combination of exploration and explanation for conflict research experts, and
- explanation for broad audiences.

We present three central research projects to address them. We summarize the corresponding research contributions below.

1.1 RESEARCH CONTRIBUTIONS

We contribute to the interdisciplinary field of data visualization by analyzing how to answer the question:

"How can interactive visualization be used to support the exploration and explanation of spatio-temporal conflict data?"

The scientific contributions across our three main projects are the following.

[APPLICATION & SYSTEM] Exploring and validating the results of a statistical method used to integrate multiple conflict data sets (see Chapter 3: EXPLORATION FOR EXPERTS).

- We characterize a collection of conflict data sets and a method used to integrate them into one holistic data set. We then abstract the domain problems that arise when trying to validate the correctness of the integration.
- We translate the domain problems into the field of interactive visualization and abstract a set of analytical tasks and an application workflow from them.
- We present the design and implementation of an application following the derived workflow to solve the identified tasks.
- We evaluate the application with five conflict researchers and analyze the corresponding results.

[APPLICATION] Using exploratory explanation for communicating the workings of a statistical method to conflict researchers (see Chapter 4: EXPLORATORY EXPLANATION FOR EXPERTS).

- We summarize the workings of a statistical method used to test causal relationships between different types of conflict events. We then present our process for designing an interactive data story to explain the method to conflict researchers.
- We evaluate different versions of the story with eight conflict researchers.
- We summarize our findings regarding how exploratory explanation can be performed to communicate scientific methods to researchers.

[Empirical Study & Design Space] Characterizing and analyzing a collection of explanatory interactive visual data stories with a spatio-temporal context (see Chapter 5: EXPLANATION FOR BROAD AUDIENCES).

- We introduce a collection of 130 web-based interactive visual data stories with a spatio-temporal context published between 2018 and 2022.
- We present a design space to characterize the 130 stories. The design space was created by combining and adapting three existing design spaces.
- We analyze the stories, deriving patterns and trends, e.g., which storytelling techniques were frequently used and how similar the stories are to each other.

1.2 PRIOR PUBLICATIONS AND AUTHORSHIP

List of first author publications related to conflict data:

[May+21] B. Mayer, K. Lawonn, K. Donnay, B. Preim, and M. Meuschke. "VEHICLE: Validation and Exploration of the Hierarchical Integration of Conflict Event Data." In: *Computer Graphics Forum* 40.3 (2021), pp. 1–12. DOI: 10.1111/cgf.14284.

[May+23b] B. Mayer, N. Steinhauer, B. Preim, and M. Meuschke. "A Characterization of Interactive Visual Data Stories With a Spatio-Temporal Context." In: *Computer Graphics Forum* 42.6 (2023), e14922. DOI: 10.1111/cgf.14922.

[May+24] B. Mayer, K. Donnay, K. Lawonn, B. Preim, and M. Meuschke. "Expert exploranation for communicating scientific methods - A case study in conflict research." In: *Computers & Graphics* 120 (2024), p. 103937. DOI: https://doi.org/10.1016/j.cag.2024.103937.

List of first author publications related to medical data:

[May+23a] B. Mayer, M. Meuschke, J. Chen, B. P. Müller-Stich, M. Wagner, B. Preim, and S. Engelhardt. "Interactive visual exploration of surgical process data." In: *International Journal of Computer Assisted Radiology and Surgery* 18.1 (2023), pp. 127–137. DOI: 10.1007/s11548-022-02758-1.

List of contributing author publications related to medical data:

[Kos+24] G. Kostiuchik, L. Sharan, B. Mayer, I. Wolf, B. Preim, and S. Engelhardt. "Surgical Phase and Instrument Recognition: How to identify appropriate Dataset Splits." In: *International Journal of Computer Assisted Radiology and Surgery* (2024). DOI: 10.1007/s11548-024-03063-9.

1.3 STRUCTURE OF THIS WORK

We first provide background on conflict research and conflict data, the properties of such data, as well as ways to visualize them and to interact with the visualizations. We continue the background by discussing how the introduced approaches are used in different settings, ranging from exploration to explanation. Afterwards, we present the three core chapters of this work. They lead from EXPLORATION FOR EXPERTS over EXPLORATORY EXPLANATION FOR EXPERTS to EXPLANATION FOR BROAD AUDIENCES. Lastly, we conclude the work, discussing our results and listing potential future work.

2 BACKGROUND

In this chapter, we first provide general information about how the field of conflict research evolved, what kind of data sources it builds on, and how conflict data is of public interest outside the field of conflict research. Afterwards, we present the properties that spatio-temporal conflict data commonly have, before giving an overview of how they can be visualized and how to interact with the visualizations. Based on the most relevant related works, we then discuss how these techniques are used when dealing with conflict data in two different fields, visual exploration and visual explanation in the form of data-driven visual storytelling. Lastly, we explain how these two fields can be combined, before summarizing how the presented information in this chapter builds a foundation for our work.

The content of this section is a combination and extension of the background provided in three of our prior works [May+21; May+23b; May+24].

2.1 CONFLICT RESEARCH AND DATA SOURCES

The quantitative study of political conflict has a long tradition in the social sciences. In line with the seminal work of Richardson [Ric48], it had initially focused on the study of large-scale interstate wars. More recently, the focus has shifted to intrastate wars, including civil war and terrorism. The main catalyst for this shift have been a number of large-scale institutional data collection initiatives. They systematically collect information about conflict events on a disaggregate level, e.g., on the level of individual attacks, whereas before, data was collected rather on higher levels of aggregation [DGB14]. In the disaggregate data sets, conflict events are geocoded, time-stamped, and annotated with contextual variables describing, e.g., the actors involved and the event type. For instance, such an event could code that on January 1st, 1997, a violent attack was committed against civilians by an armed political group at latitude 0.1337 and longitude 29.289 [Ral+10]. The leading data sets in this area are the Armed Conflict Location and Event Data (ACLED) [Ral+10], the Uppsala Conflict Data Project – Georeferenced Event Dataset (GED) [SM13], the Global Terrorism Database (GTD) [Gtdb], and the Social Conflict Analysis Database (SCAD) [Sal+12].

These data have been used to analyze a wide range of policy-relevant topics, e.g., the motivations for individual attacks, deliberate targeting of civilians, or the relationship between inequality and violence [DGB14]. The focus primarily lies on the statistical analysis of conflict patterns and the derivation of causal relationships. Since the shift towards the more detailed collection of disaggregate data, data-based map visualizations have become more expressive and, therefore, more prominent in conflict research literature. However, such visualizations are mostly static, missing out on the benefits of interactivity.

In addition to data sets collected over time and curated for research, there is also interest in data about conflicts that are as up-to-date as possible. Such data is needed, e.g., to inform the broad public about how the events in conflict-torn regions unfold. It is usually embedded in news stories by newspapers such as *The New York Times* or *The Guardian*. Journalists gather this data from various sources. They include consulting firms [Roc], satellite imagery providers [Max; Cop], news agencies [Reu], analysis projects [Ctp], and organizations like the *NATO*, the *UN*, or *NASA*. The sources of information also include defense and health ministries of the countries involved in the conflict, statements from political officials, and even blogs of so-called "military bloggers" on Telegram or Facebook.

Many of the sources can have political or economic interests in the context of the conflicts they collect data on. This highlights the importance for journalists to verify the factfulness of the data and to make the acquisition process transparent for the readers of their stories. Furthermore, data can be inaccurate even without intentional manipulation by the source, e.g., if information like the number of casualties in a large-scale conflict has to be estimated. Conveying the uncertainty of such estimations also increases the transparency and the readers' trust [Sac+16].

2.2 COMMON PROPERTIES OF SPATIO-TEMPORAL CONFLICT DATA

In the context of spatio-temporal conflict data, the frame of reference is usually *geospatial*. However, for brevity, we only use the term *spatial* in this work.

When talking about conflict data, we need to distinguish between the actual *phenomenon* that took place, like an attack on a government building, and the *data element* recorded to represent the phenomenon, like a table entry specifying the location and time of the attack [Slo+23]. In the context of our work, the corresponding data typically has the following characteristics.

Adapting the terminology of Tominski and Schumann [TS20], the **spatial** extent of recorded phenomena can be

- a *point*, like the geo-location of individual attacks,
- a *line*, i.e., a sequence of points, like a road network, or
- an *area*, described by the outline of a region of interest, like seized territory.

The spatial information is usually provided as quantitative data. It is either described by the individual longitudinal and latitudinal coordinates of each data point, or as sequences of such coordinates to describe paths or outlines of geographical areas.

According to Aigner et al. [Aig+23], the **temporal** extent of recorded phenomena can be

- *point*-based, meaning that a phenomenon was recorded with a specific timestamp at which it took place, like a specific day, or
- an *interval*, meaning that the occurrence of a phenomenon stretched out over a period of time.

Point-based phenomena are usually individual *events*, whereas interval-based phenomena can be treated as *states*, i.e., phases of continuity. However, a state can also be represented by two "interval events," its beginning and its end [Aig+23], or by its beginning and its duration. The temporal information is often encoded on a day-based resolution [Don+19] but can also be encoded on a more precise basis like hours, minutes, and seconds [Sig]. Due to the cyclical nature of temporal units like days, weeks, months, and years, conflict data can be treated as *cyclic* [Aig+23], however, they are more commonly considered as *linear* [Aig+23]. Moreover, conflict data can have *multiple perspectives* [Aig+23], e.g., when combining recordings about the same phenomenon from different information sources, like different newspaper articles covering the same protest [Don+19].

In addition to the independent variables, space and time, dependent information can be encoded. It can be quantitative or qualitative. Quantitative information includes the number of casualties from an attack or the costs of a social aid project. Qualitative information includes the event type or the actors that were involved. Examples for event types are protests or violent attacks, like bombings, and examples for actors include civilian protesters or the military [Don+19].

Moreover, graph data can be associated with conflict information, like a hierarchical grouping of the actors: Protesters are a subcategory of nonviolent civilian groups, whereas the military is a subcategory of violent political groups [Don+19].

Taken together, conflict data is typically multidimensional, i.e., with multiple independent variables (space and time), and often also multivariate, i.e., with multiple dependent variables [TS20].

2.3 VISUALIZATION OF SPATIO-TEMPORAL CONFLICT DATA

In this section, we give an overview of how spatio-temporal conflict data can be visualized. As there exist various kinds of conflict data, a plethora of ways to visualize them would be possible in theory. To keep the section in a reasonable scope, we focus on the most relevant approaches. First, we present different ways to project Earth's 3D sphere to 2D. Then, we discuss how to visually combine spatial and non-spatial information. According to Tominski and Schumann, there are two main ways: *direct* and *indirect* visualization [TS20]. We introduce the two ways, adapting Tominski and

Schumann's structure, and we combine it with Mayr and Windhager's abstraction of approaches to integrate spatial and temporal dimensions [MW18].

2.3.1 *Map Projections and Cartographic Layers*

Map displays are suitable to visualize spatio-temporal conflict data as they allow the preservation of the spatial structure [AA06]. In the real world, conflict phenomena take place in 3D space, when considering the elevation of geographic locations in addition to their longitude and latitude. Such information can be relevant to better understand the dynamics of certain conflicts, as geographical features such as the course of a mountain range or a river can be crucial for the borders between neighboring territories [Mar16]. Most commonly, this space is not visualized directly in 3D but projected into 2D. In that case, the elevation can be encoded otherwise, e.g., by mapping it to the color or texture of the terrain, or following other cartographic guidelines [Mac04]. Cartography also deals with the loss of information that the projection of Earth's 3D sphere into 2D inevitably introduces. Different projection methods were proposed, focusing on preserving different geographic aspects as well as possible.

A tradeoff has to be made between the truthful representation of the sizes of areas, their shapes, angles, directions, and distances between points. For instance, the Mercator projection preserves angles and shapes while the Albers projection distorts shapes to preserve area, see Figure 1. There are also projections that try to compromise between preserving different aspects, like area and angle in case of the Equal Earth projection [BSJ19]. More advanced projections like polyhedral projections, which slice up the globe and project it onto multiple connected faces of a polyhedron [Pę17], are rather uncommon in conflict research.

In the projected maps, additional cartographic features can be displayed aside from the elevation. Information like city names, road networks, or bodies of water can be relevant for the dynamics of conflicts. For each type of information, a layer can be generated and superimposed on the same map view. Ideally, users are provided with the option to toggle the visibility of individual layers to display only the information of interest.

2.3.2 *Direct Visualization*

In direct visualization, the non-spatial information is embedded directly in the projected geographic space. In this embedded form, the temporal information can be included in the following ways [MW18]: via *animations or slideshows*, *juxtaposition*, *superimposition*, or *stacking*. We first explain the idea behind each of the four approaches and shortly discuss their advantages and disadvantages according to Mayr and Windhager [MW18]. Afterwards, we give a brief overview of which types of visualizations can be used to display data with point-, line-, or area-based extent. In this context, we

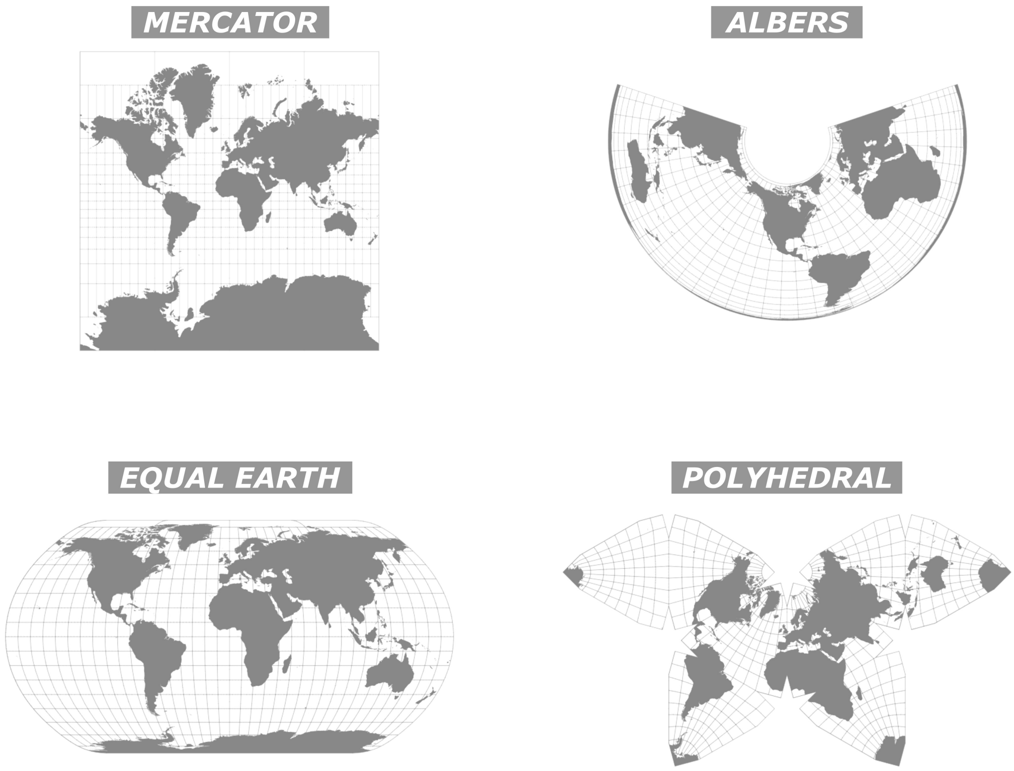


Figure 1: Four different types of map projections.

also mention which of the four kinds of approaches for encoding time is suitable for the different visualizations.

In each of the approaches, *animations or slideshows*, *juxtaposition*, *superimposition*, and *stacking*, a layer is generated for each time step in the data, but the arrangement of the layers differs between the approaches, see Figure 2.

- For *animations or slideshows*, the different layers of temporal information are displayed on the map one after the other. They can be displayed separately, with breaks in between, or in quick succession, to create an impression of fluid transitions between the steps. Animations and slideshows facilitate an effective perception of changes, however, comparisons between time steps become more difficult the more other time steps separate them from each other. Animations and slideshows can also be cognitively demanding to interpret if too many changes happen at once. In literature, the approach to animate temporal changes is also referred to as *dynamic*, in contrast to the other following approaches which are *static* [TS20].

- For *juxtaposition*, the layers are placed side-by-side, like in *small multiples*. This way, all layers are visible at the same time, facilitating an easier comparison between time steps that are further apart. In addition, the information from the layers does not have to be kept in short term memory, as it is the case over the course of an animation. However, with juxtapositions, additional effort is needed to precisely compare the information from different steps, as the information does not change in place, but across multiple separated map views.

- For *superimposition*, the different layers are merged onto a single map, while providing means for the viewers to distinguish to which time step each of the different pieces of displayed information belongs. For instance, this can be achieved by coding time via color. This approach allows the user to naturally integrate the spatial and the temporal information. A downside is that visual clutter and occlusion can occur more easily, as all information is displayed at once. Moreover, it might be difficult to encode the temporal information in a way that the different time steps are still reasonably distinguishable.
- For the last approach, which we refer to as *stacking*, a third display dimension is created on top of the map, orthogonal to the two spatial dimensions. This third dimension is used to display the different layers of temporal information. The approach is also referred to as *space-time cube* [Kra03]. Stacking, as compared to superimposition, has the advantage that both the spatial and the temporal dimensions can be distinguished equally well. However, visual clutter and occlusion can also occur in this view. Moreover, as a 3D cube is displayed on a 2D computer screen, misinterpretations about either the spatial or the temporal positions are possible due to perspective misconceptions. To avoid this, a well-chosen orientation of the cube towards the user is required and, ideally, means to rotate the view.

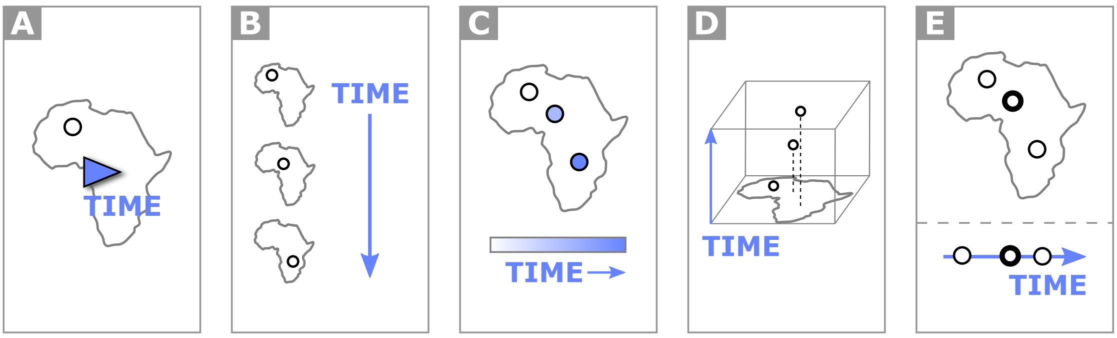


Figure 2: Mayr and Windhager's five approaches to encode space-time consist of (A) *animations or slideshows*, (B) *juxtaposition*, (C) *superimposition*, (D) *stacking*, and (E) *multiple coordinate views* [MW18]. Approaches (A)-(D) are introduced in Section 2.3.2, and approach (E) is introduced in Section 2.3.3. *(Figure adapted from Mayer et al. [May+23b], licensed under CC BY 4.0.)*

When it comes to the visualization of conflict data, some fundamental *visual variables* are used primarily to control the visual appearance of the *marks* used to display the individual elements [Mac04]. The basic marks to build visualizations are 0D points, 1D lines, 2D areas, and 3D bodies, while 3D bodies are rather uncommon in visualizations of conflict data. Moreover, the dimensionality of the point and line marks is only theoretical, as, in practice, they also require a two-dimensional extent on the screen to be visible to the viewer. The marks' visual variables, which can be used to encode values of the underlying data points, include the position, size/area, color, and shape of the mark, while color can be further separated into hue, lightness, and saturation. To visualize conflict data, the position variable is usually reserved to encode the geographic location (and the time in case of layer *stacking*). Size is often used to encode additional quantitative information, like the number of casualties from an attack. Color, particularly the hue, and shape are typically used to represent categorical information, like different event types or different actors in a conflict, whereas color can also be used to encode the temporal information. In the latter case, typically the saturation and the lightness are adjusted rather than the hue.

Depending on the spatial and temporal extent of the phenomena introduced in Section 2.2, different encodings and chart types are applicable. Note that, only because a phenomenon has a certain spatial or temporal extent, the corresponding mark does not automatically have to have the same extent. For examples of the mentioned visualization techniques, please refer to Figure 3.

Point (spatial). If the spatial extent of the depicted data elements is point-based, dot maps, glyph maps, and space-time cubes can be used.

Dot maps are usually only applicable if the temporal extent of the data elements is also point-based. To encode time in dot maps, *animations or slideshows*, *juxtaposition*, and *superimposition* can be used. The individual dot marks can depict additional information of the data elements by adjusting their size, color, or shape. If multiple variables are encoded in self-contained graphical objects, they are considered as *glyphs* [TS20].

Glyph maps can be used to depict both temporally point- and interval-based data. As glyph maps are usually more complex to read and interpret than dot maps, encoding time via *animations or slideshows* and *juxtaposition* should be treated with caution to limit the cognitive effort required for performing comparison tasks. When encoding time in the glyphs directly, they can even turn into small charts on their own, considered as *embedded charts* [AA06]. However, the larger the glyphs are, the more difficult it is to avoid their overlap without shifting their position, which would skew their geographic integrity. Glyph maps can also be used to display data that was recorded for entire areas, but since the glyphs are still placed at specific points on the map, e.g., the centroid of the area, we consider them as point-based marks.

The line between glyph maps and space-time cubes gets blurred if the glyphs are displayed as 3D objects, *stacking* the temporal information in the third dimension [TSWS05; TS12]. Aside from 3D glyphs, space-time cubes can also be used to

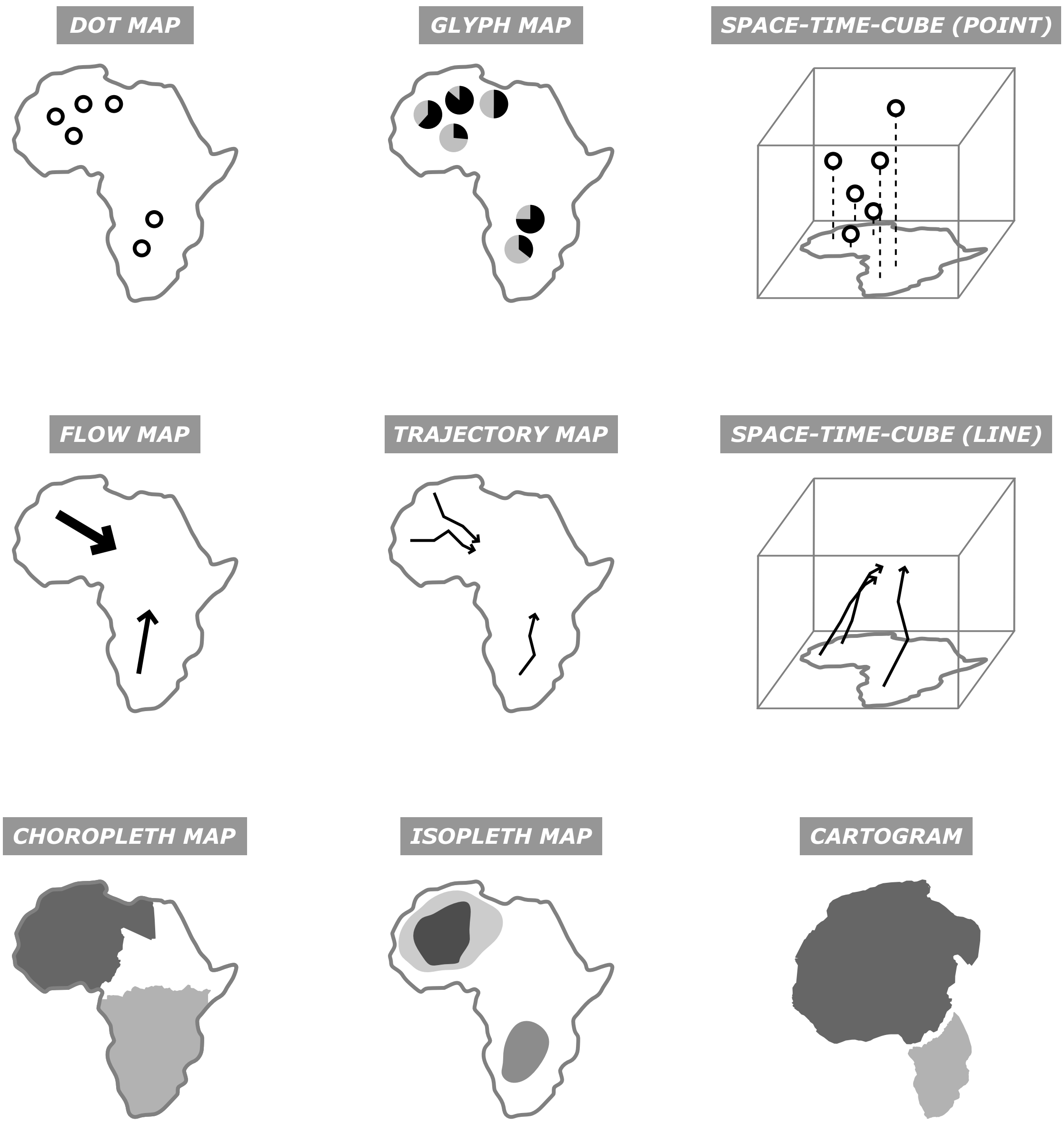


Figure 3: Sketches for nine visualization techniques used for displaying data on a map. The first row contains techniques suitable for point marks, the second row contains techniques for line marks, and the third row contains techniques for area marks.

display dot marks. In this case, however, appropriate visual guides like depth cues might be necessary to facilitate the correct interpretation. Depending on the type of mark, point-based or interval-based temporal information can be displayed in a space-time cube. Overall, however, using space-time cubes to visualize conflict data is rather uncommon. This is also the case if they are used to display line marks, which we discuss next.

Line (spatial). To visualize line-based spatio-temporal data, flow maps, trajectory maps, and space-time cubes can be used.

Flow maps show the direction of recorded movement between origin and destination, irrespective of the underlying path geometry [Jen+18]. An example would be a simplified visualization of streams of refugees from a conflict zone to other regions. As marks, typically arrows or lines are used. The movement records can be associated with respective volumes, like the number of refugees. Due to their nature, flow maps encode temporal information via layer *superimposition*, as flow marks connect destinations and origins, i.e., information from different points in time, in the same map view. In flow maps, it is advisable to use curved instead of straight flow marks to reduce clutter, and to use arrow heads to indicate direction [Jen+18]. If a large number of flows need to be displayed, aggregating them into bigger flow marks or into additional glyphs might be necessary [AA13; And+17].

If the movement paths of individual objects should be represented accurately, trajectory maps are commonly used [DN21]. In contrast to flow maps, in addition to the origin and the destination of a movement, the geometry of the remaining movement path is also relevant in this case. Similar to flow maps, when displaying large numbers of trajectories, measures need to be taken to avoid visual clutter, requiring approaches for aggregation, e.g., via clustering and edge bundling [And+17]. The *superimposition* of temporal layers in flow and trajectory maps can also be combined with other approaches, like *animations or slideshows* and *juxtaposition*, to distinguish between flows from different, larger periods of time.

In addition, *stacking* temporal layers in the third dimension is possible, again, marking the transition to space-time cubes. In fact, the original purpose of space-time cubes was to depict movement paths [Kra03]. To do so, the temporal components of the paths are commonly mapped to the third dimension, producing 3D paths. However, each path can also be flattened into 2D before stacking them on top of each other, thereby combining layer *superimposition* and *stacking* [Tom+12].

Area (spatial). Information displayed via an area mark can either be recorded for the entire area directly, or it can be the result of an aggregation of individual point-based recordings from within that area. To display such information, either choropleth maps, isometric/isopleth maps, or cartograms can be used, while the latter are rather uncommon in the context of conflict data.

In choropleth maps, data values are mapped to the color of the depicted areas [Tob73]. The areas typically describe administrative boundaries such as cities or counties, or natural geographical regions. By mapping a single color to an entire area, a uniform distribution of the corresponding data value across the entire area is (sometimes falsely) conveyed. Both temporally point- and interval-based data can be displayed on choropleth maps, while it is difficult to depict precise interval durations without additional glyphs or separate linked views. As only a limited number of attributes can be displayed in a single static choropleth map, and the area marks use up the entire map area, layer *superimposition* is not ideal to encode time-dependent information. Instead, *animations or slideshows* and layer *juxtaposition* are more suitable. Extending the area marks into 3D via *stacking* is not feasible, as it would produce a lot of occlusion.

In contrast to choropleth maps, isometric and isopleth maps encode data in area marks that do not follow regional boundaries [WGK10]. Instead, the outlines of the areas and the depicted values are determined by interpolating between values measured at different positions.

Differing from the previous approaches, cartograms do not use color as the primary visual variable to encode data. In cartograms, data values are primarily mapped to the size and shape of the geographic areas directly, distorting their shape [WGK10]. This is detrimental to their geographic integrity, which is why they are less applicable to display conflict data.

2.3.3 *Indirect Visualization*

In indirect visualization, the non-spatial information is displayed in separate views with an abstract frame of reference instead of a geographic one. These separate views are often combined with a map display on the side, but do not have to be. For instance, if the location information of the data elements corresponds to well-known geographic regions like cities, it can suffice to use the city names as labels for the corresponding visual marks in an otherwise abstract chart. Such a view might not require an additional map display. If both a map display and one or multiple abstract views are used, visual cues like color help to establish the connection between the non-spatial information and the corresponding geographic position.

If time is encoded in a separate view instead of directly on a map, Mayr and Windhager refer to it as a fifth approach next to *animations or slideshows*, *juxtaposition*, *superimposition*, and *stacking* [MW18]:

- In *multiple coordinate views*, the temporal information is separated from the map display in an additional, abstract view, see Figure 2 (E). To establish the connection between the information from the different views, visual cues like color can be used, or interaction like linked brushing [WBWK00]. In multiple coordinate views, the temporal information can be displayed more clearly and in more detail than in the other four direct visualization approaches, where it needs to be integrated into a map display. However, due to their separation, the temporal and spatial information are more difficult to integrate mentally.

To discuss in more detail how the temporal information can be encoded in the abstract visualizations of multiple coordinate views, theoretically, we could re-apply the classification by Mayr and Windhager on the abstract visualizations. For instance, consider a situation in which multiple scalar variables were measured for the same data point over time. In this case, the values from a single time step could be visualized as a bar chart. To display the values from all time steps, the different "temporal layers" of the bar chart could be *animated*, *juxtaposed* as small multiples, *superimposed* as grouped bar charts, or turned into 3D bar charts and *stacked* in a third, orthogonal dimension to reflect the changes over time. This means that the categories *animations or slideshows*, *juxtaposition*, *superimposition*, and *stacking* could, again, be applied in the category of *multiple coordinate views*.

However, we will not go into such detail in the following. Rather, we summarize the most common approaches to encode time in the abstract views of multiple coordinate views. Most commonly, time is statically displayed as the x-axis of a line chart. Aside from line charts, area charts or stream graphs can be used to visualize temporal information.

Other more basic abstract visualizations include bar charts, histograms, and scatterplots, which are a popular choice due to their simplicity. To display more complex, multivariate data, the following types of visualizations, adapted from Tominski and Schumann, can be used [TS20].

- As conflict data are often stored as tables, *table-based visualizations* are a reasonable choice to depict the non-spatial information. Instead of displaying just the textual information, the table cells can be colored based on the underlying information, or bars of different lengths can be depicted in the cells to encode quantitative information. This makes trends and relations between the data elements easier to identify, particularly when allowing users to rearrange the table rows.

- *Combined bivariate visualizations* can be used to depict the pairwise relations in multivariate data sets. For instance, in scatterplot matrices, one scatterplot is created for every pair of variables, and they are arranged in a grid. They are suitable for displaying pairwise correlations and outliers.

- In *polyline-based visualizations*, each data element is represented by multiple connected straight line segments. For each of the non-spatial variables, an axis is created on which the value of the element is encoded via position. Parallel coordinates and radar charts are common examples. In these views, relationships across multiple dimensions can be identified more easily than in, e.g., scatterplot matrices. However, the order of the axes is important as the relationship between adjacent axes can be interpreted a lot better than between distant axes.

- *Glyph-based visualizations* encode the values from multiple variables in a single glyph. They are well-suited for providing an overview, but as the individual glyphs typically do not take up a lot of display space, it can be hard to convey detailed information. While this technique can be used for indirect visualization, i.e., embedded in an abstract frame of reference, it is more commonly used as direct visualization technique, embedded directly on a map.

- *Graph visualizations* can have fundamentally different layouts. To visualize the connectivity of the nodes from graph data, *node-link representations* are a common choice. In them, the nodes are depicted as dots and their connectivity as explicit line edges. Additional variables can be encoded via the color, size, or shape of the marks representing the nodes and edges. To reduce the risk of visual clutter, the layout of a node-link diagram needs to be defined thoughtfully. *Matrix representations* encode the connectivity not as lines but as cells of a matrix. The rows and columns of the matrix correspond to the nodes, such that each cell can represent the connectivity of one pair of nodes, typically via its fill color. Depending on the arrangement of the rows and columns, interesting patterns can be visualized quite clearly, but usually not all at the same time. For *implicit representations*, the connectivity of the nodes is not conveyed explicitly. Instead, the connectivity of the nodes is conveyed through their relative position, e.g., via adjacency as in icicle plots, or via nesting as in treemaps.

The information in the separate, non-spatial views can refer to just a single data element, i.e., just a single spatial position, or to multiple data elements and positions. If only the data of a single element is displayed, the user should have the option to select the element. If the data of multiple elements is displayed, it can be aggregated or displayed individually for each element. There also exists the approach of *probing*, where a separate view is opened for each location the user selects on a map. Such additional views can also be displayed immediately upon hovering over a map, with the separate view being displayed directly at the mouse pointer position. In this case, the line between indirect visualizations and embedded charts, a direct visualization technique, becomes blurred [AA06].

To conclude this section, direct visualization should be used if more focus is put on the geospatial relations of the data, whereas indirect visualization is better suited if larger amounts of additional and potentially complex information should be displayed. The two approaches can also be combined, only showing selected additional information on the map and the remaining information in additional, abstract views.

Also note that, while we have presented various visualization techniques in discrete categories, it has become clear at multiple points that the lines between these categories are not always that clear. As visualization designers combine existing concepts in new and innovative ways, the lines between the concepts also get blurred.

2.4 INTERACTION

In visualization literature, there exist various definitions of what interaction in the context of visualization comprises. From them, Dimara and Perin have distilled a definition: "Interaction for visualization is the interplay between a person and a data interface involving a data-related intent, at least one action from the person and an interface reaction that is perceived as such." [DP20]

Providing means to interact with visualizations can amplify the user's cognition, help them to construct and refine knowledge, and make visualizations overall more accessible [DP20]. Focusing on 2D visualization, and cartography in particular, Roth identified four dimensions of interaction primitives [Rot13]. The dimensions are in line with the "stages of action" from Norman's interaction model [Nor13], and they cover system-, task-, as well as human-centric aspects [DP20].

- The first dimension refers to the **broader interaction goals**. It includes the goal to *procure*, i.e., to retrieve information, to *predict*, i.e., to forecast what will happen in the future, and to *prescribe*, i.e., to decide which actions should be taken based on the current situation and the predictions. In conflict research, all of these goals also exist independently of whether interactive visualization is applied or not, for instance, in the goal to reduce the risk of violent conflicts or to resolve ongoing conflicts.

- The second dimension proposed by Roth consists of the **objective primitives**, which refer to the tasks the users try to solve. Accordingly, users may want to *identify* certain elements, *compare* them, *rank* them, *associate* them, or *delineate* them, i.e., categorize or group them.

- To solve these tasks, users rely on the third dimension, the **operator primitives**, i.e., the "interaction techniques" [Kei02]. On the one hand, they consist of the options to *import*, *export*, *save*, *edit*, and *annotate* data. These operators are referred to as *enabling* operators. On the other hand, the *work* operators cover methods to *reexpress*, i.e., change the representation type, e.g., by switching between different ways to encode temporal information on a map [MW18]. They also include the option to *arrange* the views in a layout, to *sequence* views, to *resymbolize* them, e.g., by changing the color scheme, or to *overlay* additional information layers over a visualization, which is particularly common for map displays. With *reproject*, a different projection method can be selected for map representations, whereas the operators *pan* and *zoom* can be used to change a map's viewpoint. Roth's *filter* operator includes multiple techniques that are sometimes further distinguished in other works. For instance, in the terminology of Tominski and Schumann [TS20], it includes the techniques to "highlight" or to "dim" a selected group of elements in comparison to the unselected elements, or to remove them entirely from the view. In addition, the *search* operator allows finding specific elements or locations, while the *retrieve* operator yields additional information about selected elements or regions on a map. Lastly, *calculate* includes techniques to derive new information from existing elements, like measuring the distance between two points on a map.

- The fourth dimension, **interaction operands**, has an explicitly geospatial focus. It covers on which target an interaction technique is applied, or in the terminology of Andrienko, on which "search target" [AAG03]. The three options are *space-alone*, where the user interacts only with the geographic component of a visualization, *attributes-in-space*, where they interact to determine how additional variables vary across the geographic space, and *space-in-time* to determine how spatial patterns change over time. In addition, this dimension covers on which level (or "search level" [AAG03]) the interaction is applied: It can either be *elementary*, referring to a single map feature, or *general*, referring to all map features.

Tominski and Schumann provide advice regarding the implementation of interaction [TS20]. Accordingly, interaction should be conflict-free, meaning that each action, like a click or a drag, should be associated with a single task. Moreover, the physical and mental costs of interaction should be limited, e.g., by supporting a comparison task by automatically collecting elements that are most (dis-)similar to a selected element. In addition, the more direct an interaction is, i.e., the more closely it takes place to its target from a conceptual and spatial point of view, the easier it is for the user to build an appropriate *mental model* [BB13] of the visualization.

Building the mental model is further facilitated by ensuring the consistency and closure of the interactions and by offering informative feedback. This feedback should be as immediate as possible. If the resulting visual changes are larger, it can also help to animate them to allow the user to track the changes more easily, instead of having discrete jumps between the views. In addition, the interactions should be tailored to the target audience's experience with interactive visualization and provide the users with options to reverse their actions.

A lot of research regarding interaction revolves around mouse and keyboard as input devices [DP20]. This is also the case for conflict data. However, as there is public interest in such data, they are also often used as the basis of newspaper articles. As touch-based mobile devices nowadays play a central role in our everyday lives, such articles are also published for these devices. Accordingly, more focus is put on how touch gestures can be used to facilitate interaction. For instance, horizontal and vertical swiping and scrolling have become central input gestures for navigation [Rot21].

2.5 EXPLORATION

In the subsequent sections of this chapter, we use the foundations we have gathered so far and put them into the context of previous research.

In this section, specifically, we first discuss exploration approaches in general, including how users can be supported when using complex visualization systems. Afterwards, we narrow down on corresponding approaches working on conflict data, before summarizing how the state of the related research motivates our studies regarding visual exploration in this work.

Tasks, tools, and principles. In the following two paragraphs, we summarize the general approach of exploratory visual analysis of spatio-temporal data based on Andrienko and Andrienko's book [AA06]. There is a strong overlap between the principles described here and the more general field of *visual analytics*, in which automated analysis plays a more central role [Cui19].

Exploratory approaches are commonly used if data is of greater size and higher complexity. These approaches can be used to improve the understanding of a data set, to generate or validate hypotheses, and to construct new knowledge overall. The tasks performed in such settings can be grouped into *elementary* tasks and *synoptic* tasks. *Elementary* tasks work on individual data elements, looking up information, comparing it, or seeking for relations. In contrast, *synoptic* tasks work on groups of elements to understand *behaviors*, i.e., particular configurations of attribute values like distributions or trends. At that, *descriptive* synoptic tasks are more straightforward, including the tasks to identify, describe, or compare behaviors in the data set. In contrast, *exploratory/connectional* synoptic tasks are more open, extracting more general structural patterns and interrelations between behaviors.

To explore multifaceted and complex data, usually, multiple approaches (or *tools* [AA06]) are necessary. They range from manipulations of the underlying data to make them more workable, over the (semi-)automatic extraction of information via machine learning (like in visual analytics), to display manipulations for determining subsets of interest, like filtering and zooming. Usually, multiple approaches have to be combined, either sequentially, using the output of one approach as the input for another, or concurrently, running multiple approaches in parallel, independent of each other. Core principles include to provide both overview and detailed analysis, to provide means for identifying groups of interest and establishing linkages and structures between them, and to involve the user's domain knowledge.

Particularly the principle that the user's domain knowledge should be involved indicates that the target audience for such applications typically consists of users with a considerable amount of knowledge in the target domain that they can contribute to the analytical process. Also, a certain level of expertise is required by the users to familiarize themselves with such analytical systems, since these systems commonly have intricate functionalities and workflows. Accordingly, the audience of exploratory analysis systems is often referred to as *expert* audience, or *domain experts*. Especially the second term highlights that the users are not considered “general experts,” but experts in a certain field.

Guidance. The various ways in which data can be explored cannot be realized with only a fixed set of static visualizations. Instead, interaction plays a key role in exploratory analysis. Accordingly, corresponding systems usually contain multiple of the interaction primitives introduced in the previous section. However, the more interaction options are provided in a system, the more complex it also is for new users to familiarize themselves with the various ways in which the options can be combined to derive insights.

For this reason, Ceneda et al. examined the concept of *guidance* in the context of visual analytics [Cen+17]. They summarized ways to overcome the user’s knowledge gap by helping them to identify, judge, or execute actions. The authors distinguish three types of support with increasing degrees of guidance. First, *orienting* guides the user via visual cues, e.g., by automatically generating a map of patterns found during an analysis session. Second, *directing* suggests options in the analysis process, like a ranked list of alternative views. Third, in the most system-controlled approach called *prescribing*, the system leads the user through automatically generated visual results.

Visualization literacy. Irrespective of whether a visualization system provides guidance or not, a fundamental requirement for users to understand the system’s outputs is to have sufficient *visual literacy* [Fel08], or more specifically, *visualization literacy* [FJL22]. This is particularly important if more complex visualizations are used, which is especially the case in exploratory settings, where the analyzed data is also more complex [Cen+17]. Visualization literacy refers to “the ability and skill to read and interpret visually represented data in and to extract information from data visualizations” [LKK17]. Lee et al. showed that users’ visualization literacy correlates to their *numeracy*, i.e., their “ability to understand and process numerical information,” [Lee+19] and to their *need for cognition*, i.e., their “tendency to engage in effortful cognitive activities” [Lee+19]. Based on their findings, the authors state that improving numeracy education also benefits visualization literacy. To increase the users’ motivation for cognitively engaging with visualizations, they advocate designing the visualizations to be aesthetically pleasing, to consider the users’ personal experience, and to spark curiosity, emotion, and creativity. In addition, the authors state that, ideally, visualization systems should adapt to the individual differences between the users, automatically selecting the types of visualizations and guidance techniques accordingly.

Previous applications. We now focus on previous applications of visual exploration applied on spatio-temporal conflict data and show that, while only a limited number of such works exist, there is a large variety of data sources analyzed and visualization techniques used in them. After the original introduction of the data set GTD in 2007 [LD07], several approaches were proposed to visually analyze the patterns in it. Techniques like color strips [God+08], glyph maps [Jon+08], and dot maps in combination with parallel sets [Wan+08] were proposed. Gorecki et al. expanded on these approaches by including parallel sets and node-link visualizations of network data [GSS11].

While the previous works focused on a single data set, Weidmann and Kuse proposed a dashboard solution to overlay information from multiple data sets, including ACLED, on the same animated map [WK09]. However, as the authors strove to keep the interface as simple as possible, only limited exploration is possible. The same holds for web-based visualization dashboards developed directly for individual data sets, like for ACLED [Acl] and GTD [Gtda]. While these solutions are designed to be accessible, a certain level of expertise regarding geopolitical conflicts and visual literacy is still required to draw meaningful conclusions from them [Lee08].

More recently, Saldarriaga et al. analyzed interactive satellite imagery maps, flow maps depicting the displacement of people, and a dot map in connection with spatially referenced YouTube videos [SKB17]. From these diverse data sources, they derived insights about how conflicts shape the structure of urban regions. Moreover, Silva et al. developed a visual analytics system to explore spatio-temporal events, testing it, among others, on a data set of violent attacks against civilians in Africa [Sil+19]. Their main focus lied on exploring patterns on different levels of spatial and temporal detail, visualizing them using parallel coordinates, matrix representations, scatterplots, and timelines. Working on spatio-temporal event data streams, Robinson et al. computationally extracted patterns from RSS news feeds, focusing on reports about political or military events in Syria [Rob+17]. They explored the patterns using area charts, stacked bar charts, glyph charts, and tabular displays.

A field related to conflict data and conflict research is that of *crime mapping* [San22]. In this context, Brunsdon et al. compared multiple visualizations, like animated and juxtaposed isopleth maps, to find that there was no single best visualization to support crime analysis, but that all had their advantages and limitations [BCH07]. When developing a visual analytics system to analyze crime data, Roth et al. tried to find a trade-off between the utility and the usability of the system [RRM15]. They report on the benefits that user-centric prototyping with iterative feedback loops had on their development process.

Furthermore, data about natural disasters, like earthquakes, have similarities to conflict data, as corresponding events tend to occur situationally and in clusters. Accordingly, relationships like the dynamics between seismic activities are analyzed using statistical methods and visualizations like scatterplots and dot maps [Dzw+05].

Summary and contribution. Taken together, in this section, we showed how diverse the techniques are that can be applied when visually exploring spatio-temporal conflict data. However, it also became clear that this potential is not yet used to its fullest, as the number of corresponding applications is still limited. Moreover, such applications often only focus on one specific data set, while a few examples showed that it can be beneficial to combine complementing sources of information. However, this was done quite rarely. One reason for this can be the complexity that is introduced during the development process when combining multiple data sets in a single application. Another issue is that it is not straightforward to determine whether all elements contained across different data sets actually represent different, unique phenomena, or whether some of them represent the same original phenomenon, thereby introducing a misleading bias when they are treated as unique elements.

To this end, we contribute to the exploration of conflict data in the following way. We introduce a system that combines various interactive visualizations to explore the results and underlying workings of an existing statistical method developed for integrating different conflict data sets.

2.6 EXPLANATION

When performing visual explanation, commonly, insights have been extracted from data at a prior step and should now be communicated to an audience. The audience in these scenarios is usually rather a lay audience, but it can also consist of stakeholders more closely related to the domain where the data originated from. In both cases, the audience members should not be expected to have extensive background knowledge about the domain. Accordingly, the information needs to be broken down, focusing on the key messages. The goals when communicating data insights can be to support a decision process for the audience members, to convince them, to inform or educate them, to entertain them, or any combination of the above. To reach these goals, it is beneficial to incorporate concepts from storytelling, conveying the insights in the form of a *visual data story* [GP01]. This can increase the audience engagement, comprehension and memorization of the communicated content [Mer21; NL10]. The approach is also referred to as *data-driven visual storytelling* [Ric+18]. For brevity, we refer to it just as *visual storytelling*, even though the stories in our work are also driven by data. This abbreviation was also used in other works such as Tong et al.'s 2018 survey of the field [Ton+18].

In literature, a distinction is made between the concept of a *story* and a *narrative*. However, the corresponding definitions are not consistent. We follow Roth's definitions, who uses the term *story* as "an account of specific events, places, and people" and *narrative* as "the structure and presentation of this content that shapes the meaning of the story" [Rot21].

In this section, we provide a background on visual storytelling. At that, it should be noted that this field is younger than exploratory visualization. Accordingly, there exists a substantial number of more general works, but only few with a focus on spatio-temporal data. In the even more specific context of conflict data, no dedicated works were published so far. Therefore, the insights we present here often come from more general works, which, however, does not deter from their relevance for our work. We first present narrative *genres* and *tropes* before introducing categories of *techniques* for visual storytelling. Then, we summarize selected guidelines for writing visual data stories, before shortly discussing the concept of story authoring tools.

Please note that, in the context of visual storytelling, we use two terms when referring to the person consuming a story, based on the situation. If the person's role is more active, we refer to them as a *user*, and if their role is more passive, we refer to them as a *reader*. We also want to note that, when it comes to sensitive topics like violent conflicts, approaches from storytelling are not integrated to make the consumption of the content more "entertaining." Rather, the goal is to increase the immersion and relatability for an audience for whom, hopefully, such experiences are more distant and, therefore, more difficult to grasp.

Visual storytelling has been receiving increased attention since 2010, when Segel and Heer presented a design space and narrative structures for visual data stories under the term *narrative visualization* [SH10]. Their design space includes seven *genres* of visual data stories and a set of storytelling techniques. Over time, adaptions of the genres as well as the storytelling techniques were proposed, which we present in the following.

Genres and tropes. Roth adapted Segel and Heer's genres to map-based storytelling [Rot21] and proposed the following genres, see also Figure 4. They are all applicable in the context of conflict data.

- In a *static visual story*, the layout of the view is partitioned into different frames and the reading direction is conveyed via annotations.
- *Longform infographics* are commonly known from online journalism, where a story is usually navigated via vertical scrolling.
- *Dynamic slideshows* consist of a discrete set of slides, often navigated in a horizontal direction.
- In *narrated animations*, the story unfolds as a video or as an animated map playing over time.
- *Multimedia visual experiences* combine weblinks to multiple "chapters" of a story distributed across separate webpages. Moreover, they can use various media, like images, maps, videos, and sounds to deepen the immersion for the reader.
- In *personalized story maps*, the readers can add their own experiences to an interactive map, actively contributing to the storytelling process.

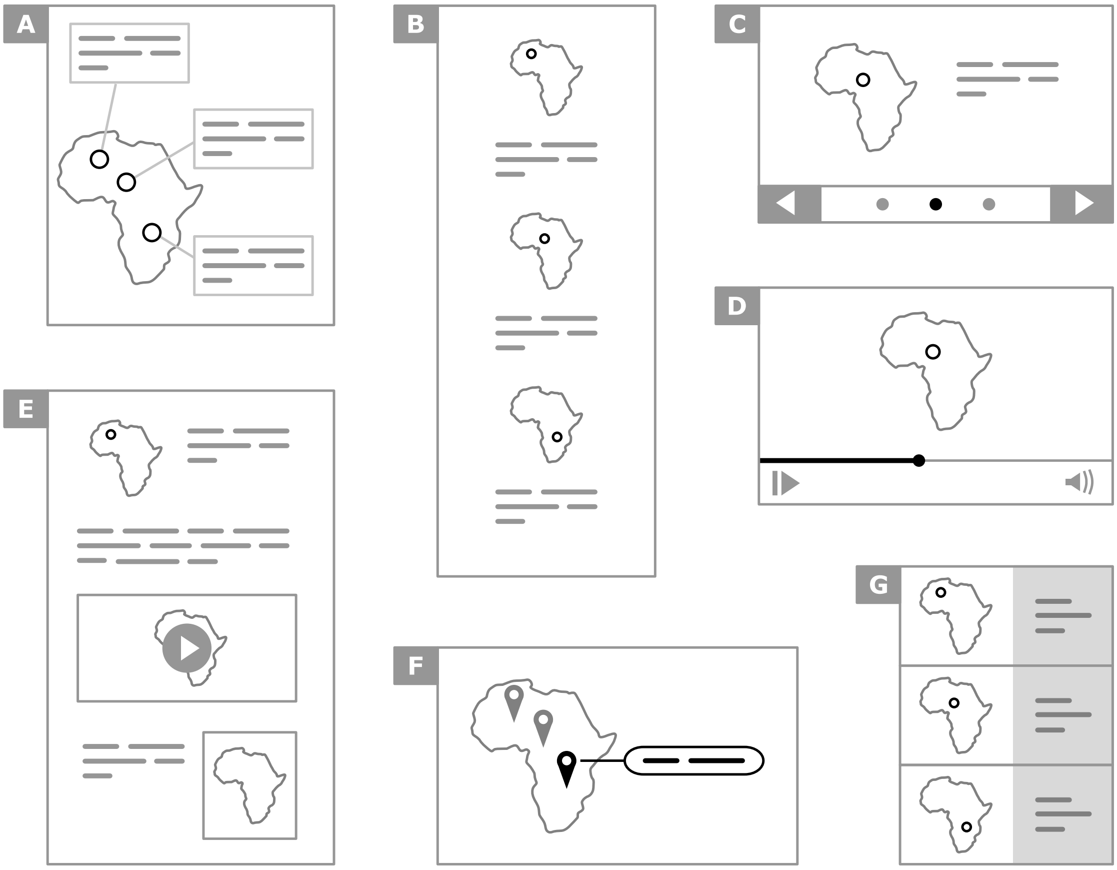


Figure 4: Roth's genres for map-based visual storytelling consist of (A) *static visual stories*, (B) *longform infographics*, (C) *dynamic slideshows*, (D) *narrated animations*, (E) *multimedia visual experiences*, (F) *personalized story maps*, and (G) *compilations* [Rot21].

- Lastly, in *compilations*, multiple reports corresponding to the same overarching event are collected on a single page in the form of visual abstracts. For instance, if a conflict lasts for a longer period of time, new ''sub-stories'' can be created regularly to keep readers updated about the latest developments. In a compilation, these sub-stories are combined into a visual overview of the entire conflict, linking to the individual sub-stories.

Under the term *narrative tropes*, Roth also summarizes techniques used to advance the narrative of a story [Rot21]. He explains that the mood of a story can be set by following a consistent design, including the color scheme, form styling like line weights, or also the font types. He also mentions how dosing the story content appropriately can help to reduce the complexity of the story, and how repeating important information can help to develop a theme for the story. Moreover, he discusses how metaphors can be used to facilitate understanding, and how embedding the voice of the author or the story characters can clarify and frame the meaning of the story.

Storytelling techniques. Expanding on Segel and Heer's work [SH10], Stolper et al. identified a set of *storytelling techniques* [Sto+16]. The authors collected four overall categories of techniques.

- The first category covers techniques that help at *communicating the narrative and explaining the data*. This includes using basic text blocks to narrate the story, or guiding the reader across a map using flowchart arrows.
- Techniques for *linking separate story elements* help to make connections, e.g., between text referring to certain data points and visual representations on a map referring to the same data points. For instance, the same color can be used for both the text and the representations on the map, or the corresponding elements can be animated simultaneously.
- The techniques in the category *enhancing structure and navigation* describe how the author conveys the order of the different parts of their story and how readers can navigate through it. A straightforward example to do so is via "next" and "previous" buttons or scrolling, while also more direct approaches can be used, like clicking on a region in a map to jump to the part of the story which plays in that region.
- Lastly, *providing controlled exploration* covers techniques for allowing readers to interact with the story content to explore further information, like filtering the elements displayed on a map and zooming into certain regions. These storytelling techniques can also be combined, e.g., by linking separate story elements both via color and simultaneous animation.

This set of techniques, which was derived from various visual data stories published online, reflects that, in visual storytelling, only a small subset of Roth's less complex interaction primitives is used [Rot13]. As depicted, interaction is primarily used for navigating through the story and for controlled exploration. This aligns with the overall approach of visual storytelling, where the story and the narrative are commonly rather *author-driven*, as compared to the more *user-driven* analysis process in exploratory visualization [SH10]. This simplicity of interaction reduces the cognitive load for the users, who might, on average, be less familiar with advanced interaction techniques than domain experts. Moreover, simplifying interaction reduces the risk of producing unwanted states in the visualization that are difficult to interpret for the users.

The same preference of simplicity holds for the types of visualizations used. Primarily, data stories use visualizations that are familiar and not too complex to understand, therefore, allowing as many readers as possible to understand the story. Examples include bar charts, line charts, and area charts, as well as map displays like dot maps, flow maps, and choropleth maps. In contrast to exploratory analysis approaches, more focus tends to be put on the aesthetics of the visualizations and texts to increase user engagement [HP17]. Through this customization, visualizations that are rather simple and "forgettable" in their base form also become more memorable [Kos16].

Story creation process. The process of creating visual data stories can be abstracted into four steps [Lee+15].

- First, the data on which the story should be based needs to be explored and analyzed to extract *story pieces*, i.e., "specific facts backed up by data" [Lee+15].
- Then, logical connections need to be made between the story pieces, and they need to be ordered to create a narrative structure.
- Based on this structure, the story can be built, including suitable visual representations.
- Lastly, the story has to be shared with the audience, ideally allowing the readers to provide feedback.

Depending on the author, the first three steps of this process often take place in a non-linear and iterative fashion, so the tools used during development should support such an iterative workflow [Ami+15].

Guidelines for visual storytelling. Structuring the narrative of a story to follow a story arc can help to engage the viewers. For instance, the popular Freytag's Pyramid arc begins with some initial exposition, then rising the tension up to a climax and delivering a resolution afterwards. This structure can also be applied to visual data stories, e.g., by providing context information for the data in the exposition, gradually introducing more visual elements to rise the tension, and summarizing the key takeaway messages and next steps to resolve the story [Yan+21].

When introducing new visual elements or changing a view in any other way, animating the transitions can lead to less attention drift and better comprehension for the users [Ami+18] and increase their engagement [McK+17]. Moreover, breaking up animations into multiple transitions can improve recall and understanding [HR07], which was also shown specifically in the context of spatio-temporal data [NPD17].

Another set of suggestions in the context of geographic data-driven stories was proposed by Latif et al. [LCB21]. Regarding the combination of visualizations and textual narratives, they suggest to use vernacular geographic language in the text, like in "the southern tip of a country." They also propose to verbally summarize observations made for groups of elements, and to explicitly mention elements which are different, like pointing out outliers or extrema and contrasting them to the norm. They also advocate providing sufficient background, making the story personal via quotes and opinions, and adding various forms of annotations on top of visualizations, like captions, annotations, and tooltips. Moreover, they suggest sequencing the visualizations to move from overview to detail, explicitly verbalizing in the text what is visualized, and placing the text close to its corresponding visualization.

In addition, *narrative design patterns* can strengthen the narrative of a story [Bac+18], e.g., via mechanisms for engaging the audience and creating an appropriate flow through the story. If a story *flows* well, it can increase the user engagement [McK+17]. The *narrative flow* can be supported by providing appropriate navigation options, immediate navigation feedback, or animating the transitions in visualizations [McK+17].

Authoring tools. The theoretical and empirical knowledge about how to write visual data stories can be used to create *authoring tools* to support story authors by simplifying and automating the story creation process. A survey of authoring tools was presented by Chen et al. in 2023 [Che+23]. However, it showed that tools to create complex stories based on spatio-temporal data are still lacking. While some semi-automatic approaches exist, like tools for designing narration-focused camera movements for geographic visualizations [Li+23], overall, there is still a gap between fully automated story generation and manually created stories [Sun+23]. Another beneficial approach is to abstract the structure of similar kinds of stories into story *templates* [Meu+22]. These templates can then be used to facilitate the creation of new stories.

Summary and contribution. In this section, we gave an outline of how visual storytelling can be performed based on spatio-temporal data. It became clear that not a lot of research exists focusing specifically on spatio-temporal characteristics, let alone on conflict data. This is in contrast to the fact that many visual data stories exist on the web reporting on geopolitical conflicts, published, e.g., by *The New York Times* or *The Guardian*. Accordingly, while general research insights can also be applied to a spatio-temporal setting, it would be beneficial to further investigate specifically how to perform visual storytelling based on spatio-temporal data.

Ideally, existing generalized literature should be adapted for this purpose, combining it with insights drawn from analyzing spatio-temporal data stories collected on the web. We did so by conducting a study in which we analyzed 130 stories and characterized them using a design space that we derived by merging and adapting three existing design spaces.

2.7 COMBINING EXPLORATION AND EXPLANATION

In the previous section, we already mentioned that explanatory approaches from visual storytelling can be enriched by interaction techniques providing controlled exploration. In this section, we discuss in more detail how to combine exploratory and explanatory approaches. We do so by depicting the influence that interaction can have on engagement, before expanding on the field of science communication, in which the combination of exploration and explanation is a powerful approach for educating an interested audience. We end by showing that such approaches are primarily used to communicate scientific insights to broader audiences, while they can also support the communication between researchers directly.

Engagement and interaction. When communicating complex topics, Böttinger et al. advocate engaging the users, proposing various approaches for doing so [Böt+20]. In general, characteristics of a visual data story that influence user engagement include the story's aesthetics, whether it can spark the users' interest, and how well it can keep their attention [HP17]. Yet, it is unlikely for a story to captivate users nonstop. Rather, they go through cycles of engagement and disengagement [OT08]. Other important aspects for engagement are how users can control the story and discover new insights [HP17]. Allowing them to interactively explore aspects of the story can be particularly beneficial. Corresponding interaction techniques were on the rise in online visual storytelling up until around 2016 [Sto+16]. However, a phase of decline was signaled by Tse's statement in 2016 [Tse16]. Tse explained that *The New York Times* reduced the amount of interactivity provided in their online stories because their readers were not using it enough to justify the additional implementation costs that interactive components require.

Yet, various points for the benefits of interactivity were made. Accordingly, Hohman et al. extensively discuss how online articles can use interaction to engage users, potentially improving learning outcomes [Hoh+20]. Hohman et al.'s work itself is an interactive article published on *Distill* [Dis]. This website, along with others, like *explorable explanations* [Exp], provides a variety of examples for articles communicating scientific content in an interactive and engaging way. Overall, the biggest problem with interactivity occurs if important information is hidden behind interaction [Ais23]. If this pitfall is avoided, interaction can be used to dig deeper into the data and build trust with the story [Ais23].

Accordingly, under the term *exploranation*, Ynnerman et al. investigated how to combine interactive exploration and explanation [YLT18]. To perform exploranation, they recommend augmenting explanatory applications like visual data stories with small, interactive environments. These *interactive microenvironments* can arouse the users' curiosity by allowing them to explore the explained content. At that, the interactions should be constrained in a way that the results can be easily interpreted by users while also being expressive. Additional visual elements like annotations can help to embed the exploration results in the overall narrative, but they should be visually subordinated to the central exploratory view [YLT18].

Ynnerman et al. also give advice on what to consider when using exploranation in a *synchronous* [Lee+15] setting, where a single presenter interacts with a story application to present insights to an audience. In this case, the audience members cannot interact with the application directly. While such settings are relevant, we do not focus on them in our work. Instead, we restrict the scope to *asynchronous* [Lee+15] and *author-driven* [SH10] settings.

Science Communication. Combining approaches from visual exploration and explanation is particularly beneficial when communicating scientific insights. This was also the setting of the applications that Ynnerman et al. reported on [YLT18]. In this context, Meredith advocates using informative and engaging visuals for communicating scientific information to improve the comprehension and engagement of the audience [Mer21]. Employing storytelling approaches to communicate science, like appropriate story arcs [GGCM18], can even improve the memorization of the conveyed information while making it more enjoyable [NL10].

While these benefits would also be interesting for communicating science between researchers, such approaches are usually rather used for communicating insights to a broader lay audience [Mer21]. Little research exists focusing on an expert audience, probably also because it is sometimes considered "unscientific" to use storytelling approaches in scientific contexts. However, writing scientific articles in a more enjoyable and interesting way does not automatically reduce the information gained from them [Hun79]. An example that comes close to such a setting was proposed by Cortes Arevalo et al., who used visual data stories to communicate science to an audience with more domain expertise [CA+20]. Yet, their audience consisted of domain practitioners, not researchers directly.

When targeting an audience consisting of conflict researchers, the users' numeracy can be expected to be advanced. However, even scientists can struggle with numeracy [Nat17]. In such situations, it is recommendable to reduce the cognitive effort, e.g., by explaining what certain numbers mean, and drawing attention to important information [Ins14]. The ways in which audience members interpret new information also depend on their beliefs, values, and ways of understanding the world, which can be shaped by various factors [Nat17]. The corresponding mental models of the audience members can be expected to better support a correct understanding of conveyed concepts if they have more domain expertise [Nat17].

Summary and contribution. While the combination of exploratory and explanatory approaches is beneficial for science communication, corresponding applications are primarily designed with broad audiences as the target. However, domain experts can also benefit from such approaches.

For this reason, we created an interactive visual data story to analyze how exploranation can be used to explain a scientific method to conflict researchers.

2.8 SUMMARY: FROM EXPLORATION TO EXPLANATION

In this chapter, we provided background regarding typical characteristics of spatio-temporal conflict data, how it can be visualized, and what kinds of interaction can be used to explore the visualizations. Based on that, we presented three directions in which these concepts can be applied to support the analysis and communication of conflict data. The three directions can be located on a conceptual spectrum regarding the *freedom of discovery* they provide, ranging from exploration to explanation, with exploranation placed between the two ends, see Figure 5. For this work, we conducted three main studies to investigate the three directions. At that, the target audience in the studies revolving around exploration and exploranation consists of domain experts, while the study dealing with explanation considers broad audiences. We re-use the corresponding iconography from Figure 5 regarding the degree of freedom of discovery and the target audience throughout this work.

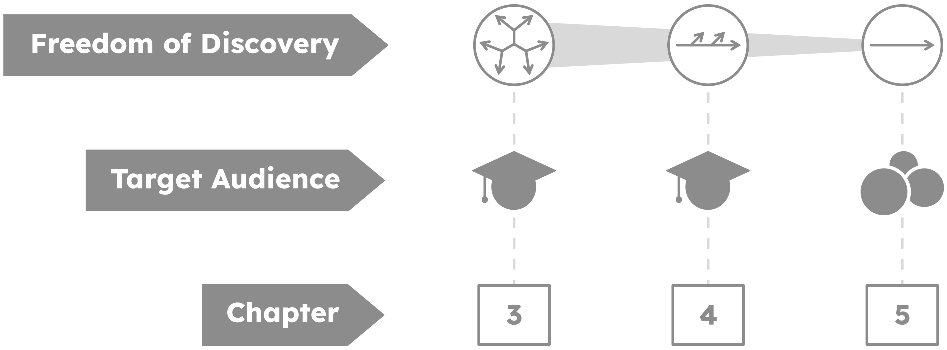


Figure 5: Across the three central chapters of this work, we focus on different degrees of free-dom (exploration, exploranation, and explanation, from left to right) and on different types of audience (domain experts (left and center) and broad audience (right)).

Part II

FROM EXPLORATION TO EXPLANATION

We continue with the core chapters of the dissertation. To maintain transparency about what was included in the underlying publications and accordingly peer-reviewed, we keep the information within the main part of each chapter close to the original publications. However, we have added an Addendum section at the end of each chapter. In it, we present further analyses we conducted as well as discussions about how the content of each chapter fits into the scope of the overall dissertation.

3

EXPLORATION FOR EXPERTS

The analysis of large-scale conflicts requires a comprehensive picture of the related events. Therefore, it is beneficial to combine the information from multiple conflict event data sets. To do so, researchers have developed a semi-automatic matching algorithm to integrate event data of different origins into one comprehensive data set based on hierarchical taxonomies. The validity of the corresponding integration results is not straightforward to assess, since the results depend on user-defined input parameters and the relationships between the original data sources. While exploratory visualization can be highly useful in this context, only rudimentary visualization techniques have been used so far to analyze the results, allowing no trustworthy validation or exploration of how the final data set is composed.

To overcome this problem, we developed VEHICLE*, a web-based application to validate and explore the results of the hierarchical integration of conflict event data sets. For the design, we collaborated with a conflict research expert to identify the underlying domain problems and derive a task and workflow description. The application combines both traditional and novel visual analysis techniques, employing statistical and map-based depictions as well as advanced interaction techniques. We show the usefulness of* VEHICLE *in two case studies and an evaluation together with five conflict researchers, confirming domain hypotheses and generating new insights.*

This chapter is based on the following contribution [May+21], John Wiley and Sons. Used with permission under License Number [5891970309247]:
B. Mayer, K. Lawonn, K. Donnay, B. Preim, and M. Meuschke. "VEHICLE: Validation and Exploration of the Hierarchical Integration of Conflict Event Data." In: *Computer Graphics Forum* 40.3 (2021), pp. 1–12. DOI: 10.1111/cgf.14284.

3.1 INTRODUCTION

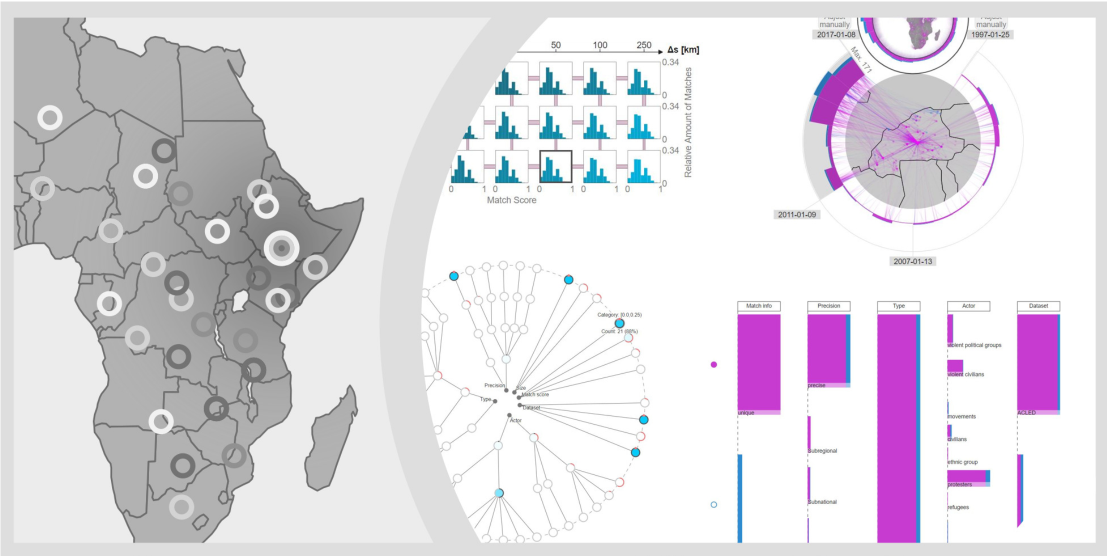


Figure 6: *VEHICLE* is a web-based application for validating and exploring the integration of conflict event data sets that were recorded by different institutes. It allows analyzing parameter influence as well as spatial, temporal, and hierarchical distributions regarding the integration process. The visualizations in the right half depict snapshots from an analysis focusing on protest events that took place in Burkina Faso. *(Figure adapted from Mayer et al. [May+21], John Wiley and Sons. Used with permission under License Number [5891970309247].)*

When and where do riots take place? Why are some protest campaigns only temporarily active, and others develop into a widespread civil war? Questions like these are of great interest to conflict researchers. Over the last decades, an increasing number of data sets have been produced that encode the emergence and progression of conflicts. Conflict phenomena are recorded as point-based events in relation to space and time. Previously, the data used for analysis was often confined to a single data set. However, for a holistic view on the course of events and, hence, a better understanding of their interdependency, it is required to consider the information from more than one data set. The different data sets, however, are not all recorded using the same coding scheme. Institutes collect information from different sources like newswire or newspaper articles to extract the event recordings for their data sets. Therefore, it is possible for recordings from different data sets to represent the same original incident while containing different information.

To solve this issue, experts have recently started to integrate the information from different data sets to receive one holistic set. The most prominent method to do so is the semi-automatic method *MELTT* [Don+19]. It is parameter-dependent and relies on hierarchical taxonomies to classify the events. However, validating the plausibility of the outputs and understanding their composition is vital as the resulting data set is the foundation for all further analysis and inference. Exploratory data visualization is an effective tool for diving into complex data and identifying patterns in them [AA06]. Yet, only basic techniques have been used, so far, to validate the results of the integration process.

Therefore, we present *VEHICLE*, a web application to analyze the results of hierarchically integrated conflict event data. We designed it in collaboration with a domain expert following Munzner's nested design model [Mun09]. It allows the inspection of the influence of input parameters on the integration as well as the characteristics and similarities among the integrated data. For this purpose, we use multiple linked interactive visualizations, such as map representations and radial glyph-based layouts, see Figure 6, right. We demonstrate the effectiveness of the application by presenting two case studies and an evaluation together with five domain experts. In summary, we make the following contributions:

- A characterization of the integrated conflict event data and the associated domain problems regarding the validation and exploration of its composition.
- A task and workflow abstraction to translate the problems into the field of interactive data visualization.
- An application design and implementation following the derived workflow to solve the identified tasks.
- An evaluation with five conflict researchers to validate the usefulness of the application.

VEHICLE is available as a web application [Vehb], and a video providing an overview of it is available on YouTube [Veha].

3.2 RELATED WORK

We put the literature introduced in the Background on Exploration more closely into the context of this chapter.

In our study, we analyzed the integration of geospatial event data with point spatial footprint and instant duration. For the integration, the events are organized in a hierarchical structure. We structure this section accordingly.

3.2.1 *Integration of Conflict Data*

The analysis of integrated data is of critical relevance for the quantitative study of conflict, given the incomplete and often complementary coverage of individual data sets [Don+19]. The hierarchical integration of event data from different sources according to Donnay et al. [Don+19] requires at least two hierarchical taxonomies describing different aspects of the data. Then, the assumption is that, the deeper in the different taxonomies two different event recordings fall into the same categories, the more likely it is that they represent the same original incident.

Based on that idea, duplicate recordings are eliminated when integrating multiple data sets. This automatic strategy is more efficient and replicable than attempting manual integration. Yet, it cannot be precluded that potential biases from a single data set might carry over into the integrated data set [Wei15]. This opens up to a more general problem, that is, that studies in conflict research typically rely on techniques to analyze the outcome of the integration that are not fully suited to capture its whole complexity. For instance, prior work has considered mainly time series visualizations and corresponding statistical analyses, or static maps aggregated over fixed time windows [DF14; Wei16; Don+19].

3.2.2 *Exploratory Analysis of Spatio-Temporal Conflict Event Data*

For the development of our application, Andrienko and Andrienko's book on the exploratory analysis of spatio-temporal data provided us with structured guidance to identify principles, tasks, and techniques to apply in our design [AA06]. Existing applications for the visual exploration of conflict events often do not go into sufficient depth to derive complex insights from them [WK09; Acl; Gtda]. There also exist approaches that are more advanced [Rob+17; Sil+19], however, they do not provide the option to jointly analyze information coming from different sources.

Applications that do allow the joint analysis of different data sources were proposed by Saldarriaga et al. [SKB17] and Lu et al. [Lu+16]. Yet, they cannot be used to merge data sets of similar shape, but they are rather used to analyze complementary types of data in combination, like satellite imagery, movement data, and videos [SKB17], or event recordings and RSS feed data [Lu+16].

3.2.3 *Visualization of Hierarchical Data*

Hierarchical data captures relationships between sub- and superordinate components. They can be visualized explicitly or implicitly [SS06]. The most common explicit representation are node-link diagrams [RT81], where nodes are connected by edges to express hierarchical relations between them.

Implicit techniques are more space-efficient as they use alignment instead of edges to encode the relationship [SHS10]. This way, they provide a space-filling view [SS06]. The visual marks of subordinate levels can be arranged side by side to their parent mark or inside it. One example of the side-by-side arrangement are icicle plots [KL83], where the hierarchy of rectangular shapes is expressed by stacking them.

To save even more screen space, visual marks can be displayed inside their parent mark, producing a recursively self-containing arrangement. The most well-known representatives of this class are treemaps [JS91]. They express the hierarchy through nested rectangles. Enhancements of the visual representation include cushion treemaps [VV99] and Voronoi treemaps [BD05]. Despite improvements, arrangements based on self-containment make it difficult to compare individual elements of a hierarchy if they are not placed close to each other. In addition, it is difficult to trace paths through the hierarchy. Hence, we rely primarily on explicit node-link diagrams and implicit side-by-side representations in our visualizations of hierarchical data.

3.3 CHARACTERIZATION OF THE DATA

The integrated data sets that we analyze in *VEHICLE* are the Armed Conflict Location and Event Data (ACLED) [Ral+10], the Uppsala Conflict Data Project – Georeferenced Event Dataset (GED) [SM13], the Global Terrorism Database (GTD) [Gtdb], and the Social Conflict Analysis Database (SCAD) [Sal+12]. We consider the events from these data sets that were recorded in Africa between 1997 and 2016, a total of 197,502 events. Out of these, 140,738 belong to the data set ACLED, 25,788 to GED, 16,928 to GTD, and 14,048 to SCAD. In the following, we define terminology to introduce the integration procedure and to classify the resulting data. The integrated data can be exported and downloaded on the same website [Vehb] as *VEHICLE*.

3.3.1 *Definitions*

In the following, the original occurrence of an event in the real world is referred to as an *incident*, while the recording of an incident in a data set is referred to as an *event*. We make this distinction as a single incident may be represented by multiple events. The set of *attributes* of an event depends on whether it is encoded in one of the original data sets prior to the integration or in the context of the integrated data. In the original data sets, events have *continuous attributes*, namely the location, i.e., longitude and latitude, and *discrete attributes*, namely the date of occurrence. In addition, they have *categorical attributes* to describe the event type (e.g., strategic development or riots), the main actor, and how precise the recorded location is.

For the integration of conflict data sets, hierarchical *taxonomies* were introduced by Donnay et al. [Don+19]. Accordingly, one taxonomy exists for each of the attributes: "event type" (*type*), "primary actor" (*actor*), and "geographic coding precision" (*precision*). Each of these taxonomies has multiple *levels*, grouping the values of the corresponding attribute from the original data sets into overarching *categories*. An example of how an event can be encoded in the event type taxonomy is given in Figure 7.

The event type taxonomy consists of four levels and 28 categories, the actor taxonomy of three levels and 21 categories, and the precision taxonomy consists of four

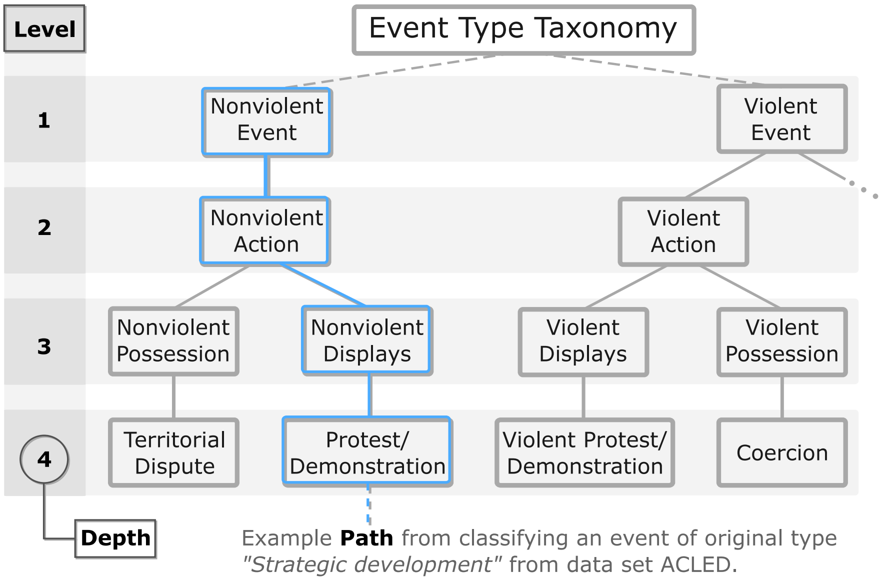


Figure 7: An extract from the taxonomy for the attribute "event type" and corresponding terminology [Don+19]. *(Figure adapted from Mayer et al. [May+21], John Wiley and Sons. Used with permission under License Number [5891970309247].)*

levels and 19 categories. The actor taxonomy branches up into violent groups and non-violent groups, which branch up further, ending in categories like "armed groups," "protesters," or "refugees" on the deepest level. The precision taxonomy branches up into multiple subtrees, on the highest levels encoding on which overall level of detail the event was recorded, ranging from "national" over "subregional" to "unknown." These subcategories, in return, branch out into more detailed subtrees, which end in categories such as whether the event was recorded in admin regions of a specific order, in a fuzzy region around a city or village, or at an exact location.

When an attribute value of an event from one of the original data sets is classified in the corresponding taxonomy, the value is assigned to one *path* of the taxonomy. A path is a sequence of categories visited when traveling from a category at a higher level to a category at a subordinate level, see Figure 7. The higher the level is in the tree, the lower the corresponding level number is. The *depth* of a taxonomy is defined as its number of levels, which also corresponds to its maximum possible path length.

The taxonomies are used to integrate the events from the original data sets, as explained in more detail in Section 3.3.2. As a result, events are either identified as *matches* or *uniques*. A *match* is a tuple of events from different data sets that were identified to represent the same original incident. A *unique* is an event that was found to be the only one out of all the events to be covering a certain incident.

In two views of *VEHICLE*, a taxonomy of primary interest can be selected for the investigation, e.g., "type." It is then referred to as the *primary attribute* and the categories at its highest level as the *primary values*, e.g., “violent events” and “non-violent events” in Figure 7.

3.3.2 *Matching Procedure*

For the application in this project, we adopted the matching procedure and the underlying taxonomies from the *MELTT* algorithm [Don+19]. It has two primary input parameters: Δs and Δt. These parameters control how far two different events may occur apart from each other regarding time (Δt) and space (Δs) to still be considered as candidates for a match.

For the events proximate to each other based on these constraints, the algorithm identifies the pairwise similarity between each of them. This similarity of two events, called *match score*, depends on how deep into each of the different taxonomy trees the two events have their deepest common parent node. The deeper the first common parent, the more similar are the events. The score produces values between 0 and 1, while a smaller score corresponds to higher similarity. Finally, based on the pairwise similarities, the events which are most similar are identified as matches by solving a stable marriage problem [GS62].

After bringing the original data sets to a uniform shape, we used this algorithm to calculate and add the matching information to the data (140MB in total). It took about three hours on a standard desktop computer.

3.3.3 *Classification of the Data*

For the *data abstraction* [Mun09], we adopt the taxonomy for time-oriented data proposed by Aigner et al. [Aig+23]. Since the data covers a sequence of consecutive years, it is *ordinal*, *point-based*, and *linear*. As multiple events can be recorded on the same day, it has *multiple perspectives* and since the dates are expressed in a calendar system, the data has *multiple granularity*. The time primitives are *instant* and *determinate*. Additionally, the data is both *quantitative* due to the continuous and discrete attributes and *qualitative* due to the categorical attributes and, consequently, also *multivariate*. The frame of reference is *spatial*. The internal time of the events is *temporal*, and the external time in our setting is *static*.

The data is graph-based due to the hierarchical information from the taxonomies and the information which events form matches together. It can be both large-scale with almost 200k events and small-scale, depending on the size of the inspected subsets.

3.4 TASK ABSTRACTION

In the following, we present our *domain problem characterization* [Mun09]. Together with a conflict research expert, we identified five domain problems that occur when analyzing the hierarchical integration of conflict event data. They consist of validating the algorithm outcome, searching for potential malfunctions, determining an appropriate set of input parameters, getting a first understanding of the patterns in the data, and exporting a subset of the data for further personalized investigation. From these problems, we distilled the overall domain-specific *analytical tasks* that may arise when trying to solve the problems.

- **T1: Understand the influence of the parameters Δs and Δt on the matching result.** An appropriate choice of Δs and Δt is not known in advance, only ranges of reasonable values can be estimated. In our case, they reach from $\Delta s = 0\text{km}$ to $\Delta s = 250\text{km}$ and $\Delta t = 0\text{d}$ to $\Delta t = 2\text{d}$. As the choice of the parameters influences the number and the quality of the matches, it is important to see how these two characteristics change over the different outcomes. In addition, it is useful to provide means to assess how strongly the algorithm outcomes differ from each other. This allows determining which parameter changes lead to the largest variation of the outcome and are, thus, interesting for closer inspection.
- **T2: Understand whether the number and structure of the identified matches are reasonable.** To assess the credibility of the matching result, the analysts need to understand the underlying structure and rules of when and where matches occur, and why it is not the case in other regions where unique events prevail. For that, they need to solve the following two subtasks.
  - **T2a: Analyze and compare the distribution of matched and unique events.** A multitude of questions may arise when investigating the character of matched and unique events. The analysts need to get an understanding of the circumstances under which matches are identified. When are they found? Where? For what kinds of events? In contrast, are there regions where multiple events occur in close proximity, but barely any matches are identified? And what is the reason for that? To answer such questions, analysts need to be able to approach their investigations from various angles. In addition, it would be beneficial to determine what the most striking differences between matched and unique events are.
  - **T2b: Determine where in the taxonomies the events are matched with which frequency.** For the analysts, it is necessary to grasp the distribution of the matches across the categories of the different taxonomies as well as to compare the frequencies of individual categories to identify both overall patterns and outliers. This way, it becomes clear in which subtrees of the taxonomies the data sets overlap and for which categories they do not.

- **T3: Inspect and export subsets of interest.** As already touched on in **T2a**, analysts need to narrow down the set of events or matches to inspect subsets in more detail. In doing so, the analysts should be able to inspect the events both on a large scale but also on a small scale of only a few events to investigate overall trends as well as small anomalies. At that, resetting to previous views is important to minimize the consequences of operating errors and to compare and refine subsets. Moreover, for subsequent personalized investigation, experts need to be able to export identified subsets of interest.

Refining the taxonomy or core aspects of the matching algorithm aside from the input parameters are domain problems outside the scope of our application. This would substantially increase the complexity of the application while it showed that not all users would want to analyze the results that deeply. For instance, the original data sets have 8452 types of actors in total. Hence, when modifying the actor taxonomy, large numbers of original actor classes might have to be reassigned, requiring dedicated solutions. We also do not aim for providing techniques for in-depth analysis of the integrated data set such as identifying causal effects. The techniques applied for such tasks vary strongly in the field of conflict research and would cause the application to become too complex.

3.5 THE DESIGN OF VEHICLE

The visual interface of *VEHICLE* comprises multiple linked components to handle the multi-faceted integration data. We developed our application in an iterative process using prototypes [LD11] and implemented it using the JavaScript library D3.js [D3]. In this section, we describe our *operation abstraction* and our *visual encoding and interaction design* [Mun09]. The different components of *VEHICLE* facilitate a workflow which we present before discussing their design in detail.

3.5.1 *Workflow*

Based on the analytical tasks from Section 3.4, we designed a workflow, including applicable *subtasks* (*emphasized* in the text below) according to Andrienko and Andrienko [AA06]. We also reference the sections describing how we implemented the different aspects of the workflow.

(A) Overview of Parameter Influence. To get an overview of the algorithm outcomes (**T1**), the analysts first inspect the matching results for different parameter combinations at a higher level (Section 3.5.2.1). To do so, they *look up* and *compare* the overall quality of the identified matches and the similarity of the different outcomes. They *characterize* the overall *behaviors* and their relations created by varying the parameters. This way, the analysts identify the parameter combination which is of most interest for a more detailed analysis. As the remaining views are still hidden at this point, the analysts are guided from overview to details. During the subsequent analysis, the analysts may come back to the overview to compare the identified patterns to the ones occurring for other parameter combinations.

(B) Analysis of All Events. For a specific matching result, the analysts *characterize*, *compare*, and *relate* the patterns describing the distributions of the events across the different attributes (**T2a**, **T3**). This includes the selection and *comparison* of subsets of events with respect to space, time (Section 3.5.2.2), and categorical attributes (Section 3.5.2.3). Additionally, they determine what the most striking differences between certain subgroups are. Throughout the investigation, they select the level of detail at which to display the taxonomy information. To avoid getting lost in abstract visualizations, they refer to a map-based view depicting the current selection. Following the above steps, the analysts also gain an impression of the extent to which the data sets overlap for the different categories. To analyze the match information in more detail, the analysts alternate between (**B**) and (**C**).

(C) In-depth Analysis of the Match Distribution. For the selected parameter combination, the analysts *look up* and *compare* where and when the matches occur (Section 3.5.2.2) as well as how they are distributed across the taxonomies (Section 3.5.2.4) to gain an understanding of the underlying patterns (**T2b**). Again, a map-based view serves as a reference frame of the current selection. Aside from gaining an overview of the match distribution, the analysts search for and investigate subsets of matches (**T3**). Interesting subsets are usually determined by categories with an unexpectedly high or low number of matches. The analysts may also investigate how the matched events of the selected subset *compare* to the unique events. In addition, they analyze the relationships between the categories to understand what the determining factors for certain kinds of matches are, e.g., for those with a particularly good matching score.

3.5.2 *Visual Design and Implementation*

The design of *VEHICLE* follows the workflow above. The view from (**A**) is always visible. Once a parameter combination is selected, the further analysis scope can be specified in the bottom bar in Figure 8. The scope can either be selected as "All Events" (**B**) or "Matches Only" (**C**). Both subviews replace each other, while each of them contains a version of the spatio-temporal reference. If the scope is set to "All Events," a primary attribute has to be selected in addition. Any performed filtering can be undone and redone. Additionally, the data of the current view can be exported via "Export Data" in the bottom bar in Figure 8 (**T3**). Moreover, all components provide a help button to explain their key functionality. We allowed direct user interaction [Shn97] wherever possible.

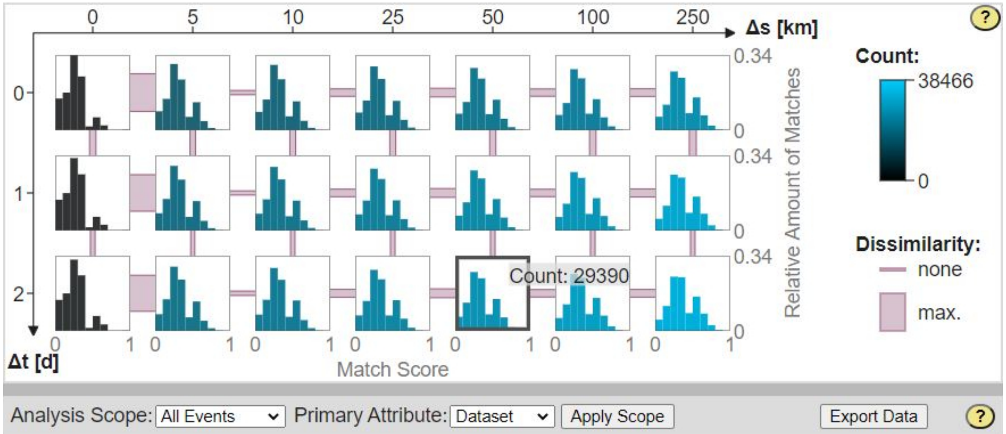


Figure 8: The ParaMultiples summarize the matching outcomes in histograms. They depend on spatial (horizontal axis) and temporal (vertical axis) constraints. In addition, the dissimilarities between match distributions of adjacent outcomes are depicted as bars. *(Figure adapted from Mayer et al. [May+21], John Wiley and Sons. Used with permission under License Number [5891970309247].)*

3.5.2.1 *ParaMultiples*

The first component provides an overview of the influence of the spatial and temporal parameters Δs and Δt on the matching results, see Figure 8. The analysts can compare how the matches are distributed for pre-defined parameter selections in small multiple histograms, the *ParaMultiples*. Each histogram corresponds to one parameter combination and depicts the relative frequency of the corresponding matching scores (mapped to the y-value) and the total number of matches (mapped to the fill color). Mapping the absolute counts to the y-values would be inappropriate as the number of matches varies strongly between the parameter combinations, resulting in very small bar heights for certain results. In addition, the absolute counts are displayed when hovering over the histograms. This way, analysts can look up and compare the overall match quality for different input parameters and their impact on the number of matches (**T1**). We selected histograms as they are particularly well-suited for those tasks [AA06].

Dissimilarity Score. Adjacent histograms are connected by bars representing the dissimilarity of the match distributions for the corresponding parameter combinations. For the dissimilarity measure, we considered that for every matching result, each taxonomy induces a weighted tree. In this tree, the categories are the nodes and the node weights are the relative number of matches identified in each category. Existing tree similarity scores [BBY03; Bil05; YKT05] are not suitable, as they focus on a more general comparison of trees where the trees are of inherently different structure. Hence, we created a score tailored for comparing two trees of identical structure where subtrees can be considered similar even if their weights are not distributed exactly across the same nodes.

With the score, two kinds of similarity should be measured (**T1**). Firstly, a vertical similarity, measuring if the matches are identified in similar depths of the trees. This reflects whether the matches have similar scores as the scores improve if the matches are identified at deeper levels of a taxonomy. Secondly, a horizontal similarity, measuring whether the matches are distributed across similar subtrees of the taxonomy trees. In the example in Figure 7, this refers to, e.g., whether both trees cover a similar number of matches of violent events in general as compared to non-violent events.

Accordingly, to calculate the dissimilarity between two weighted trees of the same taxonomy, first, an accumulated version of each tree is created:

- Starting at the deepest level, for each node, half of its weight is added to its parent node's weight.
- This is done recursively until level 1 of the taxonomy is reached. This way, the weights of the nodes at the deeper levels of the tree influence the weights on the higher levels, but for each level higher up, the influence is halved recursively.
- Eventually, the weights of the two accumulated trees are subtracted node-wise and the absolute values of the node-wise results are summed up.
- To receive the dissimilarity between the trees, the sum is normalized by dividing by the maximum possible value that can be reached for two trees of the given structure.

This way, the resulting score has a value between 0 and 1, with 1 corresponding to the highest possible dissimilarity. Since for each matching result, there are three taxonomy trees, corresponding to *type*, *actor*, and *precision*, we receive three separate dissimilarity values when comparing two matching results, one for each tree. To determine the overall dissimilarity between two matching results, we form the average of the three separate dissimilarity values, yielding values between 0 and 1. To provide as much contrast as possible, the scale for the width of the dissimilarity bars goes from 0 to the maximum dissimilarity present in the view.

3.5.2.2 *TempMap*

In this view, the spatio-temporal distribution of the events is displayed in a comprehensive way (**T2a**, **T3**), see Figure 9. The location of each event is encoded on a map. In addition, the temporal information is mapped to a radial stacked histogram that is wrapped around the map. Aggregation of time-dependent events into histograms is a common visualization technique [And+07]. If a primary attribute is selected in the "All Events" mode, the bars of each stack correspond to the different values of

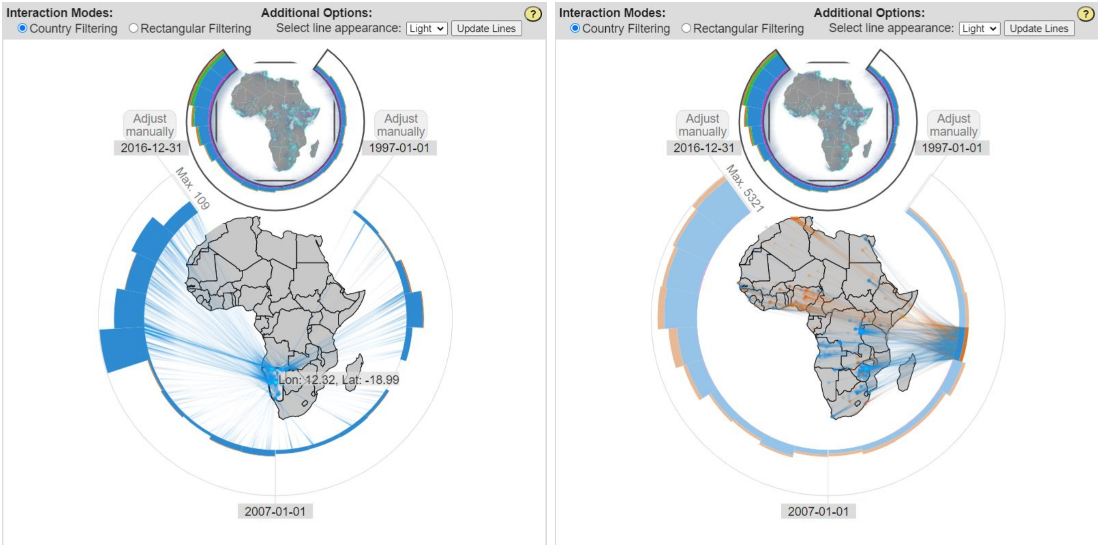


Figure 9: The TempMap shows the spatio-temporal distribution of events. Hovering over a country (left) or a temporal bin (right) highlights the contained events. Hovering over the map also displays the longitude and latitude of the hovering position (visible in the left view). In the right view, a strong spatial separation is revealed between the events encoded in orange and the events encoded in blue. *(Figure adapted from Mayer et al. [May+21], John Wiley and Sons. Used with permission under License Number [5891970309247].)*

the primary attribute. If "Matches Only" are analyzed, each stack contains only one bar. To assess which events took place at what time, each event location is connected by a semi-transparent line to the corresponding point on the outer timeline. Thus, by inspecting in which direction the lines leave the event locations, the dates of the events can be estimated. The technique is inspired by the TimeWheel [TAS04] and the RingMap [ZFH08]. It was implemented similarly by Tominski and Schumann [TS20], improving the version of Tominski et al. [Tom+12].

Line Design. As large numbers of events need to be displayed, visual clutter would be caused by representing each event with a simple point and line mark. To avoid this, we adjust the opacity of the marks depending on the number of events displayed [ED07]. With x representing the number of events in the current set, the formula to calculate the alpha values is

$$\alpha(x) = c + m(x+b)^{-k}, \text{ with } c = 10^{-3}, m = 2 \cdot 10^4, b = 2.5 \cdot 10^3, k = 1.5. \qquad (1)$$

We determined the parameters by manually adjusting the alpha values for event subsets of various sizes and fitting the function to the measured values. At that, we made a trade-off between preserving the mark visibility for small sets and reducing visual clutter for large sets. Figure 10 depicts the measured values (interpolated linearly) in blue and the alpha function that we fit to it in orange. For an idea of how the differ-

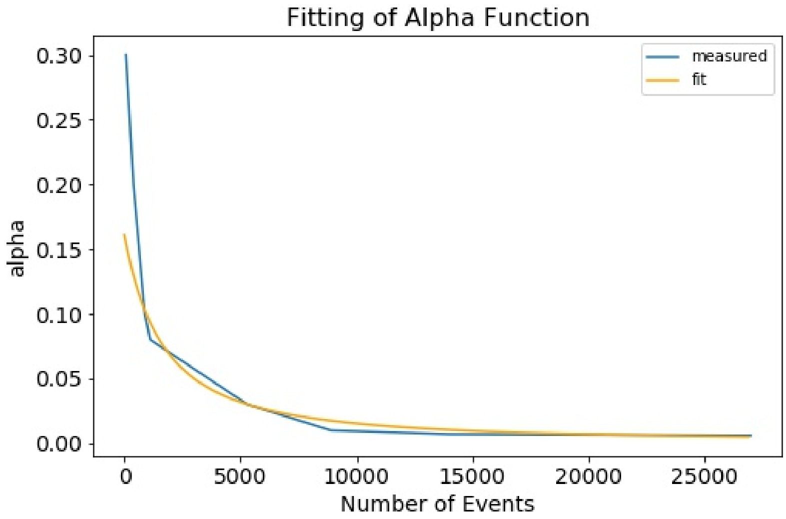


Figure 10: The linear interpolation of the measured values is displayed in blue, and the fit alpha function is displayed in orange. *(Figure adapted from Mayer et al. [May+21], John Wiley and Sons. Used with permission under License Number [5891970309247].)*

ent parameters influence the shape of the alpha function, please refer to Figure 11. In

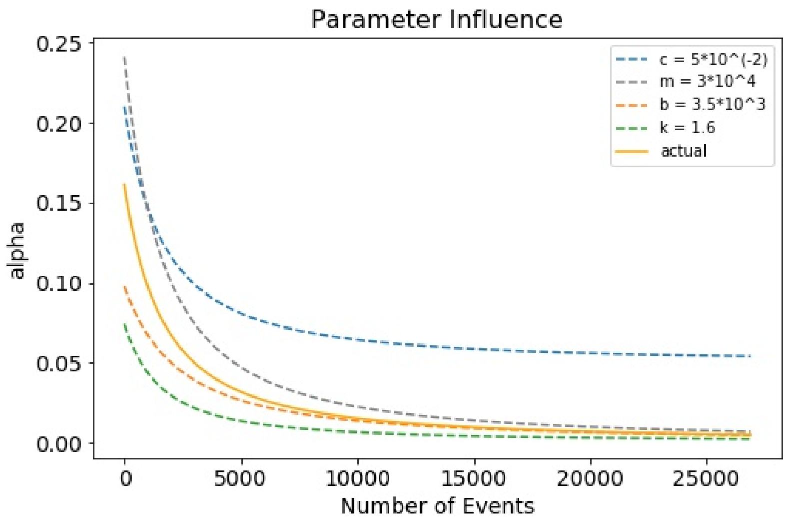


Figure 11: The influence of the different parameters on the alpha function when increasing them separately. *(Figure adapted from Mayer et al. [May+21], John Wiley and Sons. Used with permission under License Number [5891970309247].)*

addition, we imposed a lower limit for $\alpha(x)$ of 0.002.

To further reduce clutter, the line marks fade to 20% of the $\alpha(x)$ value at their midpoint, meaning they are most visible at the endpoints. The analysts also have the option to adjust the opacity to improve the visibility or to hide the lines entirely, keeping only the point marks. Moreover, a highlighting functionality is available when hovering over individual countries or temporal bins to quickly focus on the related events, as depicted in Figure 9.

Layout. We use a radial layout similar to other applications [ZFH08; TAS04]. It facilitates the search for spatio-temporally dense regions and the look-up of when individual events have occurred via the line marks. Moreover, smoothly integrating closely-related aspects of the data into a single context can facilitate the perception of the combined information as a whole [CCF95]. Thus, we support a quicker and more holistic assessment of the data by allowing to grasp the spatio-temporal context comprehensively. To prevent the temporal information from being interpreted as having a cyclic context due to the circular layout, the timeline is not closed as a ring but left open in the area where a minimap is placed.

The minimap provides an overview when filtering and zooming into the data spatially or temporally, as it maintains the initial view while indicating the zoomed regions. To filter and zoom spatially, the analysts can either use a rectangular brush or click on a country to focus only on events within it. To filter and zoom temporally, they can either select the interval by brushing in the region of the temporal labels, see Figure 6, or by entering precise begin and end dates.

Alternative Techniques. We did not implement spatial aggregation techniques to reduce the visual complexity since events should be distinguishable even in small regions. This would require a high aggregation resolution, making it ineffective. Moreover, we do not use an additional line encoding to express which events were matched. As most matched events naturally occur in close proximity, the resulting possible insights are too small to justify the associated increase in visual complexity. Another option would be to arrange the stacks of the histograms next to each other instead of on top. This would improve the separability of the corresponding lines but distort the temporal ordering. However, we want to faithfully assess the proximity of events, so temporal distortion is not an option. The same holds for spatial distortion.

3.5.2.3 *EventCharts*

To analyze how the events are distributed across the categorical attributes, the *EventCharts* use hierarchical stacked barcharts, see Figure 12. Barcharts are effective for lookup and comparison tasks [AA06] and do not require high visual literacy. The view is available when analyzing "All Events." For a selection of categories from the taxonomies, bar stacks are displayed. Moreover, the information in which original data set each event was recorded is displayed as an additional attribute as well as whether it was matched or not. As for the TempMap, the bars of each stack correspond to the primary values. For each category, we map the number of events classified as belonging to that category or to one of its subordinate categories to the width of the corresponding bar. The bar height depends on the level of the category in the taxonomy: the deeper, the smaller. The axes for the attributes are aligned parallel with the primary attribute in the left-most position.

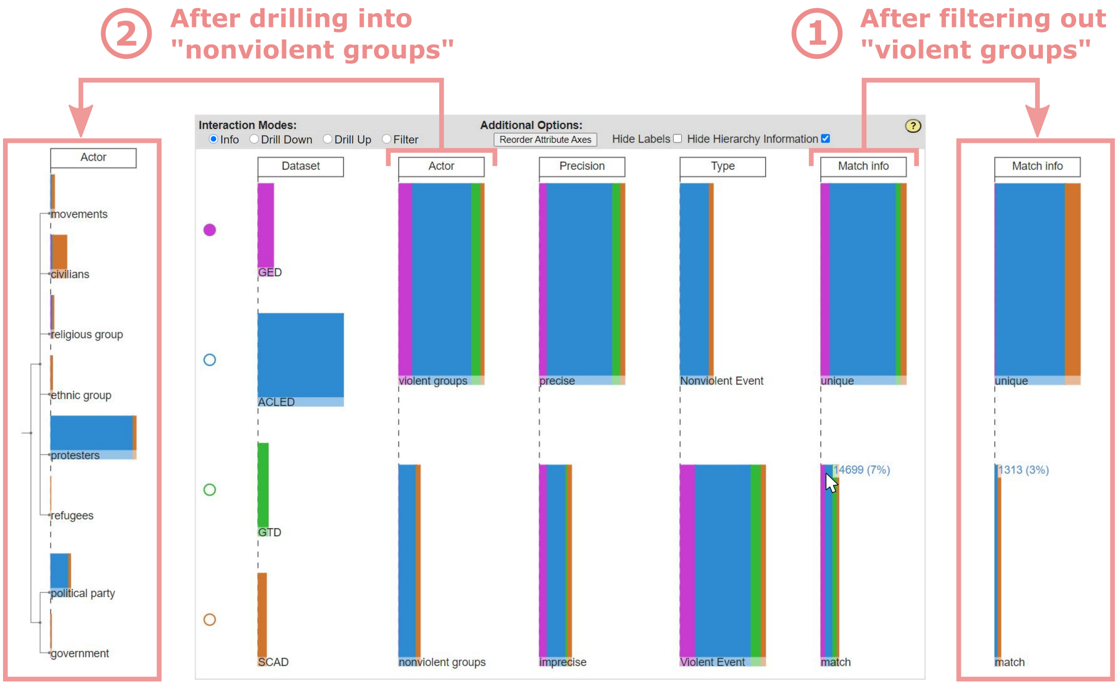


Figure 12: The EventCharts display the distribution of the events across the categorical attributes. The results of filtering (1) and subsequent drilling into the actor taxonomy (2) are shown. In addition, the hierarchical information is depicted in (2). *(Figure adapted from Mayer et al. [May+21], John Wiley and Sons. Used with permission under License Number [5891970309247].)*

Interaction. The baseline for the bar stacks can be adjusted to improve their comparability. Additionally, events can be filtered based on their attribute values (**T3**). The filtering is linked with the events displayed in the TempMap. The hierarchical aspect of the data is revealed when *drilling* into the taxonomies. The analysts can either *Drill Down* or *Drill Up*. The *Drill Down* mode works as follows:

- When hovering over a bar stack in *Drill Down* mode, the bar stacks corresponding to the underlying child categories one level deeper in the taxonomy are displayed as a preview.
- When clicking, the child categories replace the parent category. At that, all bar heights in the corresponding attribute axis are updated such that the bar heights

for categories from a certain level are half of the bar height of the categories from one level higher.

This way, the hierarchy of the displayed section of the taxonomy can be perceived by the analysts. Additionally, it can be visualized as explicit nodes and links by toggling "Hide Hierarchy Information." Once the analysts have drilled into a taxonomy, they can drill up again as follows:

- In *Drill Up* mode, the analysts can hover over bar stacks to display which other categories are siblings to the corresponding category in the taxonomy by enframing them.
- By clicking, the enframed categories are removed and their parent category is displayed.

The perception conveyed by bars corresponding to categories of different levels could be misleading if the analysts interpreted the area of the bars as the counts instead of their width. As a remedy, the number of events represented by each bar can be retrieved in the *Info* mode when hovering over the bar. For instance, Figure 12 shows a mouse cursor with the hover info "14699 (7%)" next to it, meaning that the bar represents 14,699 events, which corresponds to 7% of all events.

Separability of Primary Values. The analysts can automatically rearrange the attribute axes of the barchart according to how well they allow to separate the primary values. This can be used to determine which attribute provides the best separability of matched and unique events or what the most striking differences are between, e.g., violent and non-violent events (**T2a**). To calculate how well each attribute allows to separate the primary values in the present view, the following steps are taken:

- A binary classification is assumed for each of the categories. This means, for each stack of bars corresponding to a single category, the count of the largest bar is considered as the number of events that can be correctly classified for that category.
- This way, for each displayed category of an attribute, the maximum count of a single bar is determined and all the maximum counts are summed up.
- Finally, the sum is divided by the total number of events that are currently displayed to yield the attribute's *separability score*.

The scores of the attributes are used to determine the new axis order, arranged from highest score to lowest. Using only the present categories of each taxonomy to calculate the score allows the analysts to adjust the granularity for the calculation by drilling down or up.

Alternative Techniques. For the given task, more space-filling hierarchical visualizations such as treemaps, icicle plots, or parallel sets could be used. However, it showed that the analysts do not need to see all categories at once. Hence, our solution is more suitable as it displays the hierarchical information without overloading the screen like space-filling techniques can tend to do. Moreover, it makes it easier to compare categories from different levels with each other.

3.5.2.4 *MatchTree*

The *MatchTree* uses a radial tree layout to visualize the distribution of the matches across the categorical attributes, see Figure 13. All categories can be inspected at the same time. It is available when analyzing "Matches Only." Besides the actor, the event type, and the precision, additional information is displayed. It covers the number of events participating at each match (*Size*), the matching score (*Match score*) discretized into four equally-sized bins, and the data sets that participate in each match (*Dataset*).

Glyphs. Each category is represented by a circular glyph indicating how many matches were identified in that category, see Figure 14. The hierarchy of the taxonomies is conveyed by connecting the glyphs of sub-/superordinate categories with lines. The glyphs map the number of matches to both the fill color and the length of their surrounding arc. The encoding via the color channel allows quick identification of categories with a high number of matches. Combined with the more precise encoding via the arc length and displaying the exact count when hovering over a node, categories of interest can be identified easily (**T2b**). The maximum domain value of the color scale and the arc length scale corresponds to the maximum match count across all nodes. When hovering over a node, a dashed circle spanning all the trees is displayed, indicating its level for better comparison with other nodes, see Figure 13. Moreover, the sub-tree induced by the hovered node is highlighted.

Filtering. The filtering is designed to be consistent with the EventCharts to strengthen the analysts' mental model. It can be performed by (de-)selecting the desired categories in the "Filter" mode. In synchrony with the TempMap, this filters out all matches identified in any of the deselected categories. If the count of a node changes due to filtering, its arc is split up into colored sections to express the difference to the previous state, see Figure 14. This supports the user to track the changes across all nodes (**T3**). Being able to track the changes can yield relevant insights, e.g., when filtering out the matches with low matching scores. As the glyphs reveal to which glyphs a lot of the corresponding "bad-quality matches" were mapped, the analysts can identify to which data sets these bad matches primarily belong, i.e., which data sets are difficult to match with the others. Similarly, the analysts can also identify what types of events are more difficult to match.

If the count of a node increases after filtering, the corresponding arc section, the **gain arc**, is colored in dark green, see Figure 14. If the count decreases, the lost section of the arc, the **loss arc**, is colored in bright red. The remaining arc section, the **neutral arc**, is colored in dark grey. Since the **gain arc** is part of the current count while the

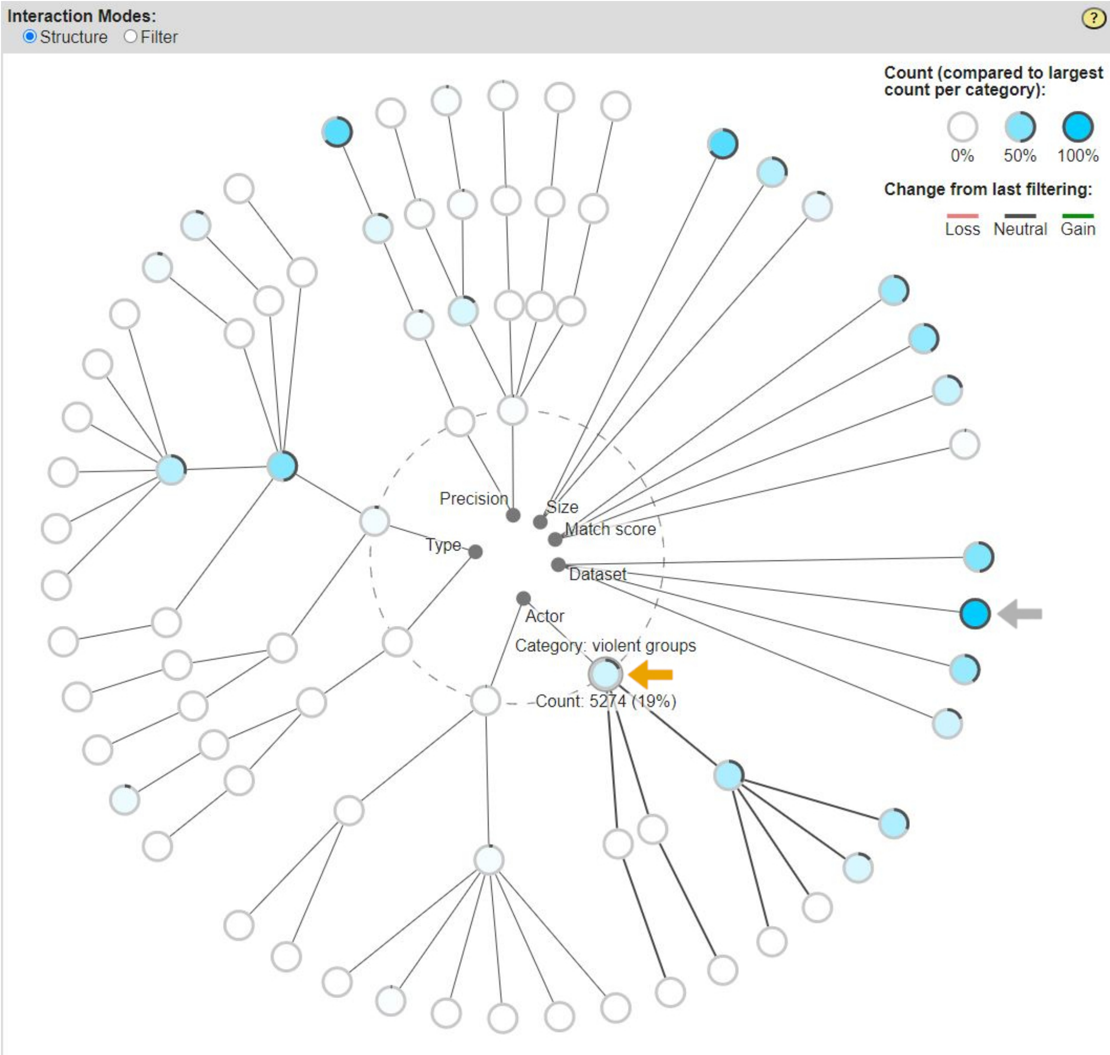


Figure 13: The distribution of the matches across the categories. When hovering over a glyph (orange arrow), a dashed circle indicates the level of the hovered node, and the induced subtree is highlighted. The color and the arc around each glyph represent how many matches were identified for the underlying category, with the corresponding scales being relative to the largest count across all categories. Accordingly, an entirely blue glyph with a full circle arc around it (gray arrow) does not reflect that all events from this category were matched, but that this category had the most matches out of all categories. The fill colors and arc lengths of all glyphs are in relation to the match count of this category. Here, the hovered glyph encodes that 5,274 matches were identified for "violent groups," which equals 19% of the largest match count in the view. *(Figure adapted from Mayer et al. [May+21], John Wiley and Sons. Used with permission under License Number [5891970309247].)*

loss arc is not, the brightness of the **neutral arc** was selected to be more similar to the brightness of the **gain arc** than to the brightness of the loss arc. If the length of the loss arc would exceed the current scale (which is possible as it refers to the count from the *previous* step), it is indicated by a red dot on top of the glyph, see Figure 14.

An additional option of interaction is to collapse sub-trees of the taxonomies. All ancestors of a clicked node are then hidden, and their counts are added to the clicked node's count. That way, the visual complexity of the graph can be reduced by aggregating information.

Alternative Techniques. We selected the radial layout over non-radial options despite potential drawbacks [BW14]. It benefits from a more compact usage of space and, hence, a more comprehensive view. This way, the distance that must be covered to compare glyphs is lower than compared to when the trees are arranged in parallel. In the latter case, especially trees in the outmost positions would require more cognitive effort to be compared [BW14].

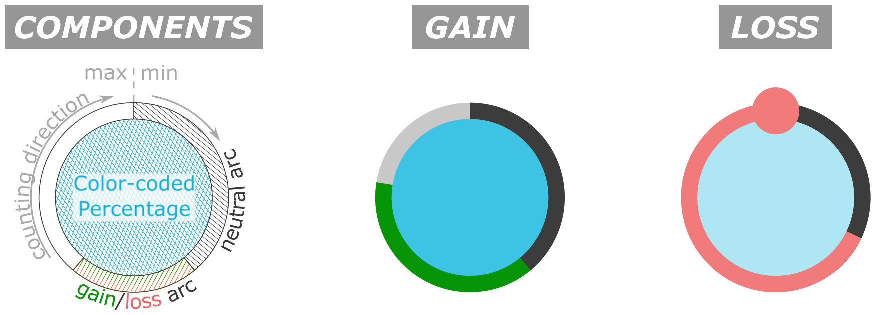


Figure 14: The sketch on the left depicts the components of a node glyph. The **green** and red arc indicate the number of matches **gained**/lost in comparison to the previous filtering state. A dot at the end of the loss arc shows if the scale is exceeded. *(Figure adapted from Mayer et al. [May+21], John Wiley and Sons. Used with permission under License Number [5891970309247].)*

3.6 EVALUATION

To show the usefulness of *VEHICLE*, we present two case studies and an evaluation conducted with conflict researchers.

3.6.1 *Case Study 1: Validating the Outcome*

This case study is based on findings we made in exploratory sessions with our collaborating expert throughout the development. It exemplifies findings of users interested in validating the algorithm outcome. In the sessions, the dissimilarities in the ParaMultiples between the different outcomes seemed quite low, except for the jump from 0 to 5km (**T1**). This is reflected both in the small dissimilarity bars and in the relatively similar histogram distributions, see Figure 8.

To investigate this behavior, the analyst inspects the MatchTree for different parameter combinations. The analyst finds that, often, a high number of matches occurs in only a small set of the nodes and that this set barely changes when varying the parameter combinations (**T2b**). The categories in this set are mainly related to violent events and events with high geographic coding precision, see Figure 13.

To examine the categories which barely have any matches, e.g., those with "nonviolent groups" (NVG) as actors, the analyst switches to the EventCharts with "Dataset" as the primary attribute (**T2a**). The bars on the "actor" axis reflect that the majority of events have "violent groups" (VG) as actors, giving the first reason for the lower number of matches for NVG. After filtering out VG, the percentage of matched events reduces substantially, see Figure 12 (1). This indicates that there must be a reason for the lack of matches in NVG other than the limited number of events compared to VG. One of them is found when *drilling down* into the actor attribute axis. This reveals that, on the deepest levels, the categories are mostly populated by events from only a single data set, see Figure 12 (2). Hence, at least at these lower levels, barely any matches can occur. Even more relevant insights are gained when looking over to the TempMap. By hovering over different countries and temporal bins, it quickly shows that especially the spatial distribution of events from different data sets varies quite a lot, see Figure 9 (right). Closer inspection by zooming into certain areas confirms this impression.

This yields two insights. Firstly, the actor taxonomy might be too fine-grained, preventing matches from being identified on deep levels of the NVG subtree. Secondly, the data sets seem to have quite different scopes for collecting events performed by NVG. In addition, this behavior, which also occurs similarly for other categories with a low number of matches, explains the high similarity between different algorithm outcomes observed in the ParaMultiples. This is the case because the reasons listed above hold for most parameter combinations.

Overall, a deeper understanding of the underlying workings of the matching algorithm could be gained than it was possible before.

3.6.2 *Case Study 2: Exporting a Subset of Interest*

This case study is adapted from the personalized evaluation sessions. The analyst wants to determine, inspect, and export a subset of interest from the integrated data, in this case, protest events in Burkina Faso (**T3**). In the first step, they refer to the ParaMultiples, see Figure 8. The color distribution of the histograms clearly shows that the number of matches increases with loosening spatial and temporal constraints, which is considered reasonable. In addition, the analyst compares the changes between adjacent histograms using the dissimilarity bars and the histogram fill colors. The analyst identifies a clear jump between $\Delta s = 0\text{km}$ and $\Delta s = 5\text{km}$ and confirms its appropriateness. Running the algorithm with $\Delta s = 0\mathrm{km}$, i.e., requiring the matched events to be recorded in exactly the same location, is the constraint with the largest impact on the number of matches, by far.

As the overall number of found matches seems reasonable, the analyst proceeds to select the most appropriate parameter combination (**T1**). To do so, they consider several aspects. Matches of high quality, conveyed through the match score distribution, make the results more trustworthy. As the overall quality reduces with loosening constraints, this speaks for stricter constraints. In contrast, the number of matches increases with loosening constraints. The higher the number of matches, the lower the chances of falsely covering an original incident twice in the final data set. This speaks for less strict constraints. However, having more matches also increases the chances of falsely matching an event that should actually be unique, excluding it from further analysis. To find a trade-off between these aspects, domain knowledge helps in combination with the examination of the match score distributions. Well-shaped distributions without strong peaks are considered more reasonable.

With $\Delta s = 50\mathrm{km}$ and $\Delta t = 2\mathrm{d}$, the analyst continues with "All Events" as the scope and "Match info" as the primary attribute, meaning that the two primary colors used across the subsequent views encode whether the corresponding events were matched or not. The analyst inspects the TempMap and filters for the country Burkina Faso, see Figure 6. They gain an impression of how the matches are distributed by using temporal highlighting (**T2a**). To inspect spatial clusters around the main capital and temporal clusters at specific times, they zoom in further and reset to the overall country view when done.

To select only events of type "protest," the analyst uses the EventCharts to drill down into the event type taxonomy and filter for protests (**T3**). They drill into other categories to better understand the distribution of the matches and use the automatic axis reorder function to see that events are best matched if they are encoded with high geographic precision, see Figure 6 (right). Using the TempMap, they also find that, for the most part, the matches were identified around the capital, increasingly since 2011, see Figure 6 (right). This means that the events recorded by the different data sets began to overlap more in the recent years and in closer proximity to the capital, which also seems reasonable.

Finally, to see how the matches of the selected subset are distributed, the analyst opens the MatchTree (**T2b**). They find that the matches are of high quality according to the match scores, see Figure 6. Satisfied with the validity of the matching outcome and the selected subset, the analyst goes back to view "All Events" and exports the integrated data.

3.6.3 *Pair Analytics Sessions*

In individual pair analytics sessions [AH+11] that lasted between 50 and 90 minutes, a total of five conflict researchers (excluding our collaboration partner) used *VEHICLE*. In the remainder of this chapter, we refer to the participants of the evaluation simply as the *participants*. With our remote assistance, they analyzed the data described in Section 3.3. We video-recorded the sessions and analyzed them afterwards. Each session started by collecting some background information about the participants. They had been working in conflict research for 3/6/9/13/19 years, had at least some programming experience with languages like *R*, and had a background in statistics. Moreover, the participants were experienced with static data visualizations but not with interactive visualizations. Two of them identified as female and three as male.

After collecting the background information, we explained the different components of *VEHICLE* to the participants. During the introduction, they already interacted with the application, and we discussed the first insights. Afterwards, they explored the application freely, if necessary with our assistance, while exchanging insights and impressions with us. Eventually, they were asked to fill out a short questionnaire. For each of *VEHICLE*'s components, the questionnaire asked whether the participants found the visualizations meaningful and comprehensible, and whether they found the interaction suitable and satisfactory. Moreover, we included questions to ask whether they found the linking of the components meaningful, whether they could export a subset of interest, and whether they would like to use the application in the future. The questions could be answered using 5-point Likert scales. We structure the results of our evaluation according to *VEHICLE*'s components.

ParaMultiples. The participants found the view "intuitive" and "useful for robustness checks and to see where the biggest change happens." They liked that both relative and absolute information is displayed and could easily assess the influence of the parameters. One participant would have liked to extend the range of Δt and to have hover information to better compare the dissimilarity counts.

TempMap. The participants found the view "useful" and "straightforward," with one even expressing that it "should be industry standard." They liked "that you can manually type the date," but one participant would have wished to also adjust the temporal bin sizes and to have a less abstract map, e.g., with terrain or cities. Another participant would have liked to see the count of individual temporal bars in the histograms when hovering over them. In addition, some participants initially experienced difficulties using the temporal brush, but were able to adapt after some time.

EventCharts. The participants found the view "handy" and "useful." They liked the drilling functionality and that the data in the view was synchronized with the exported data. One participant had issues with grasping what the axis reorder function did and stated that the view "should not be more complex" as "you have to think a little about how to interpret the information." Another participant initially struggled with processing what happened when we changed the primary attribute. The quantitative feedback reflects these difficulties, see Figure 15.

MatchTree. The participants found the view "useful" and "really interesting." They especially liked the additional information about size, match score, and data set "because it is really difficult to look at otherwise." They used this combined with filtering and inspecting the arcs. About the complexity, one participant stated: "You first have to get used to the depiction [...] but I find it cool."

Overall. In the free investigation, the participants showed excitement and curiosity for the various ways to explore the data. With one participant, we even lost track of time to which they stated "it shows how much there is you can do." The participants found the application "super useful" but said that it was not for casual users. One participant stated that it "provides a much deeper look into [the integrated data]" than possible before. This also shows in the quantitative feedback, see Figure 15. Still, one of the participants would have liked to "go into the original event text to feel psychologically closer to what the system does." At the start of the sessions, the participants required quite some assistance, but towards the end, they could do almost all investigations on their own. They expressed confidence to use the application on their own and excitement to do so in the future.

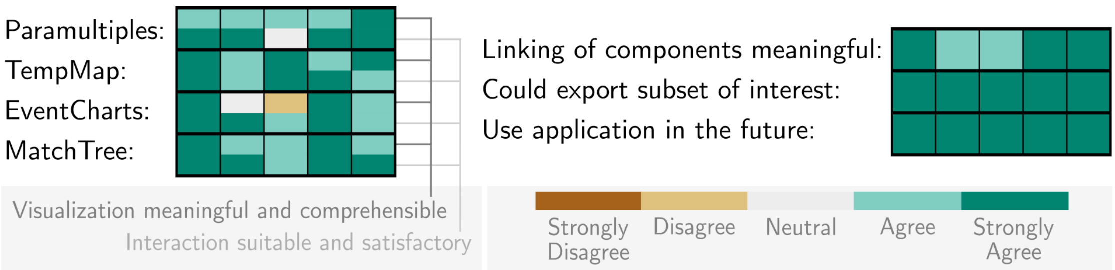


Figure 15: Each column of the two tables corresponds to the quantitative feedback of one participant after the pair analytics session. *(Figure adapted from Mayer et al. [May+21], John Wiley and Sons. Used with permission under License Number [5891970309247].)*

3.7 DISCUSSION AND FUTURE WORK

Potential improvements. Our evaluation showed that the domain experts could use *VEHICLE* to analyze the integration of conflict data and enjoyed doing so despite some drawbacks. Accordingly, they wished for more personalized adjustments. This includes viewing additional information and adjusting the temporal bin sizes in the TempMap. While these issues can be solved through smaller adjustments, a more difficult drawback to deal with is the application's complexity.

To make *VEHICLE* more comprehensible, particularly the EventCharts view, several actions should be taken in the future. To improve legibility, the same height for all category bar stacks could be used. By doing so, the different axes would have a more balanced aspect ratio and, thus, allow for a more prominent display of the hierarchy based on node-links or stacked rectangles. This way, we could reduce the confusion introduced by the differing bar heights and could potentially even rearrange the axes to save space. In addition, the axis of the primary attribute could be set apart from the other attributes to better convey its relevance and, thus, the meaning of the view. This could be done using, e.g., spacing, a frame, or a dedicated background. Highlighting the primary attribute could also make the "Reorder Attribute Axes" interaction feel more intuitive.

Overall, the application would benefit from more *guidance* [Cen+17]. While we have included means for supporting the analysts, like additional textual explanations for each view upon request, more elaborate techniques would be beneficial. Options include to orient the users via visual cues, to direct them to unusual patterns, and to suggest potential further analysis steps for them to take.

Generalizability. To extend our approach to other application areas, the corresponding data need to have the following properties. The event recordings, which come from different sources, represent their original incidents with a loss of information. Otherwise, if no information was lost in the recording step, the events could be matched directly in a non-probabilistic way. Ideally, the data have a spatio-temporal context, and they can be classified in at least two different hierarchical structures/taxonomies. The depth and number of the taxonomies may vary to a certain degree to still be displayable in the EventCharts and the MatchTree but should not be too high to prevent visual clutter. If necessary, the views could be rescaled.

To integrate and analyze such data using *VEHICLE*, analysts need to take the following steps. They need to create suitable taxonomies and classify the original data accordingly. In addition, they need to determine reasonable ranges for the parameters Δs and Δt as, naturally, different ranges might be applicable depending on how densely populated the regions under consideration are. For the ParaMultiples, this adjustment should be relatively feasible. Despite changes to the maximum values of the ranges, only a limited number of grid points for each parameter would be necessary. The underlying assumption is that larger ranges indicate higher imprecision of the measured distances, meaning their granularity can be coarser. In a case where this does not hold, panning and zooming functionalities should be provided, or the ParaMultiples would need to be changed more substantially.

After the parameter selection, a matching algorithm like *MELTT* [Don+19] needs to be run on the given data. Techniques like semantic matching [GS03] could improve the integration results, e.g., in the case where a protest is recorded as a violent event by one data set and as a nonviolent event by another one. With semantic matching, the underlying meanings of the different categories could be considered rather than just the category names. This change would require a larger adjustment of the matching algorithm and the MatchTree to indicate which matches were identified this way.

After following the steps above, the data could be analyzed in *VEHICLE* as described across the previous sections. An example of such an approach could be imagined in social media analysis when trying to collect a set of original incidents from social media posts. The different data sources would then be different channels and the taxonomies could be either manually created by domain experts or the results of different hierarchical clusterings applied on the data [Fen+15; ISB14].

3.8 CONCLUSION

We introduced *VEHICLE*, a web-based application to validate and explore the hierarchical integration of conflict event data. Such investigations were only possible in a rudimentary way so far, calling the validity of insights derived from the integrated data into question. Throughout the development of *VEHICLE*, we identified associated domain problems, characterized the underlying data, and derived a task and workflow abstraction. To accomplish this, we worked together with a conflict research expert to design and develop the application in an iterative process. *VEHICLE* consists of multiple linked views that allow analyzing the data in two main ways. The first way considers all events from the different integrated data sets, while the second way considers only the *matched* events, i.e., those events that were identified to represent the same original incident. For the identification, we adopted the matching algorithm *MELTT* [Don+19]. To facilitate the analysis, we relied mainly on direct interaction techniques and employed radial layouts in two of the views. In addition, we provided a view allowing analysts to investigate the influence of the input parameters required for the matching algorithm on its outcome.

In two case studies and a pair analytics evaluation, we demonstrated that with those design decisions, *VEHICLE* allows conflict researchers to gain new insights about the integration process and assess its validity. It also allows them to confirm existing hypotheses, to explore subsets of events, and to export subsets of interest for further analysis. At the same time, the results showed that, due to its complexity, *VEHICLE* has to be considered an expert application. In addition, the evaluation provided us with directions on how to extend the application in the future.

3.9 ADDENDUM

The study showed that interactive visualization can benefit the field of conflict research, while it is still barely applied in this domain. Seeing more such applications in the future would be great. However, in the evaluation, we also observed that gaining an initial understanding of *VEHICLE* was challenging for the participants. As the application has multiple linked views that can be explored in unison, the insights that can be theoretically gained are numerous, particularly when applying filters across multiple views. Yet, this intricacy also hampers the process of learning the application. We tried facilitating the process by providing help texts. However, these texts were hidden behind an additional button click, which might be one of the reasons why they were not used a lot.

In a separate project in the domain of medicine, we tried putting such explanatory aids more clearly into view [May+23a]. In the project, we developed a web-based application [Ive] consisting of multiple linked interactive visualizations to explore multi-faceted data about surgical workflows. Above each visualization, we placed a hideable text block briefly summarizing what is displayed in the view. Figure 16 shows a corresponding example for two views of the application, which consists of seven views in total. In addition, the application contains a sidebar for additional interaction options, like zooming, brushing, and filtering. When hovering over the corresponding segments of the sidebar, additional information is displayed about how to use the interactions. However, even with these more explicit means of explanation, an evaluation of the application showed that the participants required some time to familiarize themselves with the system and that they did not feel too confident using it [May+23a]. This also made clear that such issues persist across different application domains.

Therefore, we wanted to investigate how to better support an expert audience that has limited experience regarding interactive visualization. There already exist approaches like *guidance*, where visual analytics systems are augmented with means to guide users to take meaningful actions in their analytical process [Cen+17]. However, we wanted to go further in a direction that we have not seen research deal with a lot, so far. Namely, we wanted to investigate how techniques from *visual storytelling* can be combined with interactive visualization to target expert audiences. While such approaches are quite common for broad audiences, domain experts are rarely the target of corresponding applications.

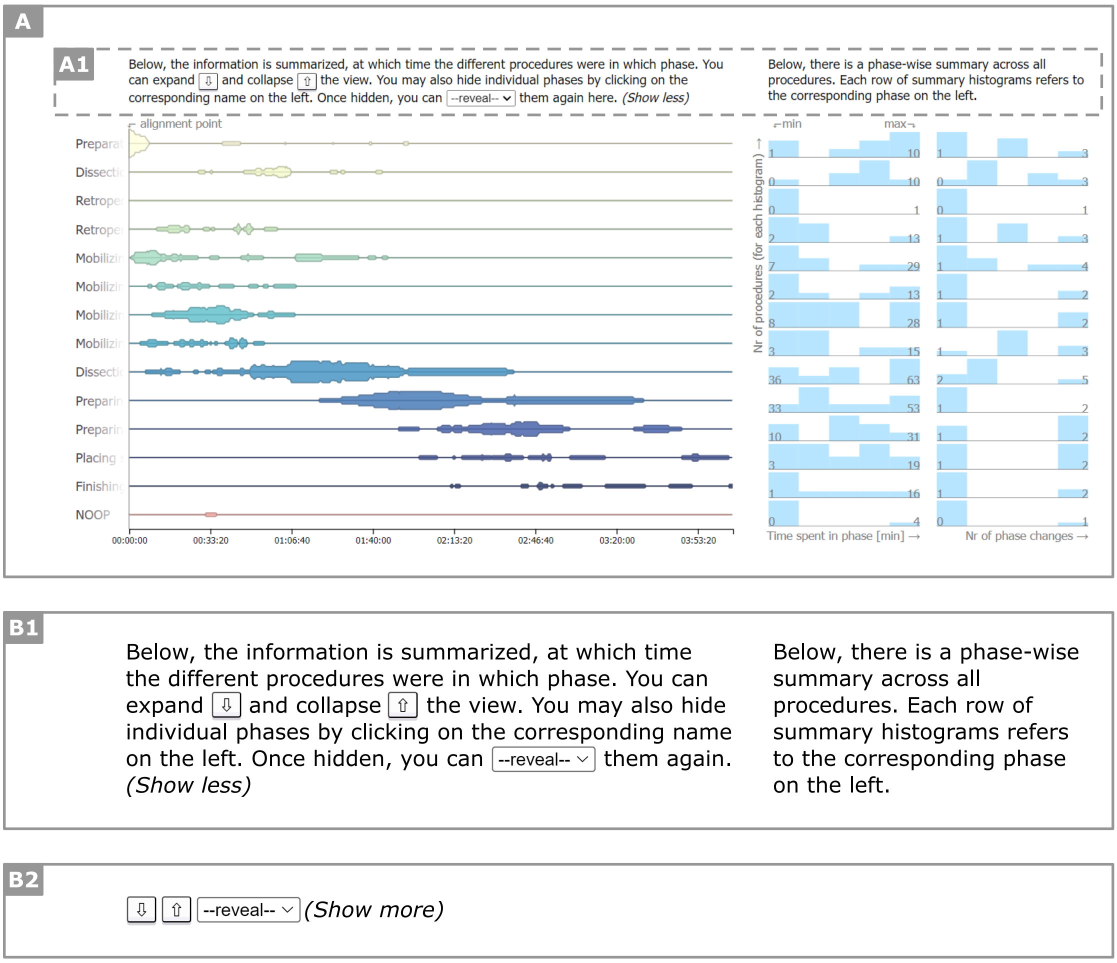


Figure 16: (A) depicts a screenshot of an application developed for surgical workflow analysis [May+23a]. It contains two main views: a column of area charts on the left as well as two columns of histograms on the right. Above both views, there is a corresponding textual explanation (framed by (A1)). For better readability, (B1) shows an upscaled version of the texts from (A1). In addition, (B2) displays a compressed version of the text from (B1) that replaces the original text above the views when clicking on *(Show less)*. This way, analysts can remove the explanatory texts and keep only the relevant controls once they are familiar with the views.

4

EXPLORATORY EXPLANATION FOR EXPERTS

Science communication aims at making key research insights accessible to the broad public. If explanatory and exploratory visualization techniques are combined to do so, the approach is also referred to as exploranation*. In this context, the audience is usually not required to have domain expertise. However, we show that exploranation can not only support the communication between researchers and a broad audience, but also between researchers directly.*

With the goal of communicating an existing method for conducting causal inference on spatio-temporal conflict event data, we investigated how to perform exploranation for experts, i.e., expert exploranation*. Based on application scenarios of the inference method, we developed three versions of an interactive visual story to explain the method to conflict researchers. We abstracted the corresponding design process and evaluated the stories both with experts who were unfamiliar with the explained method and with experts who were already familiar with it.*

The positive and extensive feedback from the evaluation shows that expert exploranation is a promising direction for visual storytelling, as it can help to improve scientific outreach, methodological understanding, and accessibility among researchers.

This chapter is based on the following contribution [May+24], licensed under CC BY 4.0:
B. Mayer, K. Donnay, K. Lawonn, B. Preim, and M. Meuschke. "Expert exploranation for communicating scientific methods - A case study in conflict research." In: *Computers & Graphics* 120 (2024), p. 103937. DOI: https://doi.org/10.1016/j.cag.2024.103937.

4.1 INTRODUCTION

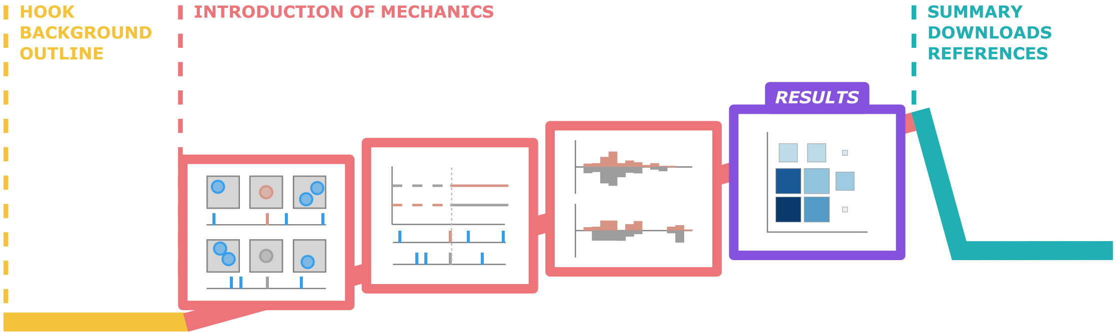


Figure 17: The illustration shows the arc of our story used to explain a method for performing causal inference to conflict researchers. The four core scenes of the story are depicted through schematic representations of the key visualizations used in the scenes.

The combination of explanation and exploration, referred to as *exploranation* [YLT18], is commonly used to convey information from domain experts to a broad audience without domain expertise [YLT18]. However, we argue that it is also beneficial for supporting the communication between experts from the same domain.

The scenario we focus on is the following: Researchers have developed a complex method for which it is not straightforward to assess how and under which circumstances it can be applied. It could support the work of fellow researchers, but they only have limited capacity to delve into the method to determine whether it would be beneficial to them. In this scenario, exploranation can facilitate the fellow researchers to gain insights into the workings of the method to assess its applicability to their own research problems.

So far, not a lot of research has been performed in this direction, while there exist several questions regarding such exploranation for domain experts, i.e., *expert exploranation* (ExEx): Are researchers interested in ExEx at all? If yes, how would they like to use it? And which potential directions do they see for it?

The process and challenges of creating ExEx stories need to be investigated, including what kind of information is necessary for other experts to understand the content in sufficient detail, how to abstract this information, and how to order it in an understandable sequence. Moreover, it is of interest how to perform effective storyboarding for producing suitable visualizations and texts using appropriate jargon.

To investigate the process and potential of ExEx in more detail, and in which aspects it differs from performing exploranation for broad audiences, we collaborated with a conflict researcher to create a story explaining one of their developed methods [SD14]. The method's purpose is to perform counterfactual inference between spatio-temporal conflict events. For instance, it was used to test which kinds of aid projects are suitable to stabilize war-torn communities, and which, instead, tend to increase anti-government activities [KS18]. We considered three such example use cases of the method. For each example case, we automatically generated an adapted version of the same original story. Accordingly, we produced three versions of the story, each explaining the causal inference method in a different application scenario [Mat]. By abstracting the corresponding design process and evaluating the stories, we aim at increasing the insights into how ExEx can be performed in the following way:

- We present the design process of an interactive story to explain a conflict research method to domain experts.
- We present the results of an evaluation of the story. The evaluation was performed with two complementary user groups – one with in-depth knowledge of the explained method and one with barely any prior knowledge.
- We summarize our learnings regarding how ExEx can be performed and which future research is necessary.

The three versions of the story can be accessed from this website [Mat].

4.2 RELATED WORK

Science communication. To explain research insights to a target audience, first, the audience needs to be understood [Mer21]. We focus on conflict researchers, ranging from graduate students to senior scientists. While their numeracy can be expected to be increased, limiting the cognitive effort is still advisable [Ins14]. As they have domain expertise, their mental models can be expected to better support a correct understanding of conveyed concepts [Nat17].

For communicating scientific information, Meredith advocates using informative and engaging visuals to improve the comprehension and engagement of the audience [Mer21]. There are many examples for communicating science to broad audiences by making use of visualization, e.g., for astrophysics [Lan+21], climate communication [WSM19], and medical communication [Meu+22]. However, for addressing expert audiences, the literature is more sparse. The most common means for communicating within the research community are traditionally abstracts, posters, oral presentations, peer-reviewed research papers, and, more recently, open access formats [UKW17]. In these formats, not a lot of focus is put on using effective and engaging ways of communication.

Story creation process. For the general steps of our story creation process, we primarily referred to five works [Lee+15; CA+20; Zha+22; SH14; Ami+15]. Lee at al. identified steps ranging from the initial ideation, over the scripting and editing, to the final presentation to the audience [Lee+15]. Contrary to our work, they did not focus on a specific audience. More closely related in that regard is the work of Cortes Arevalo et al., presenting key steps to take when creating stories for the communication between researchers and practitioners [CA+20]. Zhang et al. focused on cognitive and communication theories in their framework for data-driven visual storytelling [Zha+22]. They adapted their design process from Satyanarayan and Heer [SH14]. Both groups of authors [Zha+22; SH14], as well as Lee et al. [Lee+15], assume the design process of a visual data story to begin with the exploration of a data set in search for interesting facts to communicate.

In contrast, in our scenario, the key goal of the story is already known, i.e., explaining a scientific method. In this case, the beginning of the design process is less a task of exploring data, but more a task of identifying which core mechanics of the method should be explained in what detail. These considerations are part of the larger process of storyboarding, which was analyzed by Amini et al., revealing the non-linear and iterative order of the steps involved [Ami+15]. In our work, we present the similarities and differences between our expert-focused story creation process and the five works listed above.

Story structure and layout. Our story follows the traditional Freytag's Pyramid story arc, for which Yang et al. analyzed how it can be applied to data stories [Yan+21]. Regarding the overall layout, our story would be classified as a "dynamic slideshow" [Rot21] or "interactive slideshow" [SH10], as it is navigated discretely through a set of slides and provides interactions at key points throughout the narrative [SH10]. Hullman et al. investigated how to sequence slideshow-style presentations for improving their comprehensibility and memorability [Hul+13].

Design spaces and automation. For advice regarding concrete storytelling techniques, we referred to the design spaces by Roth [Rot21], Stolper et al. [Sto+16], and Mayr and Windhager [MW18]. Such design spaces often lead to the development of authoring tools, based on the identified dimensions. However, in their survey of authoring tools from 2023, Chen et al. came to the conclusion that tools to create complex slideshows are still lacking [Che+23]. Overall, there is still a gap between fully automated story generation and manually created stories [Sun+23]. To properly automate a process, it should first be thoroughly understood. Accordingly, as there is not much research dedicated to expert exploranation, it is necessary to understand how it can be performed in general before trying to automate it.

Structuring visualization and text. Previous works inform on how visualizations and text can be integrated effectively [ZOM19], some even focusing on spatio-temporal stories [LCB21]. Taken together with McKenna et al.'s narrative *flow factors*, including aspects like how the reader can navigate through the story [McK+17], these works provide a foundation for structuring and improving the story reading experience.

Exploranation. Under the term *exploranation*, Ynnerman et al. introduced design principles for how to combine interactive exploration with explanation [YLT18]. In applications of the exploranation approach, it was claimed that corresponding systems can be used for lay audiences as well as expert audiences, only by adjusting the underlying story [Boc+18; Bro+23]. Similarly, we also show that generating multiple versions of the same base story with differing content and level of detail can be beneficial. However, in contrast to the two works [Boc+18; Bro+23], we argue that dedicated investigations focusing explicitly on experts are still necessary. In the two works, the focus was put on contextualizing and exploring scientific data sets while using the same underlying system, *OpenSpace* [Boc+18]. However, instead of patterns in a data set, we aim at explaining the workings of a scientific method, for which several concepts need to be introduced, building on each other. This requires customized sequences of explanations and corresponding views. Accordingly, the requirements for the story differ from those of a lay audience, which would be interested in the results of the method rather than its individual mechanics.

Another related approach are *interactive lecture demonstrations*, where lectures are enriched by allowing students to get more actively involved with the presented information [ST04]. However, this approach is typically applied in synchronous settings with a live presenter, whereas our stories should work in a standalone, asynchronous setting.

In summary, we use insights from science communication to create a visual data story for communicating the workings of a scientific method between researchers. For the design, we build on a range of works providing guidelines for the creation of slideshow-based, interactively explorable stories. With our work, we do not intend to replace traditional means for communicating scientific results between researchers [UKW17], but to enrich the existing set of options with a more accessible and immediate way to learn about new scientific methods.

4.3 THE COMMUNICATED METHOD

We give a brief outline of the method that we introduce in our expert exploranation, called *matched wake analysis* (MWA) [SD14]. To do so, we follow the example of how MWA was applied in an analysis performed in a separate study [KS18]. Below, we refer to this example application as *EA*. The EA study focuses on Afghanistan to analyze whether aid projects that exclude parts of a given community can lead to an increase in anti-government (i.e., *insurgent*) activities in that area. An example of an aid project excluding part of a community would be funding a selected private group of residents from a larger community. An example of an insurgent activity would be a violent attack by a terrorist group against the military or civilians supporting the government. The goal of the study was to not only identify the correlation between aid projects and insurgent violence, but rather to determine whether the exclusionary application of aid *causes* an increase in insurgent attacks. To answer this question, the authors made use of MWA to perform a corresponding causal inference.

MWA relies on three types of events: *treatment* events, *dependent* events, and *control* events, which we explain below (see also Table 1 for an overview). The goal is to determine the effect that the occurrence of treatment events has on the occurrence of dependent events. To show actual causality, this effect is compared to the effect that the occurrence of corresponding control events has on the occurrence of dependent events. In the example of EA, *aid projects excluding parts of the community* are the treatment events and *insurgent activities* are the dependent events. The control events need to be selected such that they are reasonably comparable to the treatment events but still structurally different. In EA, the control events are *more inclusive aid projects benefiting the whole community*, like improving a village's infrastructure. Taken together, the treatment and control events represent the *intervention* events.

Event type	**Overarching term**	**Purpose**	**Equivalent in the example application (EA)**
Treatment	Intervention	Test effect of treatment events on dependent events	Aid projects excluding parts of the community
Control	Intervention	Control group to compare effect of treatment events against	More inclusive aid projects benefiting the whole community
Dependent	—	Events based on which the impact of the intervention events is measured	Insurgent activities

Table 1: A summary of the terminology we use for the different event types required for the causal inference performed in MWA.

Simply comparing the effect of the treatment and control events would not constitute a convincing causal inference, yet. Instead, it needs to be ensured that the events that are compared took place under conditions that are as similar as possible. Otherwise, confounding factors could skew the results. The conditions under which the intervention events took place are described by the so-called *matching variables*. They have to be provided by the *communicating researcher*, i.e., the expert whose application scenario of MWA is communicated.

For instance, if the treatment events took place in areas that are generally more affected by conflicts than the control events, it is also more likely for insurgent activities to occur in the same area, irrespective of the type of aid project conducted. In addition, the population density of the area in which the intervention events took place may play a role, or which ethnic groups live there. These and other conditions are collected for each intervention event and encoded as matching variables used to perform *statistical matching* [IKP12]. The matching is used to ensure that only those treatment and control events are compared in the final inference that are similar enough with respect to all provided matching variables. Accordingly, only those interventions are compared that took place under similar conditions. Figure 18 (S3) shows the visualization from our story illustrating this matching.

One special matching variable already touched on in the examples above is the *trend*. It is calculated inside MWA directly and does not have to be provided by the communicating researcher. The trend describes the dynamic of the occurrence of the dependent events prior to the occurrence of the intervention events. In the example of EA, it captures to which degree the number of insurgent activities were already increasing or decreasing prior to the start of the aid project. In the lower half of the visualization depicted in Figure 18 (S2), the trend calculation is illustrated above each timeline.

The last core mechanic of MWA deals with the question of how large the area around the intervention events should be when counting dependent events in their proximity. This refers to both the spatial and temporal distance. The observed effects of the intervention events may vary if, e.g., only five days before and after the start of an aid project are considered instead of 30 days when counting the surrounding insurgent activities. To solve this issue, MWA runs the causal analysis multiple times using different combinations of spatial and temporal distances. The results from the different analyses can then be compared. The corresponding visualization from our story is depicted in Figure 18 (S4).

The story walks the user through all these mechanics. Like MWA, the story can be applied for different data sets and scenarios, adjusting automatically. Three such examples are provided on this website [Mat]. While MWA was originally introduced in the field of conflict research, it can also be applied to comparable problems in other domains, along with our story.

4.4 DESIGNING AN EXPERT EXPLORANATION

In this section, we present our insights into how ExEx can be performed. The insights are based on our experience from designing the story to communicate MWA. Our audience ranges from graduate students to senior scientists in the field of conflict research. Their numeracy can be expected to be advanced, and as researchers with a statistical background, they can be expected to be familiar with basic visualization and interaction techniques.

4.4.1 *Goals and Challenges*

Goals. Our goals, determined with our collaborating domain expert, were to provide researchers an

- *accessible* and
- *engaging* way to

- *understand* the core workings of MWA and
- allow the users to judge whether the method is *applicable* in their own research.

Moreover, the story should be

- *adaptable* to different application scenarios of the method.

This allows the integration of the story not only along with the original method it explains but also in the context of research projects that used the method for their own research. For instance, the version of our story following the example of the aid projects (EA) can be included on the website for the EA study. There, it can allow anyone interested in the study to understand more easily how MWA was used in its context. This way, also the scientific outreach of MWA can be increased.

Challenges. ExEx comes with certain challenges that go beyond the typical challenges of exploranation for broad audiences. The jargon of domain experts is more specific than the "lay language" of a broad public [Mer21]. This makes the story designers more dependent on the experts' input when producing text for the story. Another point that requires additional input is to understand which domain-specific concepts need to be explained and, in contrast, what prior knowledge the users, who are other researchers from the field, can be expected to have. However, among experts, the prior knowledge can also vary, depending on how familiar they are with the concepts explained in the context of the method. To account for this, users should be provided with the option to skip parts that they already know.

Moreover, when using ExEx for explaining a method, interaction is not primarily used for exploring data sets, as it tends to be the case when addressing broad audiences. Rather, it is also used to better convey certain mechanics and the implications they have for the workings of the method. We provide examples for this later in Section 4.4.3.

Another point to consider is that designing visualizations and arranging them in a sequence that effectively conveys information becomes more challenging the more complex the conveyed information is. In addition, it requires more effort for visualization designers to understand a method from another research domain deeply enough to explain it to other researchers from that domain, as compared to explaining it to a broad audience on a much higher level. Lastly, if the method to be explained is available as a software package, as much access as possible to the underlying code is desirable.

Overcoming these challenges requires close collaboration with domain experts. The increased scientific outreach and benefits for the research community that such stories can have can be valuable enough for the collaborating researchers to contribute the time necessary [Bey+20].

4.4.2 *Steps of the Story Design Process*

In this section, we describe the steps we followed to design our story. We derived them by merging existing approaches [Lee+15; CA+20; Zha+22; SH14; Ami+15] to design data stories and adjusting them to the requirements of ExEx settings. The relationship between the corresponding works and our approach is shown in Table 2. The process of collaborating with our project partner was inspired by Lloyd and Dykes's analysis of approaches for human-centered geovisualization design [LD11]. A central part played a Miro [Mir] board, which we will refer to as our "storyboard." Of course, the design process was not as separated and streamlined as conveyed through the sectioned structure below, but rather required several iterations.

4.4.2.1 *Understand and abstract the method*

The first step for the designer of an ExEx story should be to familiarize themselves with the method to be explained. When designing stories for a broad audience, the corresponding process is often more open at this point, exploring data to extract meaningful pieces of information from them. In such approaches, the corresponding steps would be to *explore and analyze* [Lee+15] the data, or to perform *exploration to uncover interesting stories in data sets* [SH14], see Table 2.

Understand the method. To understand the method, we referred to the corresponding paper [SD14] and discussed uncertainties with our project partner. This also required reading and discussing additional literature on which the method builds. While the underlying paper represented the most important source of information, for certain detailed questions, it was even necessary for our domain expert to dig into their code. Sometimes they also had to adjust it to let the corresponding package [Mwa] return the necessary information, highlighting the importance of having access to the code of the communicated method.

Abstract the key mechanics into story pieces. To keep the story at a reasonable length, not all details of its working and corresponding parameters could be explained. The abstraction process was based on iterative discussions and refinements to decide which mechanics are the most relevant for the method. To collect and organize the corresponding information, we used the storyboard. The result was a set of *story pieces* [Lee+15].

4.4.2.2 *Arrange the story pieces*

In this step, the story pieces are arranged in an educationally meaningful order, similar to the steps *ordering the story pieces* [Lee+15] or *composing the information units* [Zha+22].

Determine the interdependency of the story pieces. To do so, first, the logical relationship between the story pieces needs to be determined. Which concepts should

Approach	Steps						
ExEx	**Understand and abstract the method**		**Arrange the story pieces**		**Implement the story**		**Review and iterate**
	Understand the method	Abstract the key mechanics into story pieces	Determine the interdependency of the story pieces	Arrange the pieces in an educationally meaningful order	Determine the actor constraints	Draft text and visuals	Review and iterate
Lee et al. [Lee+15]	Explore and analyze		Make logical connection	Order story pieces	Build presentation		Respond to input
Cortes Arevalo et al. [CA+20]	Prepare concept	Define parts				Draft text and visuals	Review and iterate
Zhang et al. [Zha+22]	Plan the message		Compose the information units		Map the composition into visual		*Not explicitly mentioned*
Satyanarayan and Heer [SH14]	Exploration to uncover interesting stories in data sets		Drafting to prototype ways of communicating the stories found		Production to develop the final interactive		
Amini et al. [Ami+15]	Reading and interpreting data	Selecting data	Crafting the narrative structure			Integrating strategies to engage viewers	Non-linear and iterative process

Table 2: The relation between the steps of expert exploration and other visual data story design process abstractions.

be explained first? Which mechanics are necessary for explaining other mechanics? Flowchart sketches allowed us to discuss these questions.

Arrange the pieces in an educationally meaningful order. The determined dependencies between the story pieces help to bring them into an order in which they can be understood as easily as possible. In addition, user engagement and understanding can be increased by structuring the story along a story arc. For ExEx, the Freytag's Pyramid arc for visual data stories [Yan+21] can be adapted as described in the next paragraph, and as illustrated in Figure 17. For a more detailed view of the visualizations used to explain the key mechanics of MWA, please refer to Figure 18.

Note: The color coding used in the following paragraphs corresponds to the color coding from Figure 17 and Figure 18. We adapted it from Yang et al. [Yan+21].

At the beginning of a story, the *setting* is established [Yan+21]. In ExEx, this can include an initial question to **hook** the users, e.g., in case of the EA scenario: "What is the impact of aid projects excluding parts of a community on insurgent activities in Afghanistan?" Similar to applications for broad audiences, the hook is used to generate curiosity, e.g., by showcasing what kind of scientific questions can be answered using the explained method. Also, **background** can be provided regarding the purpose of the method and an initial **outline** of how it works. In our story, we present these three aspects in an introductory pop-up window.

In the subsequent core part of the story, *rising tension* leads up to a *climax* [Yan+21]. In ExEx, this means that the key **mechanics** of the method are introduced step by step, building up to a comprehensive picture. In the *climax*, the **results** are revealed, including the answer to the initial question from the **hook**.

Lastly, in the *resolution*, the story is concluded [Yan+21]. In ExEx, this can be a brief **summary** of the method, options to **download** the results, and a list of **references**.

In addition to the overall arc across the entire story, individual sections of it may have peaks on their own. Amini et al. noted this as beneficial if multiple aspects of a topic need to be explained [Ami+15]. To describe how we implemented this, we use the following terminology: The core of our story consists of four main sections, represented by the four screenshots in Figure 18. Inspired by cinematographic language, we refer to each such section as a *scene*. Each scene consists of a discrete sequence of views, which we refer to as *shots*, see Figure 19.

In each scene, a key mechanic is explained, following the same miniature arc in all four scenes: Every scene begins with a relatively plain view. Then, information is continuously added with each shot until the view reaches its peak complexity in the final shot of the scene, as sketched in Figure 20 (D). With this *staging*, we aimed at making the explanations as easy to understand as possible [Ami+18]. We also tried to keep the story visually consistent, even from one scene to the next, to maintain *continuity* [Rot21]. Such interconnectivity can make stories easier to follow and memorize [Zha+22]. Moreover, the final shot of each scene allows the user to interact with the visualization to engage more deeply with the information conveyed in the scene, deepening their understanding. Following the approach of *exploranation* [YLT18], these *interleaved explorative microenvironments* are designed to be expressive while also constraining the interaction to limit the risk of producing views that are difficult to understand.

With this approach, we also tried to account for the cycles of engagement and disengagement described by O'Brien and Toms [OT08]. Reengagement with an application occurs when the user invests themselves in the interaction beyond a routine level.

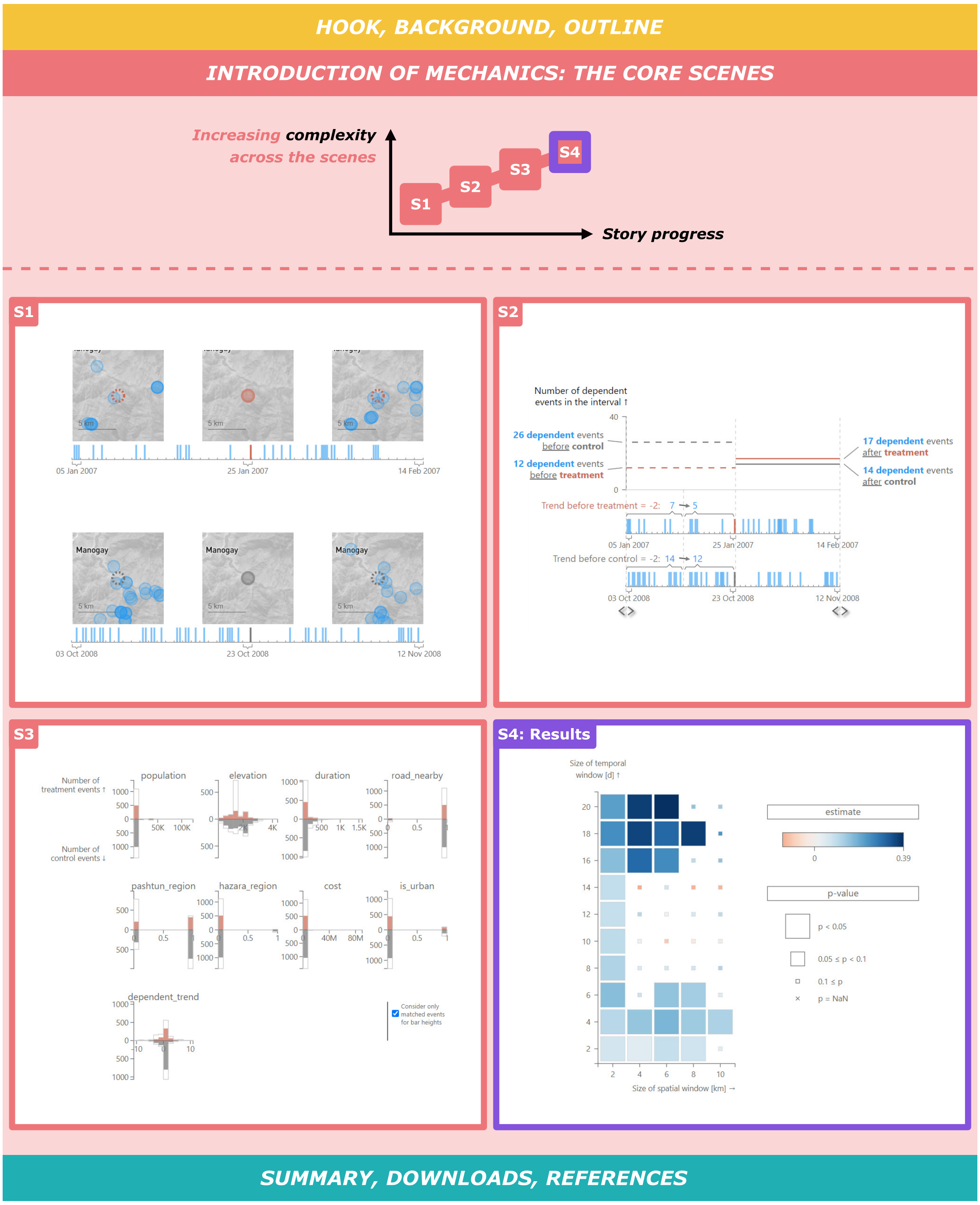


Figure 18: The sections of our story with the final shot of each core scene (S1-S4). Larger screenshots are provided in Figure 22 (S1), Figure 23 (S2), Figure 24 (S3), and Figure 25 (S4). *(Figure adapted from Mayer et al. [May+24], licensed under CC BY 4.0.)*

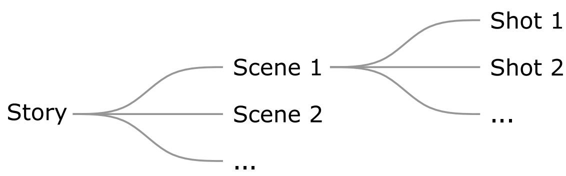


Figure 19: Decomposition of the story into scenes and shots.

Hence, the interactive shots at the end of each scene encourage the user to reengage with the content and reflect before transitioning to the next scene. For planning the arrangement of the story pieces, we used basic flowcharts in our storyboard.

4.4.2.3 *Implement the story*

When implementing a story, thorough storyboarding and prototyping can help reduce the need for time-consuming re-iterations at later stages [Hla+20]. For storyboarding, we referred to our Miro board. While it was primarily used for flowcharts in the previous steps, for this step, we used the components depicted in Figure 20. For each scene, the board contained an area for the following components: (A) A textual summary of what should be explained and visualized in the scene, in terms understandable to the domain expert. (B) A list of constraints and considerations for the scene, e.g., regarding important points to mention, potential special cases, visuals, and interaction. (C) A sketchboard area for designing and refining the different components of the visualizations, and (D) an area where the carved-out sequence of individual shots from the scene can be arranged. (E) In addition, the designers and the domain expert were able to place and reply to sticky notes wherever they needed to clarify something.

For prototypically implementing the views that were crafted in the storyboard, we used the platform Observable [Obsc] to create a collection of cell-based online notebooks. The collection is available on this website [Obsa]. As the notebooks are web-based, they can easily be shared with collaborators. This allowed us to perform fast implementations and exchange feedback early on in the design process, as advised by Hlawitschka et al. [Hla+20]. Such exchange about the possibilities on the implementation level is key for close collaboration [Bey+20]. Observable provides easy inclusion of inputs like sliders to modify various aspects of a visualization and test border cases, which we made use of in our notebooks [Obsa]. This was beneficial as our story should be applicable to various data sets. The code from the notebooks also provided a solid foundation for implementing the actual story, which we did using React.js [Rea] and D3.js [D3].

Determine the actor constraints. To make the story easier to follow, we focused on selected example events from the underlying data set when explaining the different mechanics. These example events play the main roles in the story, so we refer to them

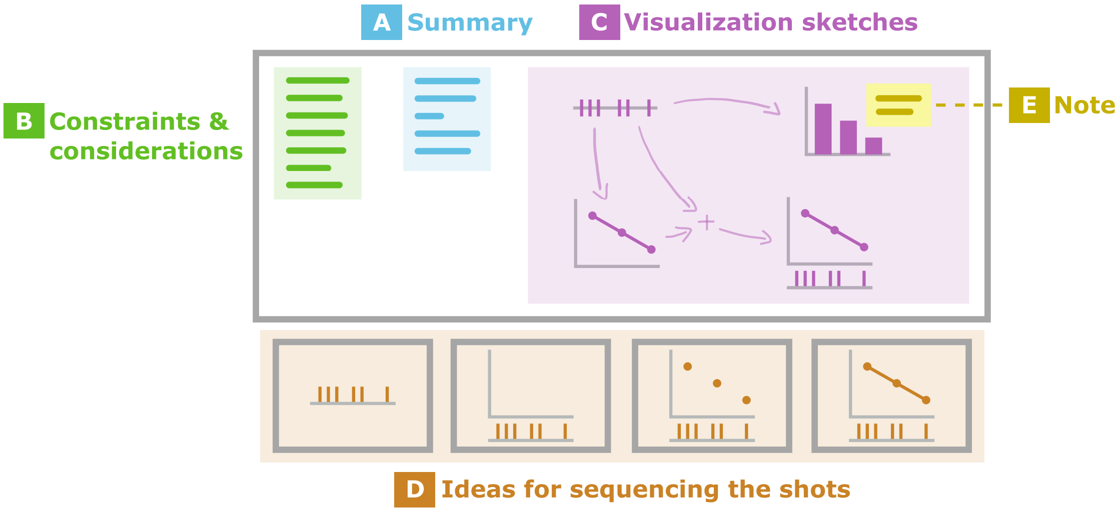


Figure 20: An illustration of what the components of our storyboard looked like for a single scene. We used these components for each scene. *(Figure adapted from Mayer et al. [May+24], licensed under CC BY 4.0.)*

as the *actors*. Drawing the actors from the data set instead of creating ideal artificial examples makes the underlying data more tangible.

Moreover, we wanted to keep the same actors in focus across the scenes. This way, it is easier to follow the story's main thread and grasp the interaction between the different mechanics. However, when drawing examples from the actual data set, it is not trivial to ensure that the examples are suitable to explain all mechanics equally well. Therefore, we thoroughly examined the story pieces to determine the constraints that the actors should fulfill. We weighted the constraints to be able to automatically determine the most suitable actors in arbitrary data sets.

Draft text and visuals. As visualization researchers, this step was the most extensive. We group our remarks based on the topics *genre*, *text*, *text-vis combination*, *staged animations*, *visual style*, and *flow factors*.

Genre. We created the story as a *dynamic slideshow* [Rot21] with the corresponding slideshow-style layout, see Figure 21. Slideshows provide control over the story pacing as they limit users' skimming and allow to deliver key plot points more clearly [Rot21]. In addition, they allow to keep content persistent across slides to incrementally build up more complex views [Rot21], and they provide a clear mapping between text and visualization [ZOM19].

Text. We drafted text at different stages of the design process. Initially, we outlined the scene contents and collected key points on the storyboard. During the development, we roughly decided what text should be provided in unison with the visuals. In the end, we finalized the text drafts with our domain expert, making sure to follow data provenance guidelines like providing data sources and additional refer-

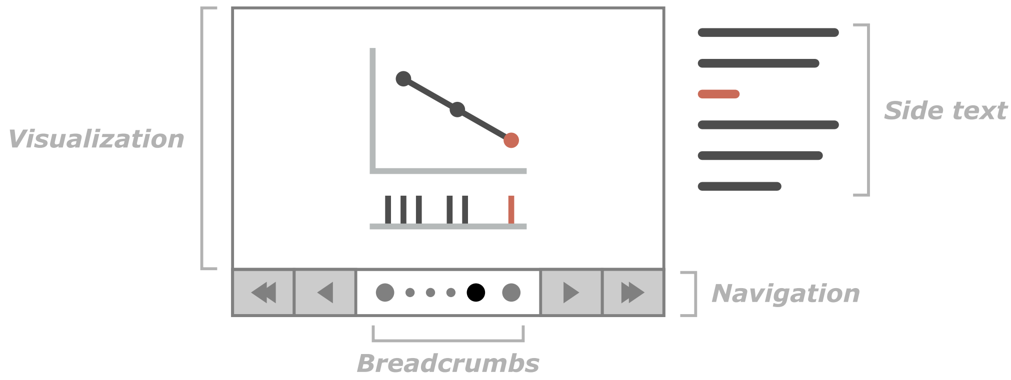


Figure 21: A single shot of our story consisted of a visualization and navigation component on the left and a text component on the right. The red color exemplifies a linking between visuals and text. *(Figure adapted from Mayer et al. [May+24], licensed under CC BY 4.0.)*

ences [HD11]. As the story should be adaptable to different data sets, certain parts of the text had to be kept variable, including automatic interpretations of the intermediate and final results.

Text-vis combination. A dedicated visualization panel took up the larger fraction of the screen space, with a smaller text panel accompanying it to its right, see Figure 21. This way, the visualizations received the main focus to "foreground the topic through perceptual layering" [YLT18]. By revealing each new view in synchrony with the corresponding text, their connection was established, while linking individual textual and visual elements via color [Sto+16], like in Figure 21. In general, we often verbalized what was depicted in the visualizations and stressed visual differences in the text, e.g., between certain groups of data points [LCB21]. The most common types of annotations we used in the visualizations were "text," "shapes" like arrows, and "highlights" of important areas [Ren+17].

Staged animations. When building up each scene into a final view, most of the visual changes in our story are animated. This can lead to less attention drift and better comprehension for the users [Ami+18] and increase their engagement [McK+17]. We also break up the animations into multiple transitions between the individual shots, see Figure 20 (D). Such staging was shown to improve recall and understanding [HR07]. Within each scene, we tried to keep the transition costs minimal, i.e., the number of visual changes when going from one visualization to the next [Hul+13].

Visual style. Despite the positive effects of embellishments on memorability [Bat+10], for ExEx, they should be used thoughtfully to maintain a professional and credible appearance. However, a visually attractive style can let users have more patience and look at visualizations more deeply [CM07]. As depicted in Figure 18, we aimed for a clean and aesthetically pleasing style but, overall, the preference regarding the visual style may vary depending on the targeted research community. In addition, there is the question of how closely the visualizations in the story should stick to the style and content of the figures used in the underlying paper, in which the explained method was originally introduced. High similarity makes it easier to connect the information from the story and the paper. However, it may result in missed opportunities for improving upon the design from the paper, including non-static options like interaction and animation. We took inspiration from the paper but reworked two visualizations (Scenes 1 and 4) and created two entirely new visualizations (Scenes 2 and 3).

Flow factors. Flow factors capture how the navigation of a story interacts with its visual components [McK+17]. As navigation control, we provided buttons, see Figure 21. They allow the user to either transition to the next shot with an animation or to skip the animation and jump directly to the next shot. The navigation progress could be tracked via so-called *breadcrumbs*, which could also be used for navigation, see Figure 21. The story progression was intended to be linear, going from one shot to the next, but the user also has the option to skip ahead using the breadcrumbs if they are already familiar with certain concepts and want to save time. Overall, our story would be classified as a *stepper*, which was the most engaging story style in the experiments of McKenna et al. [McK+17].

4.4.2.4 *Review and iterate*

As in all visualization projects, reviewing and iterating on our story was essential. This was particularly relevant for ensuring that the designers had correctly understood and explained the relevant mechanics of the communicated method. Having to communicate such complex information, we did not decouple the implementation of the visual aspects entirely from the narrative structure, contrary to other approaches [SH14]. Creating preliminary visualizations was necessary as a basis for discussion. However, depending on how big the resulting changes are, they can be quite impactful on the story, since the explanations build on each other, as do the visual design decisions. Therefore, good abstractions and storyboarding processes are still important to keep the later changes at a minimum. For storyboarding, Amini et al. observed a highly iterative and non-linear process [Ami+15]. In their study, participants often alternated between *reading and interpreting data*, *selecting data*, *crafting the narrative structure*, and *integrating strategies to engage viewers*. This was also the case in our design process, for which our storyboard and prototyping notebooks were of great help.

Note: Some of the approaches from Table 2 also include the evaluation of the story as part of a storytelling design process [Lee+15; CA+20]. We describe our evaluation in a separate section, as it is more extensive than what will be possible under the circumstances commonly present when performing ExEx. The reason for this is that we could refer to two complementary user groups which will not always be possible.

4.4.3 *Visualization Design*

Above, we have explained our decisions regarding the *story design*. In the following, we provide an overview of the central decisions regarding the *visualization design*, and how they supported the explanation of the method.

Scene 1. As depicted in Figure 22, three map views are displayed for the treatment actor and the dependent events around it, and three map views are displayed for a matching control actor. The maps to the left display the dependent events before the intervention, and the maps to the right the dependent events after the intervention, each with a dashed circle as a reference to the location of the intervention event. This way, the different time intervals are separated and the treatment and control scenario are contrasted. In addition, timelines represent the dates of the events. By creating separate maps for the events before and after the interventions, we intended to communicate more clearly which events took place in which interval.

The stripes on the timeline are binned based on their time of occurrence. The bin size corresponds to the temporal base unit of the data. In conflict research data, this unit is often day-based. Moreover, we selected the gray color for the control event to clearly communicate that these events are only of secondary interest in the statistical analysis, whereas the red treatment events are of primary interest.

We opted for disaggregated representations to allow the users to target individual events in the interaction of the scene's final shot. Accordingly, when hovering over individual dots on the map, the corresponding stripe on the timeline is highlighted and additional information is displayed, see Figure 22. This linked highlighting also works in the other direction, highlighting the corresponding dot on the map when hovering over a stripe on the timeline. In the context of the example application EA from Section 4.3, the additional hover information includes details such as what type of aid project an intervention actor refers to and how expensive it was. Moreover, synchronized panning and zooming of the maps corresponding to the same scenario allow closer examination of the events.

Scene 2. We display the aggregated counts of the dependent events before and after the interventions, see Figure 23. The timelines were carried over from Scene 1. Visual aids above the timelines explain how the trend is calculated, including an arrow illustrating the direction of the trend. Annotations highlight the key information and adjust upon interaction.

To interact, users can change the size of the temporal window by dragging the handles at the bottom of the view to see how the counts and the trends would change if a different temporal window was used, see Figure 23. Based on this adjustment of the temporal window size, the trends between the treatment and the control scenario can become so dissimilar that MWA would not consider them a valid pair for compari-

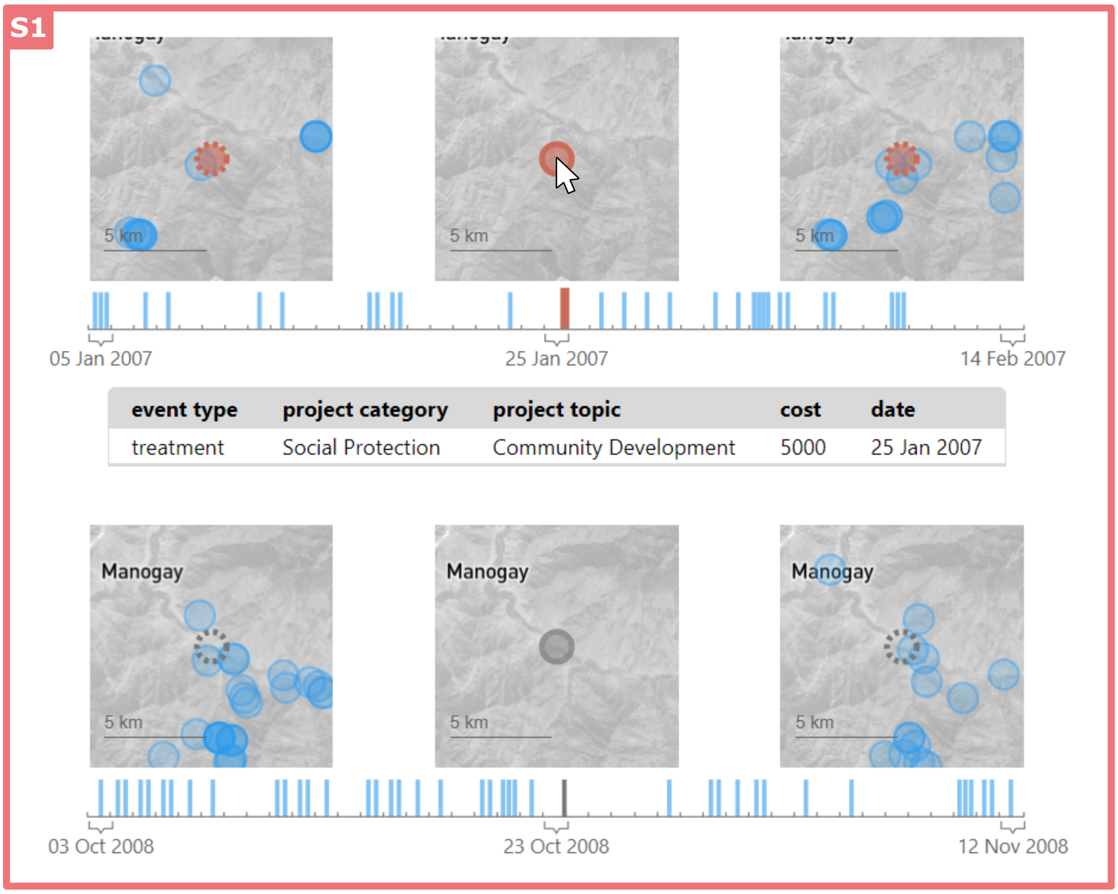


Figure 22: Hovering over a dot or stripe glyph in the final shot of scene 1 (as indicated by the mouse cursor) highlights it and reveals additional information. *(Figure adapted from Mayer et al. [May+24], licensed under CC BY 4.0.)*

son anymore when running the analysis for the corresponding temporal window size. While the details of this similarity calculation are explained only in the next scene, this exception is already indicated in this scene. Accordingly, if the trends become too dissimilar, the arrow indicating the trend turns orange.

Scene 3. As depicted in Figure 24, one pair of histograms is displayed for each *matching variable*. As explained in Section 4.3, matching variables are additional data provided separately by the communicating researcher to cover for potential confounding factors, like the population density of the area surrounding the intervention events. The histograms provide an overview of how the treatment and control events are distributed across the matching variables. This is relevant because events can only be *matched*, i.e., included as valid pairs for performing inference, if they are similar

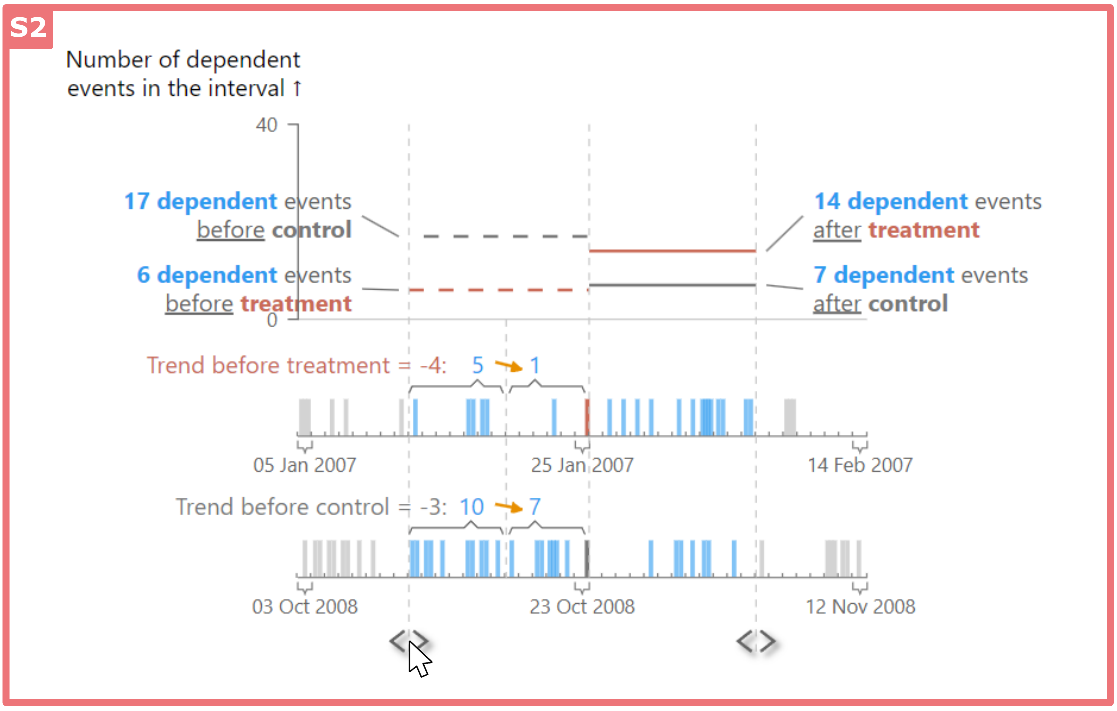


Figure 23: Dragging the handle below the timelines in the final shot of scene 2 (as indicated by the mouse cursor) adjusts the size of the temporal window across the entire view. If this causes the *trend* before the treatment event to become too dissimilar from the trend before the control event, the arrows indicating the respective trends above the timelines turn orange. This is the case in the figure, where the trend before the treatment event is -4 and, accordingly, too dissimilar from the trend before the control event, which is -3. *(Figure adapted from Mayer et al. [May+24], licensed under CC BY 4.0.)*

enough across all matching variables. The trend, explained in Scene 2, is one of the matching variables.

The stripe glyphs representing the intervention actors in the previous scenes are carried over to Scene 3. However, they are replaced by the aggregated histograms throughout the scene and, accordingly, not visible in the scene's final shot depicted in Figure 24. This transition comes with a more general shift of perspective in the story, away from the individual example events towards an aggregated view of the data, as this aggregated view is more in line with how the actual inference works.

We chose histograms to represent the distributions as they are a common technique to visualize this type of information. Moreover, the fact that histograms are made up of multiple adjacent bars was convenient as we could set the bar widths to represent more clearly how the underlying similarity calculation is performed. To determine if two events are *similar* with respect to a matching variable, the underlying matching algorithm [IKP12] divides the range of all values observed for this matching variable into equally-sized bins. All events whose values regarding a certain matching variable fall into the same bin are considered similar with respect to that matching variable. We used these bin widths as the widths of the histogram bars. If two events are similar regarding all matching variables, they are *matched*.

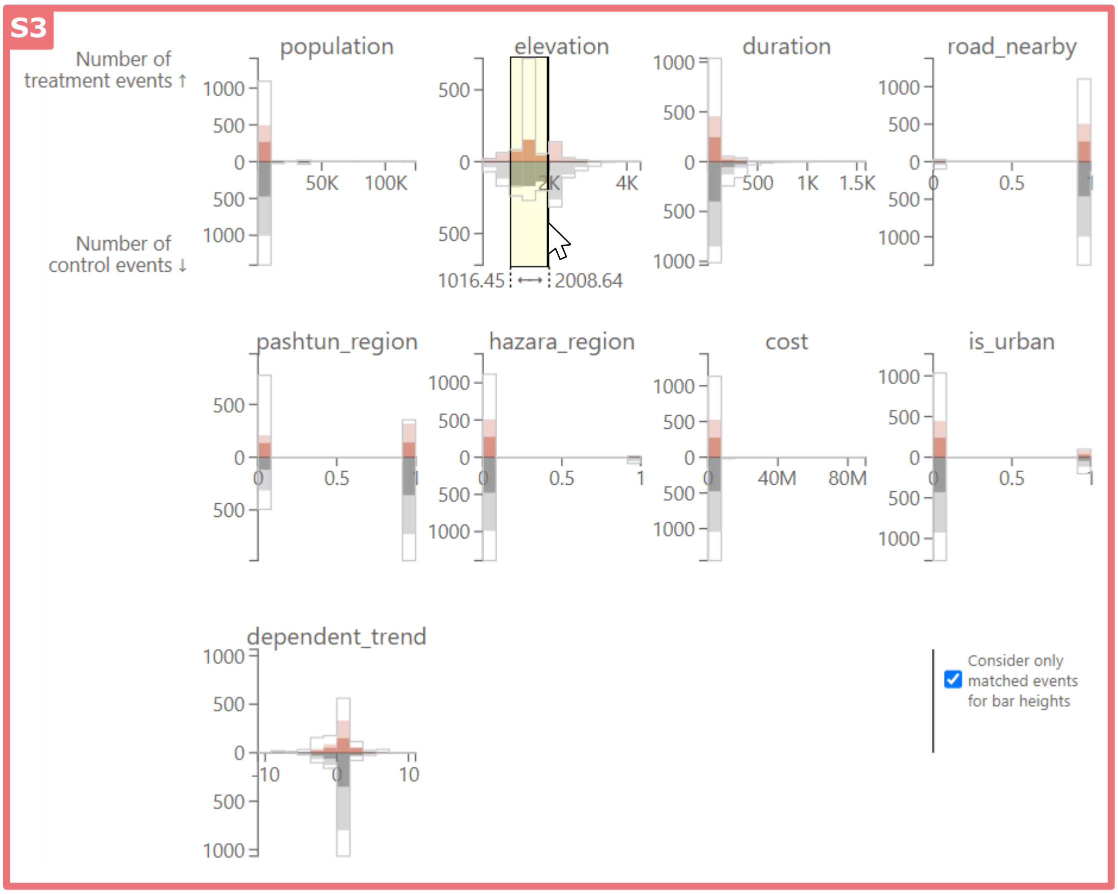


Figure 24: The final shot of scene 3. The yellow rectangle on the “elevation” histograms represents a filter that spans roughly three bins of the histogram. Across all histograms, the bars that represent events which do not lie within the “elevation” filter are represented in a lighter shade of their base color. The distribution of the events which were not matched is indicated only by a light gray outline across all histograms. *(Figure adapted from Mayer et al. [May+24], licensed under CC BY 4.0.)*

Aside from that, we placed the histograms for the treatment and control events separately below each other instead of stacking them, as it makes it easier to assess their distributions separately. Moreover, stacked representations tend to be more difficult to understand [Lee+19], while we aimed at facilitating understandability. Regarding interactivity, the views can be toggled to only display the matched events. In addition, on each pair of histograms, a filter can be applied to remove the events across all histograms that do not fulfill the filter constraint, see Figure 24. With these interactions, users can get a better impression of how the data is distributed, and it also becomes clearer how the likelihood of events being matched is reduced the more matching variables the communicating researcher includes.

Note that the x-axes of the histograms are not labeled directly, only via the respective chart titles, which correspond to the names of the matching variables in the data set. Accordingly, no separate information about the units is provided if the communicating researcher does not include it in the variable names. This solution was preferred by our collaborating expert as it keeps the additional information minimal that needs to be provided when loading a new data set into the story.

Scene 4. As depicted in Figure 25, a heatmap gives an overview of the results of the analysis when performed for different spatial and temporal windows. The effect size estimate is mapped to the color and the significance of the results is mapped to the size of the tiles. This allows to quickly detect patterns in the results and to assess for which window sizes the effects are meaningful. A similar figure was used in the original paper [SD14] and can be produced using the MWA *R* package [Mwa]. One key difference is that, in the original figure, the color scale is a sequential grayscale, whereas we use a diverging color scale going from red over white to blue. This way, our figure has two advantages over the original figure. On the one hand, the zero point of the effect estimate scale can be identified more easily, and, on the other hand, positive estimates (in shades of blue) can be distinguished from negative estimates (in shades of red) more clearly.

Another difference to the original figure is that it encodes the significance of the results, i.e., the p-value, not via the size of the tiles but via a hatched texture. Accordingly, all tiles have the same size, and one of three types of textures is overlaid over each tile. If $p < 0.05$, no hatching lines are overlaid and the tile is plainly visible. If $0.05 \leqslant p < 0.1$, dashed hatching lines are overlaid, and if $0.1 < p$, solid hatching lines are overlaid. In contrast, with our size-based encoding, the combinations of spatial and temporal windows pop out more clearly for which the results were significant. In addition, those combinations for which the effect estimate carries no meaning, as the results were not significant, receive a lot less visual attention.

To interact with the view, the user can hover over the tiles. This reveals the precise effect estimate for the corresponding spatio-temporal window. The value is displayed in the "estimate" legend such that the corresponding number can be better put into reference to the full range of computed estimates, see Figure 25.

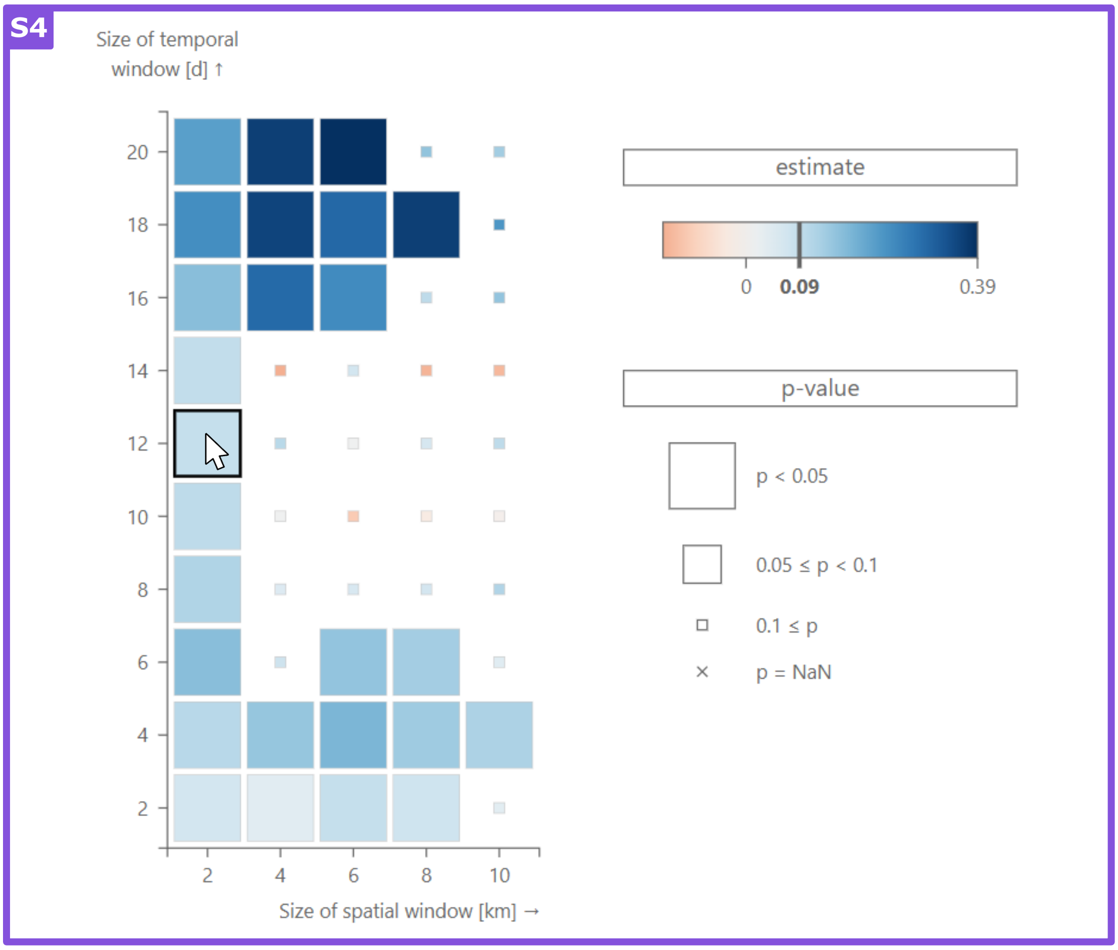


Figure 25: When hovering over a tile in the final shot of scene 4 (as indicated by the mouse cursor), the tile is highlighted and the underlying value is displayed in the "estimate" legend, here 0.09. *(Figure adapted from Mayer et al. [May+24], licensed under CC BY 4.0.)*

4.5 EVALUATION

For the evaluation, we invited eight peace and conflict researchers. Four of them had used MWA for at least one of their papers before, the other four had not. Hereafter, we refer to the first group as *experts* and to the second group as *semi-experts*. The semi-experts were two Ph.D. students and two post-doctoral researchers. The experts all had a Ph.D. and between three and 13 years of experience in conflict research. All participants stated that they see data visualizations on a regular basis, with all but two of the experts producing visualizations themselves.

For each expert, we created a version of our story based on one of the papers in which they had applied MWA. This resulted in three stories, as two experts had worked on the same paper. With each participant, we scheduled an interview in which, after a brief introduction, they were given the link to one version of the story. For the experts, we used the story corresponding to their paper, and for the semi-experts, we used the story based on the example study EA [KS18] introduced in Section 4.3. While following the story, the participants shared their screen and could ask questions or give feedback. Afterwards, we collected additional verbal feedback and sent them a questionnaire to fill out later. The interviews lasted 45 minutes on average (with a median of 38 minutes) of which reading the story took 23 minutes on average (with a median of 20 minutes). In the evaluation, we wanted to validate two main aspects:

(A) How engaging and how understandable do the participants find the story and its content?

(B) What do the participants think about the approach of ExEx in general?

Note: While memorability is also an important aspect in visual storytelling, we did not focus on it, as it was not central to our overall narrative intent of explaining the method.

Regarding (A): To test the engagement, we referred to a set of 5-point Likert scale questions proposed by Hung [Hun19]. As Hung focused on measuring engagement for individual visualizations and we wanted to test an entire story, we replaced the occurrences of "visualization" in the items by "tutorial." For the evaluation, we chose the term "tutorial" over the term "story" to prevent the participants from expecting a (fictional) story in the sense in which the word is rather used in everyday language. The resulting items are listed in Table 3 (E1-E11).

To test how well the participants understood the content of the story, we adapted the approach by Burns et al. [Bur+20]. They used the *taxonomy of educational objectives* originally proposed by Bloom [Blo56] to evaluate six different dimensions of understanding in data visualization. For our work, the dimension of *comprehension* was the most central. It refers to how well "learners understand the underlying information as a whole" [Blo56]. Burns et al. evaluate this by asking the participants in an open-ended question how they would summarize the conveyed information to explain it to a friend. We adapted this example and asked the participants the following question after they had finished the story: *"Imagine that you want to explain the content of the tutorial to a peer without showing them the tutorial. How would you describe the method explained here, Matched Wake Analysis, in your own words?"* (U0)

The participants were instructed to answer question U0 verbally. Aside from *comprehension*, we were also interested in the dimensions *knowledge* and *application* [Bur+20]. *Knowledge* refers to whether a user is able to "recall or recognize factual information" [Blo56]. *Application* describes whether a user is able to "break down a topic into parts and understand the relationship between each part" [Blo56]. We covered these two dimensions in our analysis of the answers to question U0 as the question also depends on these dimensions. In addition, we wanted to know whether certain aspects in the story were unclear or explained in too little or too much detail. Accordingly, for each of the four core scenes of the story, we included items U1-U6 from Table 3. For each scene, we also included items I1-I4 to test how well the interactive components were perceived.

Regarding (B): To investigate the participants' general interest in expert explorana-tion, we included items G1-G6 from Table 3. In addition, we collected verbal feedback during the interview.

4.5.1 *Results of the Evaluation*

In this section, we use the item IDs depicted in Table 3 and Figure 26 when referring to specific items from the questionnaire. Overall, the participants received the stories very well. They liked the direction of conveying methodological understanding in this way and mentioned that the "narrative aspect of it helps to contextualize [the presented information]." They also found the visualizations, animations, and interactive elements "extremely helpful" and "very valuable" as they "create an intuition that just cannot be created verbally." They said that the direction was "incredibly important" but "underexploited" and that "the future [lies] here." In the following, we group the quantitative and qualitative feedback based on overall topics. We primarily discuss the critical feedback, though the participants expressed that their complaints were already at a high standard.

4.5.1.1 *Engagement.*

As depicted in Figure 26, the overall engagement was quite high, with the following exceptions.

New knowledge and concepts. The question whether new knowledge and concepts were acquired was rated mixed, as the experts already knew MWA and two of the semi-experts had already read the paper before. However, participants who had already read the paper stated that the story helped them to recall information that, for some, was multiple years old. In addition, for those semi-experts who had not read it before, the learnings were large, allowing them to summarize the workings of MWA in surprising depth after using the story. Remarkably, also one of the experts gained new insights, as Scene 3 made them think about the mechanics of MWA in a way they "didn't [before], because we never made that graph."

Entertainment. The rating of one expert showed that they were not entertained during the story. This expert also remarked that the pacing of the story was generally too slow for them as they were already familiar with a lot of MWA's underlying mechanics. They suggested that a more time-efficient version of the story would suit them better. Other participants suggested that, at the beginning of each scene, a high-level

	Id	Item
Engagement	E1	My use of the tutorial is continuous and smooth.
	E2	I feel motivated while using the tutorial.
	E3	I DON'T feel frustrated when using the tutorial.
	E4	I enjoy exploring the tutorial.
	E5	I have acquired a new concept or new knowledge from the tutorial.
	E6	I think the tutorial effectively delivers its main concept or idea.
	E7	I think the tutorial is telling a compelling story.
	E8	I think the tutorial sparks my creative thinking.
	E9	I feel entertained when using the tutorial.
	E10	I feel absorbed by the tutorial while using it.
	E11	The look and feel of the tutorial is pleasing to me.
Understanding	U1	I think there are unclear concepts or explanations in this section.
	U2	I think the following concepts or explanations are unclear. *(Free text answer)*
	U3	I think there are aspects that should be explained in more detail.
	U4	I think the following aspects should be explained in more detail. *(Free text answer)*
	U5	I think there are explanations that go into too much detail.
	U6	I think the following explanations go into too much detail. *(Free text answer)*
Interaction	I1	I found the interactions appropriate.
	I2	I found the following interaction options unnecessary. *(Multiple choice)*
	I3	I would have liked to explore certain aspects in more detail.
	I4	I would have liked to explore the following aspects in more detail. *(Free text answer)*
General	G1	I would recommend the tutorial to a fellow researcher who wants to familiarize themselves with MWA.
	G2	I would use a tutorial like this instead of reading the underlying paper.
	G3	I would use a tutorial like this in addition to reading the underlying paper.
	G4	I think that the tutorial can help researchers assess whether MWA is applicable to their research questions.
	G5	I think the following information is missing from the tutorial to help researchers assess whether MWA is applicable to their research questions. *(Free text answer)*
	G6	In my research, I see the following other applications where this kind of tutorial can be beneficial. *(Free text answer)*

Table 3: The questionnaire consists of 5-point Likert scale items (aside from certain exceptions declared in brackets). For semi-exerts, G4 and G5 were phrased to refer to the participant directly and not to researchers in general.

summary would help, stating which scientific problem is solved using which approach

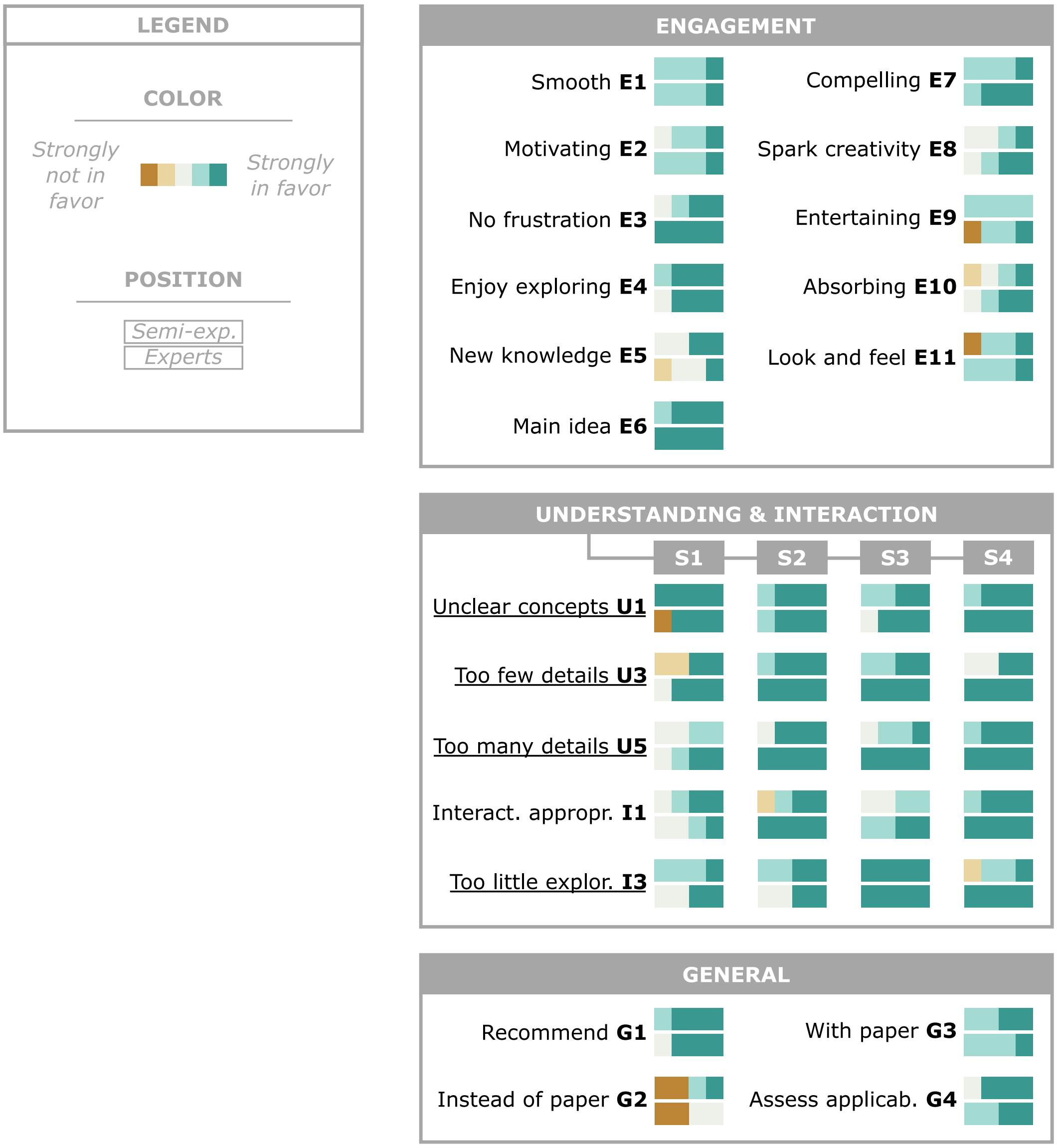


Figure 26: The ratings from the Likert scale questions. The bars for the ratings of the semi-experts are always depicted above the bars for the experts' ratings. The bold item IDs correspond to the IDs in Table 3. For negative items (which we underlined), we have flipped the color scale, so green corresponds to answers in favor of our stories across all items. The abbreviations "S1" to "S4" stand for "Scene 1" to "Scene 4." *(Figure adapted from Mayer et al. [May+24], licensed under CC BY 4.0.)*

in that scene. For instance, in Scene 3, a "potential selection bias" is prevented using "statistical matching." This way, users already familiar with the underlying approach of statistical matching could skip ahead and save time.

Look and feel. A semi-expert found the aesthetics of the story unpleasing (E11). They expressed that they strongly disliked the red color used for treatment events and would have preferred a monochromatic color scale across the entire application. In contrast, an expert said that the choice of colors was good and "made absolute sense." In general, the participants perceived the look and feel of the story very well, describing it as "amazing," "very cool," and "very good." They also used all navigation options, going forward and backward with and without animation options, using the breadcrumbs to jump to specific shots, and even following links to external websites with further information. However, the linear navigation through the story using animation was the default for all participants, showing that the animations were relevant for them. In addition, the decision to change and add new visualizations compared to the figures from the underlying paper was appreciated. One expert even stated that when they had worked with MWA, they "used the default output visualization of MWA [which was also used in the original MWA paper], but yours [in Scene 4] is much more clear." Moreover, the fact that the visualizations in Scene 2 and 3, which we had created entirely new, led to new insights for some experts, also showed the benefits of going beyond the figures used in the original paper.

4.5.1.2 *Understanding and Interaction.*

Regarding the request to summarize MWA after using the story (U0), the participants answered quite similarly. Overall, the key points related closely to the four main mechanics presented across the four core scenes of the story. The relationships between the different parts were reproduced well, speaking for good *application* [Blo56]. Regarding the *knowledge* dimension [Blo56], all recalled information in the summaries was factual. The key difference was that some participants were more explicit about certain aspects than others. Most remarkably, one semi-expert and one expert did not explicitly mention that there are two different types of intervention events, the treatment and control events. Upon request, the participants who were already familiar with the method said that it substantially helped them to refresh the memory of the method, and that they could not have summarized it in such detail anymore without using the story.

Generally, the *comprehension* [Blo56] seemed very good. This is also reflected in the self-reported understanding from the quantitative feedback. It was very positive, with a few exceptions that we discuss in the following. Overall, Figure 26 conveys that most understanding issues occurred in Scene 1. However, the corresponding free text feedback revealed that the remarks rather required more general context regarding the overall story instead of criticizing Scene 1 specifically.

Context. Contrary to the point mentioned before about improved time-efficiency for experts with strong prior knowledge, it was also remarked that, for a certain part of the scientific audience, the prior knowledge might be relatively low. Conflict research is special in that regard, as there is a divide between qualitative and quantitative scholars. According to one participant, to allow also experts with little quantitative background to follow the story, aspects like the basic idea of *scientific comparison* between treatment and control groups would need to be explained. Accordingly, in the beginning, the big picture could be summarized in a more abstract way, e.g., that the intuition behind the approach is that "in the data, there are some natural experiments hidden that have to be found, and the right things need to be compared for doing so." This could "buy the attention" of a less quantitatively-oriented audience. In addition, a participant requested whether the final textual interpretation of the results could be made even more clear regarding which units the calculated effect size is represented in.

Text structure. It is not visible from Figure 26, but in the third scene, some participants had minor issues following the main thread of the scene. It was caused by some of the text blocks being quite large, with too little visual structure. Three participants suggested adding more paragraph breaks or using a bullet point text structure. One participant also asked whether the text could be embedded more directly into the visualizations.

Interaction. Overall, the provided interaction options were perceived very well. They were explored by all participants, most deeply the interactions in Scene 2 and Scene 4. Some participants did not see the necessity to pan and zoom the maps in Scene 1 and to draw filter constraints into bar charts in Scene 3. The more straightforward interaction options were adopted better, i.e., hovering over elements for additional information (Scene 1 and Scene 4) and dragging a handle (Scene 2). However, despite their ease of use, these options yielded substantial additional information. In contrast, the option to (un)tick a checkbox in Scene 3, which did not reveal entirely new information, was seen as unnecessary by three participants.

4.5.1.3 *Applicability*

Based on the replies to G3, all participants expressed the desire to use a story such as ours in addition to the underlying paper for familiarizing themselves with a new method. However, the preferences regarding when to use it (before, while, or after reading the paper) differed. All were mentioned in the oral feedback, but, primarily, the participants would use it while reading the paper, particularly when reading the method section. However, a difficulty mentioned was that, to do so, people would need to be made aware that the story exists before reading the paper. One way would be to make sure that it can be found when looking up the method on search engines.

As expected, in most cases, the participants would not use the story as a replacement for reading the paper (G2). They would only do this if they wanted to get an overview of multiple existing methods, but not if they actually wanted to apply the underlying method in their own research. They expressed that the reason was not an issue of trust, but the fact that the paper contains more details than the story, which are necessary when actually intending to apply the method. Overall, though, the participants thought that the story can help researchers assess whether MWA is applicable to their research (G4) and that they would recommend it to fellow researchers (G1). All participants expressed a strong desire to use a tailored version of the story to explain to other researchers how MWA was used in their own research. One expert even asked whether they could share their version of the story on X (formerly Twitter).

Aside from MWA, the participants immediately came up with other methods that would benefit from ExEx, including complex analytical methods in general, like causal and spatial analysis, but also concrete methods [Kel19]. They also saw great potential for ExEx stories in teaching, primarily for students in their Master's and onward. In addition, it was mentioned that such stories can help less technical project partners understand how a complex method was used in a collaborative project.

Adding to that, one expert said that ExEx could even help those researchers who have already understood and applied the explained method. Accordingly, the expert who received a new perspective on MWA when they used our story said that such a story "can challenge you on how much you really understood - that really happened to me." Moreover, two experts said that such a story can be helpful for reviewers when submitting a paper that used the method explained in the story as "you can send the link with it like: here, in a nutshell." For such cases, when the story is used by experts with a strong background, and time efficiency plays a bigger role, again, the approach mentioned before could be beneficial: Equipping the story with a technical summary at the beginning of each scene would help to allow users already familiar with the explained mechanic to skip ahead and save time. In this case, though, the users should be able to select whether they want this more advanced version of the story or just the standard version. This would help to prevent the technical summaries from scaring away users with less prior knowledge.

Lastly, one participant mentioned that it would also be nice having the option to export animated data-GIFs [Shu+21] from certain transitions, or to export individual (static) views, including the option to load new data into the view.

4.6 REFLECTIONS

We summarize the insights obtained from our evaluation and from the previous works we used as the foundation of our design process. In that, we refer to the participants of the evaluation simply as "participants," and we do not list all the references to the previous works again that were already listed in Section 4.4. Of course, our suggestions should not be taken as strict rules. Further projects and evaluations are necessary to validate the applicability of the suggestions on a broader set of scenarios.

Don't assume knowledge from the paper. The feedback showed that potential users may want to use an ExEx story while or even before reading the corresponding paper.

Therefore, no prior knowledge from the paper should be assumed in the story. Moreover, the story does not have to replace the paper, as most participants expressed that they would rather use it to complement the paper which is likely more detailed than the story.

Collaborate closely. As in other interdisciplinary visualization projects, close collaboration with domain experts is key. In ExEx, it is relevant for making sure that the story designers have properly understood the underlying method, to identify which mechanics are most relevant to explain, and for drafting the final story text. Moreover, as much access as possible to the implementation of the explained method is beneficial, and iterations will most likely be necessary.

Prototype thoroughly. When designing an ExEx story, it is important to prototype it thoroughly, as changing substantial aspects at a later stage can be very costly since the different parts of the story (including visual design decisions) build on each other. Abstracting the method well and determining good actor constraints at an early stage is key, as is the selection of efficient prototyping tools. For bridging the gap between prototyping and implementation, the platform Observable served us well.

Use an appropriate layout. Based on the insights from previous works, we selected the slideshow genre for our story. It allows for a clear mapping between visualization and text, the possibility to progressively build up views, and straightforward navigation. We strengthened the mapping between text and visuals by linking them via color, annotations, and verbalizing the visual information. However, as suggested by one participant, interleaving text and visuals even more might reduce the effort for shifting the focus between visuals and text back and forth with each new shot. Moreover, the accompanying text should be structured well to avoid big blocks of text.

Stage the scenes. In each scene, we build up a visualization step-by-step, breaking down a complex mechanic and a correspondingly complex visualization to make it more accessible and easier to understand. Animating the transitions helps to make the visual changes easier to follow and was received well.

Provide interactivity. The participants adopted the interaction options well and gained new insights from them. The options that were straightforward but at the same time yielded new information were perceived best.

Dare to invent visualizations. When creating visualizations, we found that sticking exactly to the figures used in the underlying paper was not necessary. We based some of our visualizations on figures from the paper, but also created new ones. This helped the participants to see certain insights more clearly or to even gain new insights they had not gained when working only with the paper and the underlying method.

Provide clear interpretations. While we already provided automatic interpretations, it showed that they could have been even more precise, stressing the requirement of having clear automatic interpretations of visualizations and results.

Provide context and outline. Both at the beginning of the story and at the beginning of each scene, a clear outline should be given of the information to be presented. Also, the context in which the explained method can be applied should be provided. Moreover, if the method is explained by following an example application, also the context of this example application should be given. We tried to do so, but the participants would have liked even more such overview information. It was also confirmed that, even for an expert audience, the prior knowledge of the users can vary strongly. Accordingly, to not overwhelm users with limited prior knowledge, there should not be too many technical terms in such overviews. At the same time, to not lose users with a strong background, we suggest the next point.

Consider a time-efficient version. Particularly for experts with a strong background, it may pay off to provide a version of the story where the contents are summarized on a technical level at the beginning of each scene. This way, more knowledgeable experts would be able to progress through the story faster, only stepping through a scene in detail if they are unfamiliar with the explained concepts. The detailed version of the scene could be the same as for the original stories, so only little additional implementation effort would be required. At the beginning of a story, users could be given the option to chose whether to use the more advanced version or the original one. However, it needs to be noted that experts can gain new insights even if they are already familiar with a certain topic, as long as they are open for it, as it was the case with one participant.

Consider the benefits. As with all visual storytelling applications, a considerable amount of work goes into the design and implementation of an ExEx story. This also includes the domain expert whose close collaboration is needed, see the earlier point. However, the broad applicability of such stories and the corresponding increase of outreach can help to motivate experts to put in the extra work necessary [Bey+20]. Accordingly, the participants already had a number of additional applications in mind for which the stories could be used, ranging from teaching to supporting analysts that are already using the explained method.

Following Meredith's line of argument, putting effort in the creation of such stories can pay off even more [Mer21]. Accordingly, additional audiences aside from the intended research colleagues could be reached, like potential students or stakeholders, if the jargon and the level of detail are adjusted. A natural next step in this direction would be to conduct a study with students to investigate how strongly an ExEx story would need to be adjusted to support them in their education.

Automate. To reduce the work of the designer, research on automating the process of creating an ExEx story is crucial. While we have identified potential challenges and guidelines for ExEx in this work, more automated processes need to be built on them to make the creation of ExEx stories more accessible and economically reasonable on a larger scale. Current research, e.g., regarding more complex slideshows, is still lacking [Che+23]. A good first step for facilitating the story creation is to provide structured access to the byproducts from the creation of the content to be explained, i.e., in our case, the scientific method. Code-based notebooks are a great way to do so, as we also benefited strongly from the notebooks we had created in our ExEx process.

Make it visible. Informing potential users of a method about the existence of an accompanying ExEx story in time is key for supporting their familiarization with the method. The feedback showed that some users might use such a story even before reading the paper, and most while reading it. Providing a link to the story on the method's project website or in a corresponding software package would be beneficial. Sharing it on social media can also help, as one participant offered to do. Ideally, the story could be mentioned already in the paper, though, at the point of submitting the original paper, such a story probably does not exist, yet.

4.7 CONCLUSION

We analyzed how exploranation can be applied to support the communication of scientific methods between researchers. We focused on explaining a method for spatio-temporal causal inference between conflict events by producing a story that can automatically adapt to specific application scenarios of the method. We abstracted our design process and evaluated different versions of our story with eight conflict research experts, of whom four were already familiar with the method.

The results show that expert exploranation is a promising direction, as the experts received the story very well, seeing various applications for such stories. Their feedback allowed us to derive suggestions for how expert exploranation can be performed, ranging from the visual and textual design of the story to making it more efficient for users with differing background knowledge. Yet, substantial research is still necessary, both for validating our findings to produce more generalized recommendations, and for allowing automation in the creation of expert exploranation stories to make it more time-efficient.

4.8 ADDENDUM

In our study, it became clear that ExEx is a promising direction for visual storytelling, but also that the implementation effort for producing such stories is still very high. Especially creating the framework to navigate between shots, including animations, skipping of animations, and jumping to non-neighboring shots, required substantial effort. Therefore, in a student thesis, we abstracted the framework created in the context of our ExEx study and extended it, producing a revised framework called GSAP-ASEQ [Gö23]. The student uploaded the code for the framework to GitHub [Gsab], and provided an example story using it on this website [Gsaa]. The framework supports the creation of animated slideshow-style stories and can be used together with D3.js [D3]. It provides the following functionalities, see also Figure 27:

- **Animation.** According to its main purpose, the framework supports the creation of animations for transitioning from one shot to the next.
- **Bidirectional navigation.** The user can navigate from one shot to the next in both directions. When going backwards, the corresponding animation is played in reverse.
- **Handling of active animations.** Consider the scenario in which an animation from shot t to shot $t+1$ is currently playing. If the user clicks to proceed to shot $t+2$ while the animation is still playing, the currently active animation is finished at increased speed, and the subsequent animation follows seamlessly at normal speed. In contrast, if the user clicks to go back to shot t while the animation from shot t to $t+1$ is still playing, the currently active animation reverses right from the frame that it is currently at. Both of the described behaviors work analogously when the currently active animation that is interrupted is a reversed animation, transitioning from shot $t+1$ to t.
- **Jumping to shots.** The user of the story can jump from shot t to a shot further away than the neighboring shots, i.e., to any shot $t+k$ with $|k|>1$. In that case, all animations are skipped for the following reasons. When jumping, the main intention of the user is to save time. Therefore, the options are either to play multiple animations at increased speed or to skip them entirely. Playing them at increased speed is likely to confuse the viewer, so skipping is the better option.
- **Expressiveness.** The framework is designed to pair well with D3.js, which allows the creation of highly customized and expressive visualizations. Accordingly, the animations produced with the framework can also be customized to produce expressive transitions.
- **Staggered animation.** If several visual changes should occur in a single transition, it can be beneficial to *stagger* the animations to reduce the perceived com-

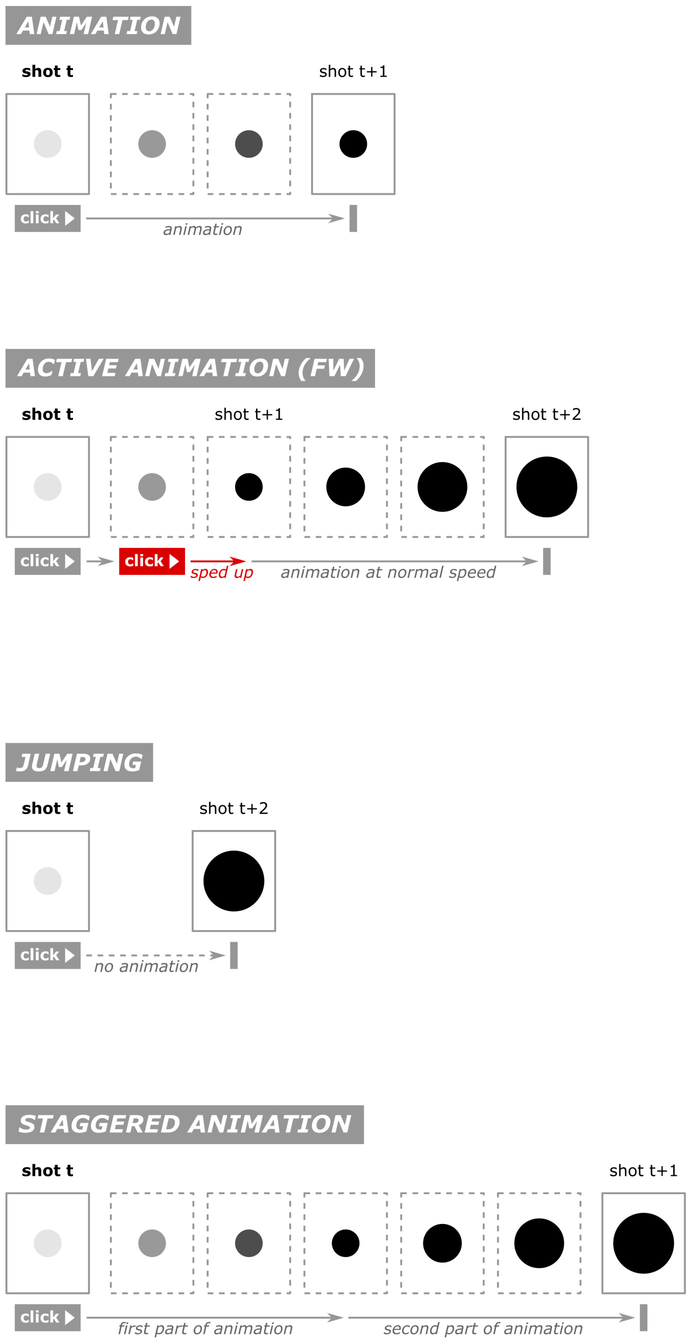


Figure 27: An illustration of the functionalities provided by GSAP-ASEQ [Gö23]. Rectangles with dashed outlines represent example frames from a fluid animation. The abbreviation "FW" stands for "forward," and "BW" stands for "backward."

plexity for the user [HR07]. Accordingly, the framework allows creating staggered animations.

As the example story illustrates [Gsaa], transitions created with the framework can be synchronized with the transitions of accompanying side texts. Moreover, the example shows that individual shots, like the final shot of the story, can be enhanced by including interactivity.

In addition to the functionalities listed above, the framework comes with the following benefits regarding the implementation of transitions.

- **One-directional programming.** The transitions only have to be implemented in one direction. When implementing them without the framework, using only D3.js, a lot more effort would be required. To provide animations in forward and backward direction, both directions would have to be implemented in a naive approach. When using a sequence of staggered animations for a single transition, this would become even more cumbersome, as the order of the entire sequence would have to be reversed.

- **Learning effort.** When using GSAP-ASEQ, only a limited number of functions have to be imported. They are used to facilitate the navigation between shots and to store the transitions. Optionally, information about the progress of running animations can be retrieved to convey to the user how far into an animation they are at any given moment. With this limited number of functions, the effort required for learning the framework is also reasonable.

As a consequence of the two benefits listed above, the amount of code that has to be written reduces drastically when using the framework, as compared to following a naive implementation approach without employing further libraries.

To provide the listed features, GSAP-ASEQ relies on the library GSAP [Gsac]. It is a library supporting the flexible creation of animations using JavaScript. However, the library is only free to use for non-commercial purposes, which poses a limitation for the use of GSAP-ASEQ.

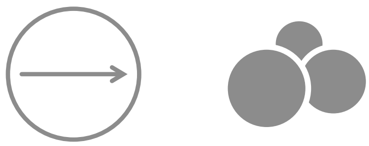
Across this chapter, we have seen the benefits of using techniques from visual storytelling when communicating scientific information to experts. However, only few of the works we could build on actually dealt with spatio-temporal data directly. Most of them had a more general scope. Accordingly, studies investigating how to perform visual storytelling in a spatio-temporal context could not only benefit broad audiences, for whom such stories are typically created, but also expert audiences, in the context of expert exploranation. For this reason, in the next chapter, we present a study in which we investigated which techniques are commonly used in data-driven online stories with a spatio-temporal context.

5 EXPLANATION FOR BROAD AUDIENCES

Large-scale struggles with a spatial and temporal context such as the COVID-19 pandemic, the war against Ukraine, and climate change have given visual storytelling with data a lot of attention in online journalism, confirming its high effectiveness and relevance for conveying stories. Thus, new ways have emerged that expand the space of visual storytelling techniques. However, interactive visual data stories with a spatio-temporal context have not been extensively studied yet. Particularly quantitative information about the used layout and media, the visual storytelling techniques, and the visual encoding of space-time is relevant to get a deeper understanding of how such stories are commonly built to convey complex information in a comprehensible way.

Covering these three aspects, we propose a design space derived by merging and adjusting existing approaches, which we used to categorize 130 collected web-based visual data stories with a spatio-temporal context from between 2018 and 2022.

An analysis of the collected data reveals the power of large-scale issues to shape the landscape of storytelling techniques and a trend towards a simplified consumability of stories. Taken together, our findings can serve story authors as inspiration regarding which storytelling techniques to include in their own spatio-temporal data stories.



This chapter is based on the following contribution [May+23b], licensed under CC BY 4.0:
B. Mayer, N. Steinhauer, B. Preim, and M. Meuschke. "A Characterization of Interactive Visual Data Stories With a Spatio-Temporal Context." In: *Computer Graphics Forum* 42.6 (2023), e14922. DOI: 10.1111/cgf.14922.

5.1 INTRODUCTION

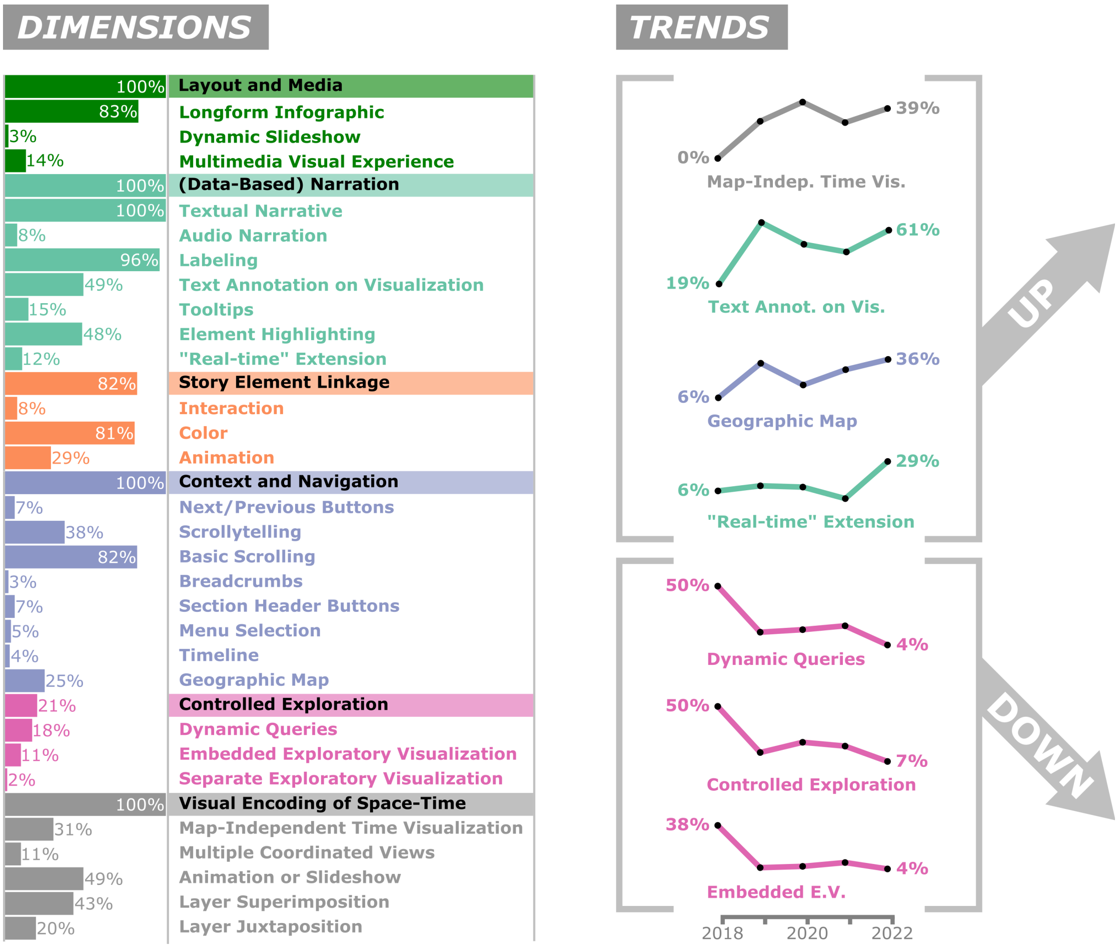


Figure 28: On the left, we display the summarized distribution of 130 collected stories after classifying them with our design space. On the right, the strongest trends in the data are displayed. *(Figure adapted from Mayer et al. [May+23b], licensed under CC BY 4.0.)*

The COVID-19 pandemic has brought interactive data visualizations front and center. These proved to be an effective way to display the spread of the virus in near real-time [LR20]. Graphical concepts such as “flattening the curve” became common terms. Maps, in particular, were used as a standard tool to show the distribution of COVID-19 cases, deaths, and vaccinations [All+20]. Russia’s invasion of Ukraine has also drawn the attention of many data teams [Eun22]. For instance, *The Washington Post* created “Four maps that explain the Russia-Ukraine conflict” [KM22]. Moreover, reporting on climate change makes considerable use of visualizations to convey the urgency of its messages, like in "A Quarter of Humanity Faces Looming Water Crises" [SC19].

These examples highlight the relevance and effectiveness of data- and visualization-based online journalism, whose traditional form has always been concerned with communicating information in an understandable, interesting, and relevant way [Ric+18]. By providing interaction, e.g., navigating via scrolling, the reader may perceive the story according to their own preferences.

Existing work on visual storytelling deals with the exploration of these new forms of communication [Bac+18; SZ18; SH10]. However, so far, little attention has been paid to data stories with a spatio-temporal context [Rot21; LCB21; MW18]. To better understand how such stories are built, we aim to contribute to their systematic characterization in the following way:

- We collected 130 web-based interactive visual data stories published between 2018 and 2022 with a spatio-temporal context.
- We derived a combined design space from three existing design spaces [Sto+16; Rot21; MW18] and adapted it based on our collected data to characterize the 130 stories.
- We analyzed patterns and trends regarding the characteristics that the stories have shown, e.g., which techniques were frequently used, as displayed in Figure 28, and how similar the stories are to each other.

The results of our analysis can serve story authors as a source of inspiration for which techniques might work well for their story. We also summarized the results in an interactive story notebook [Obsb] which we briefly introduce in this video [Not]. Our collected data set is not exhaustive and, thus, cannot represent the entire corpus of spatio-temporal visual data stories. Yet, we think that the analysis results provide valuable insights into the general trends in modern-day spatio-temporal storytelling.

Please note: In the main part of this chapter, we look at spatio-temporal data stories more broadly, not only focusing on stories based on conflict data. This makes the findings more generally applicable, but it can also benefit the insights gained for conflict stories: Often, an area can be better understood by contrasting it with comparable areas. In our case, one area would be stories about conflicts, and the other areas would be stories about different topics that also have a spatio-temporal context. Moreover, stories from related areas can provide inspiration for how to create conflict stories.

Nonetheless, to put sufficient focus on conflict stories, we have added a dedicated analysis comparing them in more detail to the other areas in Section 5.9.

5.2 RELATED WORK

First, we embed the three design spaces we built on directly in our study in the broader context of related works, and we highlight the differences to our approach. Then, we discuss what other aspects of spatio-temporal data stories were analyzed so far. Lastly, we set our contribution in relation to the current state of authoring tools developed for spatio-temporal data stories.

Foundation of our design space. Compared to approaches investigating how to tell visual data stories more generally [SH10], Stolper et al. focused on identifying more concrete *storytelling techniques* used across 45 stories [Sto+16]. Such techniques can give authors a more practical idea of how to craft stories. Only a few of Stolper et al.'s techniques directly consider a spatio-temporal context. However, after some adjustments, we could use them as a foundation for our design space, as many of their more general techniques are also useful for crafting spatio-temporal stories. Aside from the focus on spatio-temporal stories, our work differs from Stolper et al.'s approach in that we performed our analysis in a more quantitative way, and that we further subdivided and compared our collection based on the overall themes of the stories.

There also exist design spaces focusing explicitly on spatio-temporal storytelling that we could use as foundation. One of them consists of different ways to encode space-time, proposed by Mayr and Windhager [MW18], which we introduced in the BACKGROUND chapter. With some modifications, explained in Section 5.4.2, it was a suitable classification scheme to represent the different kinds of visual encodings used across our stories.

Moreover, we adapted the narrative *genres* from Roth's design space for map-based visual storytelling [Rot21] that we also presented in the BACKGROUND chapter. However, not all of the genres were relevant for our scope. We excluded *static visual stories* and *narrated animations* as they involve too little interactivity, and *personalized story maps* as they are not author-driven. Moreover, we did not include Roth's narrative *tropes* in our design space, as they deal with more high-level structural aspects than the aspects we focused on.

The three approaches we selected as foundation for our work are well-established and, taken together, provide an extensive overview of concrete design decisions for creating stories of various styles. Yet, none of them was applied to analyze such an extensive and recent set of stories as in our collection.

Further aspects. Roth's genres and tropes [Rot21] were further analyzed by Song et al. who studied what their impact is on readers' retention, comprehension, and reaction [Son+22a]. Moreover, Biriukov proposed a classification of maps in the context of interactive storytelling, along with recommended tools and libraries as development guidance [Bir21]. However, he considers maps on their own, independent of how they are embedded in a textual narrative, while texts were a key component of our collected stories. In contrast, Latif et al. explicitly examined the interplay of textual narratives and visualizations in geographic data-driven stories [LCB21]. In a qualitative manner, their analysis focused on corresponding strategies that they observed in 22 visual data stories. In comparison, our work aims to reveal insights about the use of visual storytelling techniques in a larger set of space- and time-dependent stories.

To engage readers, it can help to first make them reflect on what kind of patterns they believe to be underlying a certain issue before revealing the actual patterns to them. For this concept of *belief elicitation*, Mahajan et al. proposed a design space [Mah+22]. In the terminology of Bach et al., the concept is most closely related to the *make-a-guess* narrative pattern [Bac+18]. We did not incorporate it in our design space as barely any of our stories used such belief elicitation.

Authoring tools. While authoring tools are not the focus of this work, we still think that our findings are of use to their future development. So far, only few approaches with a spatio-temporal context exist. They include solutions to produce linked views [LJ12], space-time cube representations [KK17], video tours or slideshows from satellite images [Hsu+18], or camera movements for map views [Li+23]. However, all these approaches are quite narrow in their applicability. Therefore, we think that our findings can support the development of a more diverse set of authoring tools for spatio-temporal stories by helping developers to identify which techniques might be relevant to be provided in such tools.

Overall, though, in the terminology of Lee et al., the primary target groups of our study are the story *scripters* and *editors*, who plot the story and prepare its visual presentation [Lee+15]. In the following, we use the overarching term *story authors* for these roles.

5.3 DATA COLLECTION PROCESS

We present the sources from which we collected the stories and the criteria they had to fulfill to be included in our data set.

5.3.1 *Story Sources*

Following related approaches [Bre+16; Hul+13; McK+17; Sto+16], we relied primarily on online journals: *The New York Times*, *The Guardian*, *The Washington Post*, *The Economist*, *FiveThirtyAight*, *Bloomberg*, *The Pudding*, *Reuters*, *ProPublica*, *Berliner Morgenpost*, and *Zeit Online*. Most of them are commonly used sources for visual data journalism [Sto+16]. In addition, we included *National Geographic*, which Roth [Rot21] used as a representative source for spatio-temporal stories, along with the interactive *Manhattan Population Explorer* website [Fun18] and the *EcoWest* website [Econa] tracking environmental trends. As a time period, we selected the years from 2018 to 2022, thereby including a period before the COVID-19 pandemic and, hence, more diverse stories than the ones primarily related to the pandemic.

Among others, we considered the top-rated stories of these years listed by the community in popular blog lists of *FlowingData* and *VisualisingData* as "Best Data Visualization Projects of..." [Flo] or "Best of the Visualisation Web..." [Vis]. To also look more specifically for stories with spatio-temporal data, we searched for technique-related keywords ("Maps," "Timeline," "Interactive," "Graphics," and "Animation") and topic-related keywords ("War," "Climate change," "Wildfires," "Elections," "Protests," "Education," "Health," and "COVID-19") on the journal websites, the community blogs listed above, and via advanced Google searches. In addition, the links to related articles at the end of a story sometimes also led to other relevant stories.

5.3.2 *Criteria for Including a Story*

Each story had to contain at least one space- and one time-dependent visual component, e.g., a map with temporal information or a timeline visualization with spatial information in it. These components were allowed to appear in the same view or separated from each other. We did not require the presence of an actual map if spatial information was visually encoded in other ways, like via multiple area charts for different regions as in "There Has Been an Increase in Other Causes of Deaths, Not Just Coronavirus" [Lu20].

Moreover, the stories had to be data-driven and were required to integrate at least one interactive component that goes beyond clicking the play button in a video, leading to the exclusion of data videos. In contrast, we considered the interactivity sufficient if a story could be navigated by scrolling up and down. This is arguable, but as scrolling allows readers to consume stories at their own speed and easily navigate back and forth, we considered it justified. Moreover, we wanted to investigate the assumption sparked by Tse that the use of more elaborate means of interactivity was declining since around 2016 because "readers just want to scroll" [Tse16]. To allow for a corresponding analysis, our data also had to contain stories that use scrolling as their only form of interaction.

This way, we collected 16 stories from 2018, 21 from 2019, 32 from 2020, 32 from 2021, 28 from 2022, and one continuously updating story without definite year information [Econa]. For the example stories we refer to in this work, we also added the citation identifier in Figure 29 next to the row in which the story is classified. For a combined overview of all story titles and their classifications, please refer to the Appendix: Full List of Stories or our accompanying web-based story notebook [Obsb].

5.4 A DESIGN SPACE FOR INTERACTIVE SPATIO-TEMPORAL DATA STORIES

We created our design space by leveraging three related approaches [Sto+16; Rot21; MW18] and adjusting them based on the stories we had collected. In this section, we describe the classification process and the final design space.

5.4.1 *Story Characterization Process*

For the coding process, we took inspiration from existing approaches [SH10; Hul+13]. We first created an initial version of the design space, consisting only of the dimensions of the previous approaches. Then, three coders participated in the classification of the stories. First, one coder classified the collected stories. Throughout the first round of classification, potential issues and ambiguities with the design space were resolved through discussions of the three coders until a mutual decision was found. This also led to the modification of the design space by redefining, adding, or removing dimensions.

Once the classification by the first coder was done, the second coder went over all classifications, compared them to the original stories, and highlighted all classifications they did not agree with. Then, the third coder, the author of this dissertation, assisted to mediate discussions between the other coders, again, until a mutual decision was found. This resulted in the final design space and classifications. By relying on three coders, we aimed at reducing subjectivity but, of course, we cannot fully avoid it.

5.4.2 *The Design Space*

Below, we present the final dimensions of our design space, see also Figure 29. Throughout the classification, some border cases occurred. For full reproducibility, we wanted to draw a clear line in such cases. To do so, we referred to the original definitions from the underlying works as far as they seemed appropriate. However, they were sometimes not clear enough. In these cases, we drew the line as we considered it most reasonable. For disclosure, we include these cases in the descriptions below.

Layout and Media

To classify the techniques used for structuring the overall story layout and media selection, we took inspiration from the visual narrative genres proposed by Roth [Rot21]. Since we focused on interactive and author-driven stories, only three of them were relevant in our case. In **longform infographics**, the main story flows linearly in a vertical reading direction, usually navigated via scrolling. Text, graphics, and multimedia usually only have a limited width optimized for mobile devices. Maps, timelines, or other graphics can either flow smoothly along the entire story or be fragmented across different sections of the narrative. Despite the vertical arrangement of the stories, individual visualizations may allow to explore developments over time by scrolling in a horizontal direction [Tre+21] or through animations [GY20].

Dynamic slideshows create linearity through a series of slides or discrete visual panels of uniform size and format. The information is usually moved across the screen in a horizontal direction rather than in a vertical direction like in longform infographics.

Figure 29: Our design space and the story classifications with selected stories highlighted. Each row represents a story and the rows are grouped by theme. In each theme group, the stories are ordered according to their publication year: the earlier, the further up. *(Figure adapted from Mayer et al. [May+23b], licensed under CC BY 4.0.)*

To count as **multimedia visual experience**, a story had to contain at least one video or audio that is supportive for the transformation of the story's message. Such stories may include the perspectives of the people involved, along with a rich integration of images, maps, videos, or sounds, to create a deep sense of place. Accordingly, a single teaser video at the beginning of a story did not suffice, as it increases immersion only marginally. Deviating from Roth's original definition, we did not require the linearity of a story to be created through anchor tags or hyperlinks as this constraint restricted the definition quite heavily.

The next four categories are based on Stolper et al.'s work [Sto+16]. In our design space, we incorporated 18 of the 20 techniques they had proposed, and we added two new techniques into the appropriate categories. We also renamed one category to refine its scope.

(Data-Based) Narration

We adapted this category from Stolper et al.'s category *Communicating Narrative and Explaining Data* [Sto+16]. It contains techniques that facilitate the narration of the story, either based on data or by advancing the overall story without directly relying on data. In a **textual narrative**, the main points are conveyed through a text body. The text can either accompany a visualization on a slide or just have supporting visualizations embedded at intervals.

Audio narrations allow elements of the story to be tied more closely. We also included audio sequences that make the story more immersive, such as the songs of certain bird species [Ili18]. With **labeling**, we refer to textual support above or below a visualization or as headings for a slide or section. In contrast, **text annotations on visualizations** are part of a visualization and direct the reader's attention, e.g., to where on a map an attack took place. **Tooltips** also provide additional details. However, the reader needs to hover over the visualization to reveal them.

With **element highlighting**, attention is drawn to elements of a visualization, e.g., by coloring them, or by adding new elements, like a frame around an important region on a map. However, we required the highlighting to be author-driven to support the narrative, so simply highlighting a territorial border when the reader hovers over it did not suffice.

We added the technique of **"real-time" extension** to cover a special case of how the narrative can be communicated over time: Some stories were repeatedly extended or updated, therefore, conveying their narrative in almost real-time. This can be done, e.g., by adding new sections to a story [Sta22] or updating maps [Rot+22] to extend the story as new events unfold. The technique was inspired by Roth's genre of *Visual Story Compilation* [Rot21], in which different events on a given topic are presented as a set of visual summaries that link to separate, more detailed pages. This set of summaries is updated in near real-time as new events unfold. However, we rather witnessed such extensions within full articles and not as links from a summary page to external pages, so we included it as an individual narrative technique.

Lastly, Stolper et al. also identified the technique *Flowchart Arrows* [Sto+16] to connect parts of the story when the intended ordering may be unclear. However, as we never saw it used in our stories, we excluded it from the design space.

Story Element Linkage

In their category *Linking Separate Story Elements*, Stolper et al. identified three basic techniques to explicitly connect elements of a story, like text in various forms and visualizations [Sto+16]. The first technique is through **interaction**. This can be done by synchronizing a point in time across multiple timelines when hovering [Sta21], or via dropdown menus that adjust the entire story based on a selected location [Jac+20].

Elements can also be linked through **color**, using a consistent color mapping between elements in multiple visualizations or between text and visualizations. For instance, text referring to certain regions can be colored like the corresponding regions on a map [Lam19]. In addition, **animation** can be used to maintain the connection between elements that are used in multiple consecutive views, smoothly transitioning between the views, e.g., by transforming region glyphs from a map view into circle glyphs in an abstract 2D view [Lam19]. Scrolling is a common cue to trigger such transitions.

Context and Navigation

We renamed Stolper et al.'s category *Enhancing Structure and Navigation* [Sto+16] to *Context and Navigation*. In this context, *enhancing the structure* of a story usually meant informing the reader about *where in the story* the part is located that they are currently reading. However, we found that the techniques from this category were also used to inform the reader about *where in space and time* a certain part of the story took place. So, in the first case, *context* regarding the position inside the story was provided, and in the second case, *context* regarding the position in space and time. To capture both of these cases, we exchanged *structure* for *context*. In addition, we do not require a technique from this category that provides context to also be usable for interactive navigation.

The first technique, though, facilitates such interactive navigation: With **next/previous buttons**, the reader can navigate between discrete parts of a story, sometimes embedded in a continuous longform infographic [Ban+19]. Next, we split up Stolper et al.'s technique of *scrollytelling* into two separate techniques: First, **basic scrolling** describes the linear navigation through a story via scrolling, where the different text blocks and visualizations of the story stay in the same relative position to each other. Second, **scrollytelling** describes the case when scrolling is also used to trigger changes to visualizations or to move texts and visualizations across the screen at different speeds relative to each other, e.g., sliding text over a visualization or at its side. Sometimes, also both techniques are used in the same story [Ili18].

Breadcrumbs are used to let the readers know where in the story they are and, if interactive, to provide direct access to different parts of the story. Traditionally, they are represented as small dots, but they were also integrated more creatively, e.g., functioning at the same time as the temporal legend of a visualization [Tim+18]. Context and navigation can also be provided through **section header buttons**. For this, different sections of the story are titled and navigation is enabled by listing these titles at the beginning of the story. However, if too many options need to be listed, **menu selection** becomes more appropriate. We also included menu selections that update the entire story instead of just jumping to a specific section.

To represent at which time a certain part of a story took place, a **timeline** can be used, highlighting the corresponding time point or interval. Likewise, to communicate the location at which a certain part of the story took place, a **geographic map** can be used. However, to be included in this definition, a map's main purpose had to be to provide context, and not to display additional data. Such maps were often small, displaying only territorial boundaries and the current location of the story [Son+22b]. The other option for a map to be included here was if it provided means for navigation. However, only one map allowed such navigational aid, transporting the reader to different parts of the story when clicking on the corresponding regions [Econa].

Controlled Exploration

The last category, adopted one-to-one from Stolper et al. [Sto+16], covers different ways in which readers are allowed to explore visualizations in a constrained way that maintains the linearity of the story. They include **dynamic queries**, e.g., for filtering or adjusting parameters of a visualization, like the displayed attributes, or by flying to selected locations on a map [YSC19]. In addition, **embedded exploratory visualizations (e. v.)** provide more than the usual interaction options in a less restrained way, like panning and zooming in a map [YSC19]. Such a visualization can also be extracted as a **separate exploratory visualization (e. v.)** on a separate web page.

Visual Encoding of Space-Time

To characterize the encoding of space and time, we referred to Mayr and Windhager's visualization framework [MW18] after removing one of their proposed techniques and adding a new one. Example sketches for our set of techniques are displayed in Figure 30.

The first technique in this category, **map-independent time visualization**, was not part of Mayr and Windhager's framework. We added it as we found multiple stories in which time visualizations were not connected to a map but still included spatial in-

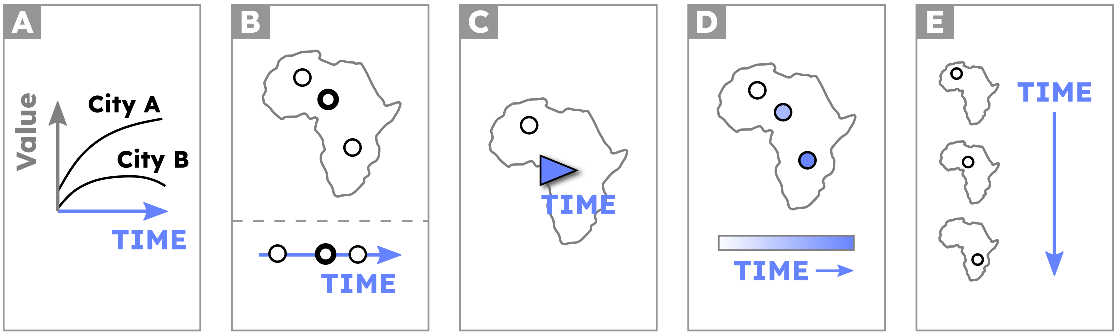


Figure 30: The figure is an adjusted version of Figure 2 from the BACKGROUND chapter, showing only the techniques for encoding space-time that we included in our final design space. The techniques encompass *(A) map-independent time visualization* with some additional value on the y-axis, *(B) multiple coordinated views*, *(C) animation or slideshow* with a blue triangle representing a play button, *(D) layer superimposition*, and *(E) layer juxtaposition*. *(Figure adapted from Mayer et al. [May+23b], licensed under CC BY 4.0.)*

formation, e.g., in the form of a line chart where the lines are labeled with the names of corresponding cities. With **multiple coordinated views** we refer to multiple visualizations that are linked through interaction, consistent use of color, or animation. These visualizations usually separate the spatial and temporal information into different views. **Animations or slideshows** represent the change of information over time as a smooth animation or a discrete sequence of steps in a slideshow. These representations can also be interactive to allow the readers to navigate back and forth. As an individual view, this technique can be used with any type of story layout.

In **layer superimposition**, multiple time-dependent information layers referring to the same spatial region are merged into an integrated view. At that, readers need to be able to recognize the temporal sequence of the events. This can be achieved, e.g., by using a time-dependent color coding, or trajectory glyphs indicating movement. In contrast, a choropleth map aggregating the temporal information for a region into an average value would not suffice. With **layer juxtaposition**, the spatio-temporal data is split into temporal slices referring to the same region. The slices are arranged side-by-side, often in reading direction. Examples include *small multiples*.

Mayr and Windhager included the *space-time cube* as an additional encoding option [MW18]. Since we did not encounter it in any of the stories, probably due to too high complexity for a broad audience, we excluded it from the design space. Moreover, we did not include a column for the case when no map representation was used at all; but for completeness, we list the five corresponding stories here [BPP21; LN20; Mal19; Rou+20; Lu20].

Meta Information

All of the mentioned techniques can be encoded as binary values: Either a story uses a certain technique or not. We also collected further information. First, the story **theme**, which is either *society & health*, *election*, *climate*, or *conflict*. We kept *society & health* as one theme as we encountered a lot of COVID-related stories where the transition between the impacts on health and on society was fluid. Moreover, we covered in which **year** a story was published and by which **source**, e.g., *The New York Times* or *The Guardian*.

5.5 ANALYSIS OF THE COLLECTED DATA SET

To analyze the data set, we pursued the following questions:

- **Q1:** General usage: Which techniques are used the most overall? Are there theme-specific differences?
- **Q2:** Similarity: How similar are the stories to each other? And how similar are the stories from different themes and sources?
- **Q3:** Trends: How did the usage of the different techniques evolve from 2018 to 2022?

To answer these questions, we relied on interactive visualizations that we implemented using the notebook-based platform *Observable* [Obsc]. We summarized the findings in an interactive story notebook [Obsb]. In this section, we present our findings regarding the analysis questions and a summary of the key takeaway messages for story authors creating their own stories.

5.5.1 *Q1: General usage*

We provide an overview of the most-used techniques, grouped by the main categories, with the complete list of technique usage frequencies being displayed in Figure 28. Afterwards, we report on noteworthy theme-based deviations from these patterns.

Most-used techniques. Regarding *layout and media*, **longform infographics** are by far used most frequently (83%), probably as they have a straightforward layout for online and mobile-based journalism [Rot21; SZ18].

For *(data-based) narration*, most authors used **textual narrative** (100%) and **labeling** (96%), compared to **audio narration** with 8%. The reason might be that textual elements are both easier to produce and more efficient to consume than audio narrations.

The *story element linkage* was most often implemented using **color** (81%) and least often using **interaction** (8%), possibly because interaction can be cumbersome to implement, especially across platforms.

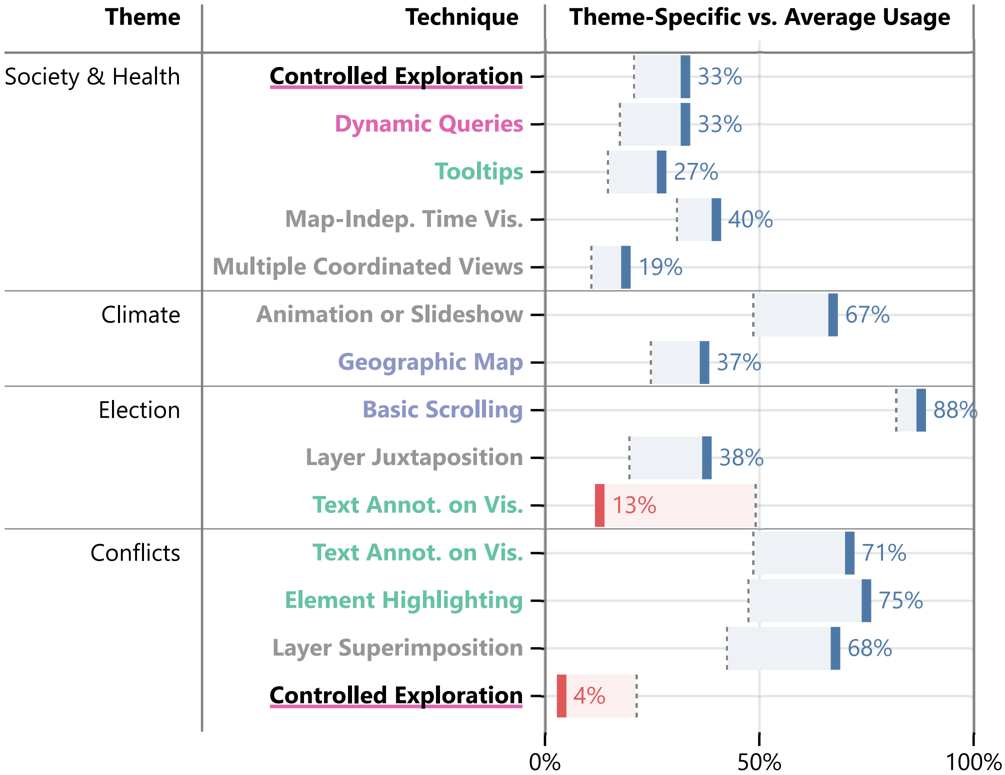


Figure 31: For selected cases, we depict how strongly the usage of certain techniques deviated from the respective overall average (represented as dotted lines) for the different themes. The annotated percentage values correspond to the absolute usage frequencies, not to the deviations from the average. *(Figure adapted from Mayer et al. [May+23b], licensed under CC BY 4.0.)*

Context and navigation was most commonly provided through the familiar interaction of scrolling, with 82% of the stories using **basic scrolling** and 38% the more advanced **scrollytelling**, sometimes both mixed in the same story (23%).

In contrast, techniques for *controlled exploration* were overall rarely used (21%) while, among them, **dynamic queries** were the most common choice (18%). The reason could again be the increased implementation effort, but also that exploration is more complex for readers than linear consumption.

The technique usage for the *visual encoding of space-time* was quite evenly distributed, with more easy-to-consume **animations and slideshows** taking the lead (49%) and the more cognitively demanding **multiple coordinated views** coming in last (11%).

Theme-based deviations. The deviations are summarized in Figure 31 in the same order in which we present them in the following.

- When considering only stories with the theme *society & health* (48 of 130), we found that they provided *controlled exploration* more often than average (+13%), particularly **dynamic queries** (+16%). Stories from the corresponding theme often cover regions in which the readers of the stories actually live, compared to reports on, e.g., climate disasters or violent conflicts. Hence, it is more meaningful to allow personalization of the story to the reader's home region. The most obvious examples of this are COVID dashboards, which also often include **tooltips** for more detailed regional information (+12%). They also use **map-independent time visualizations** (+9%) to depict infection trends and **multiple coordinated views** (+8%).

- *Climate* stories (46 of 130), on the other hand, focus strongly on *encoding space-time* as **animation or slideshow** (+18%). According to Harrower, "animated maps are better suited to depicting geographic patterns (and changes in those patterns) rather than specific rates" [Har03]. And illustrating a change in patterns is often a goal in climate stories. In addition, in some stories, larger parts play in a specific region particularly affected by climate phenomena. In those cases, providing *context* via **geographic maps** can be beneficial, especially if the regions are not generally well known. This was also done more frequently (+12%).

- *Election* stories (8 of 130) focused almost exclusively on **basic scrolling** (+5%) for providing *navigation and context*. Regarding the *encoding of space-time*, they used primarily **layer juxtaposition** (+18%). This can be convenient to compare election results from different years. A technique that was used a lot less is **text annotations on visualizations** (−37%). Usually, the information that is conveyed in such stories is not as distinguished and location-specific as, e.g., movements of troops. Thus, detailed annotations are not often required.

- *Conflict* stories (28 of 130), on the other hand, used a lot of **text annotations on visualizations** (+22%) along with **element highlighting** (+27%). This way, information like territory gains and losses, or troop movements can be conveyed more clearly. To visualize such rather spatially limited data, authors most often used **layer superimposition** (+25%), which allows to merge all information into a single view. In contrast, barely any *controlled exploration* was provided (−17%). In stories about conflicts, often very specific messages should be conveyed. In this context, allowing users to explore the data would be detrimental.

5.5.2 *Q2: Similarity*

We defined a *similarity score* to quantitatively compare the stories with each other. For two stories, s and s', it is defined as

$$\mathrm{sim}(s, s') := 1 - \frac{\sum_{i=1}^{n} \|s_i - s'_i\|}{n} \tag{2}$$

where $n = 29$ is the number of different techniques in the design space (excluding the six main category titles) and the classification of a story is represented as $s \in \{0, 1\}^n$. At that, $s_i = 0$ means that story s does not use technique i, and $s_i = 1$ means that story s uses technique i. With this definition, the similarity score sim has a maximum value of 1 if two stories are classified in exactly the same way, and a minimum value of 0 if their classifications differ for every technique.

Overall similarity. Figure 32 (A) displays the distribution of the pairwise similarities between all stories. They range from a minimum of 28% (one story pair) to a maximum of 100% (18 pairs) around a mean of 75%. This means that any two stories have the same usage pattern for at least 28% of the techniques, and that 18 pairs of stories use exactly the same techniques. Together with the bell-shaped distribution, this suggests that the majority of the similarities is homogeneously distributed.

This is also conveyed by Figure 32 (B). To produce it, we used the dimension reduction method *multidimensional scaling* (MDS) [HPG13], as it only requires a distance measure and no exploration of appropriate hyperparameters, contrary to other dimensionality reduction techniques like t-SNE [MH08] or UMAP [MHM20]. MDS takes the pairwise distances between high-dimensional elements as an input to produce a low-dimensional (here 2D) representation of the elements. At that, the coordinates of the projected points are optimized to reflect the distances between the original high-dimensional points as well as possible. As the distance measure between two stories, s and s', we used

$$\mathrm{dist}(s, s') := 1 - \mathrm{sim}(s, s') \in [0, 1]. \tag{3}$$

The resulting projection in Figure 32 (B) shows that the majority of the stories form a dense group of similar stories in the center. The other stories can be roughly characterized as follows.

- **G1** gives examples for animation-based stories as all of them use **animation or slideshow** to encode space-time and **animation** to link story elements. Moreover, all of them refer to **scrollytelling** for navigation and **element highlighting** for communicating the narrative.
- **G2** consists, among others, of COVID dashboards. 12 of the 15 stories are about *society & health* and all of them provide *controlled exploration*, along with **tooltips** in 10 of the 15 cases.
- **G3** represents stories that employ little to no advanced combination of text and visualization. None of its stories uses any *story element linkage* and the only form of exploration is provided via **separate e.v.** Moreover, all the stories use **basic scrolling** for navigation and none uses **scrollytelling**.

- **G4** seems to consist of stories that should draw the reader in more deeply, as all of them are **multimedia visual experiences** and use **audio narration**.

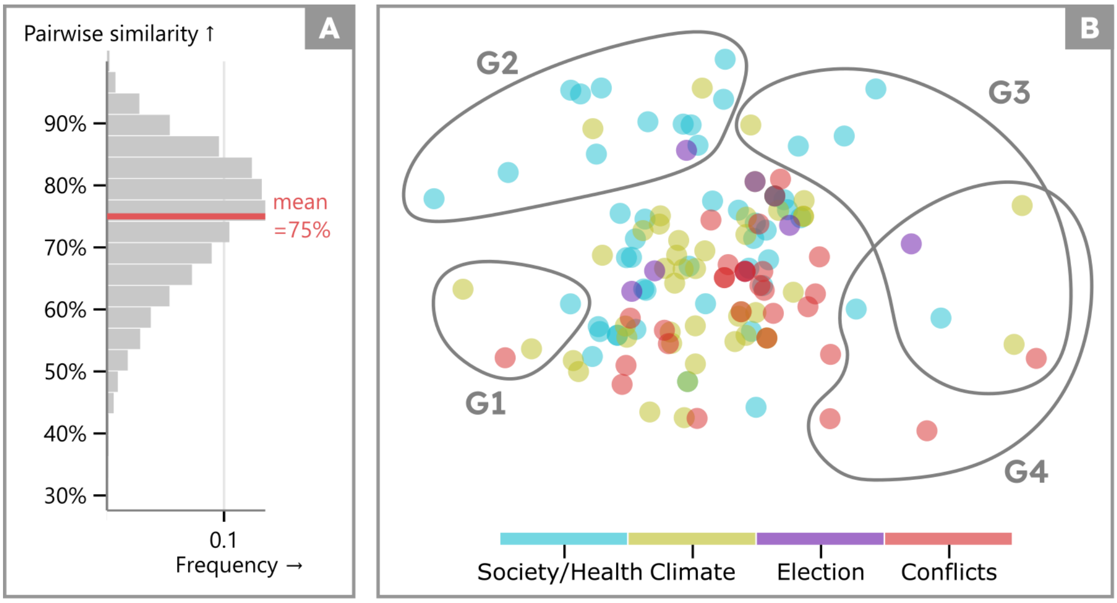


Figure 32: (A) The overall distribution of the pairwise similarities between all stories and the resulting mean value. (B) The stories as points projected into 2D using MDS. The closer two points are, the more similar the corresponding stories are. The point color depicts the theme of the story. *(Figure adapted from Mayer et al. [May+23b], licensed under CC BY 4.0.)*

Between-group similarity. To get a better understanding of the themes and sources present in the data set, we summarized the average usage statistics for each theme and source to compare how similar they are to each other. The average technique usage for the different themes turned out to be quite similar, with *society & health* and *conflicts* being least alike (86%), and *society & health* and *climate* being most alike (92%). These high similarities are also conveyed by the strong overlap of points representing stories from different themes in Figure 32 (B).

In contrast, the differences between the sources are stronger. The most significant patterns are highlighted in Figure 33 (A). There, the blue points represent all pairwise comparisons between the following sources:

- *The New York Times*,
- *Bloomberg*,
- *National Geographic*,
- *The Washington Post*, and

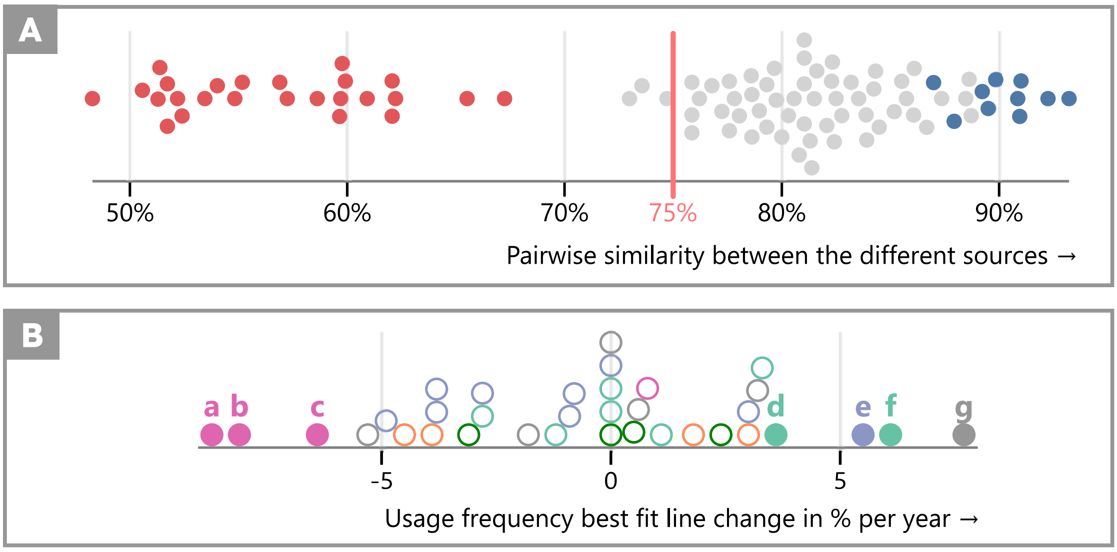


Figure 33: The beeswarm plots display distributions in a non-aggregated way. In (A), the pairwise comparisons of all sources are displayed. A group with very high similarity is colored blue and a group with very low similarity is colored red. The red line indicates the overall average similarity between all stories (75%). (B) shows for each technique the slope of the best fit line regarding the usage trend of the technique from 2018 to 2022. The strongest changes are annotated: (a) −8.7% per year, (b) −8.1% per year, (c) −6.4% per year, (d) +3.6% per year, (e) +5.5% per year, (f) +6.1% per year, and (g) +7.7% per year. *(Figure adapted from Mayer et al. [May+23b], licensed under CC BY 4.0.)*

- *Reuters*.

As they have similarities between each other of around 90%, these sources are most similar to each other. In addition, in Figure 33 (A), we have colored the points red that represent all comparisons between sources of which at least one was either the website

- *EcoWest* or
- the *Manhattan Population Explorer*.

This reveals that these two sources were very dissimilar from the rest, reflected through the lowest pairwise similarity scores.

5.5.3 *Q3: Trends*

To compare the trends, we calculated the best fit line for each technique regarding its year-wise usage frequencies from 2018 to 2022 and ordered the techniques accordingly.

The strongest *decrease* in usage over the years occurred for

- **dynamic queries** (−8.7% per year),
- *controlled exploration* (−8.1% per year), and
- **embedded e.v.** (−6.4% per year).

In Figure 33 (B), these correspond to the points **a**, **b**, and **c**. The respective R^2 of the estimations was 63%, 61%, and 49%. The underlying trends are shown in Figure 28.

The techniques whose usage *increased* the most are

- **map-independent time visualization** (+7.7% per year),
- **text annotations on visualizations** (+6.1% per year), and
- **geographic maps** (+5.5% per year) for context and navigation.

The corresponding points in Figure 33 (B) are **g**, **f**, and **e**, and Figure 28 shows the entire trends. The R^2 of the estimations was 51%, 27%, and 48%, respectively.

Possible explanations for the increases include the following. *Map-independent time visualizations* are often used in COVID-related stories, e.g., for comparing infection rates in different regions. The role that *text annotations on visualizations* play in depicting troop movements and territory changes in *conflict* stories was already discussed, as was the role that *geographic maps* play for *climate* stories to convey which specific regions are affected by climate phenomena. As all these issues have become more severe over the recent years, the corresponding techniques show a strong upward trend.

Another noteworthy trend occurred for **"real-time" extensions**. While its increase is lower than the last-mentioned techniques (+3.6% per year), it is still striking that it has the fourth-strongest increase (with $R^2 = 27\%$), see point **d** in Figure 33 (B). Naturally, fewer stories using this technique are created since such stories are usually extended when new events occur instead of creating an entirely new story. Therefore, this increase is more meaningful.

5.5.4 *Takeaway Messages for Creating Stories*

When searching for inspiration to create a story, the previous sections can help as follows. In Section 5.5.1, for each technique category, we list the most-used techniques in addition to the entire distribution provided in Figure 28. This serves as a starting point for which techniques might be suitable. The information about theme-specific deviations can help to narrow down the technique selection further. However, Section 5.5.2 showed that the theme alone does not automatically suffice to tell which techniques would be best to combine. In it, we presented exemplary groups of stories belonging to different themes but sharing key characteristics. Accordingly, it would be helpful for authors to not only get technique recommendations based on the theme but also based on the techniques they already know they want to use.

Going into detail about all possible technique combinations is beyond the scope of this work, but we provide a corresponding section "Inspiration for your new story" in our story notebook. There, authors may choose a theme and iteratively select techniques they want to use. After each selection, a view displays how often the other techniques were used in combination with the ones selected so far. In addition, a table lists all stories using the selected techniques so the author can get hands-on inspiration for how the techniques were combined in practice.

Example Scenario. Ayla wants to write a story about the deforestation taking place near her hometown. She has collected the data for the past four years about which areas were deforested in each year. Now, she wants to decide which techniques to use to tell her story. In the corresponding notebook section, she filters for the theme *climate* to find that, in such stories, the spatio-temporal data is often visually encoded as an animation or slideshow. She selects the technique to filter for all stories that used it.

She finds that, more often than usual, it is combined with scrollytelling for navigation. As she thinks that this would be an engaging way to navigate, she also selects that technique. This leads her to find that she should consider linking the separate story elements via animations and color. This way, she continues until she has decided on all the techniques she wants to use. She then refers to the table below the visualization and clicks on one of the stories that fulfill all her requirements to get a real-world example, e.g., an article by *ProPublica* discussing pollution [YSC19]. She likes how the article animates between different areas and decides to use it as an inspiration for her own story.

Aside from the steps described above, it can also help to consider the trends explored in Section 5.5.3 when deciding which techniques to use, even though they were only roughly estimated. It should also be noted that just because certain technique combinations were used often, they are not automatically the best. They might also just be, e.g., the most economical options for the publisher. Yet, we think that this can serve as a helpful source of inspiration and reflection.

5.6 Summary and Reflections

In this section, we summarize and reflect on the key findings from the analysis and discuss potential shortcomings of our approach.

Large-scale struggles shape storytelling techniques. The increasing trends discussed regarding **Q3** show that global issues like COVID-19, wars, and climate change drive the use of corresponding storytelling techniques. They also lead to the development of new techniques, like "real-time" extensions, which are particularly applicable for providing a more comprehensive picture of longer-lasting crises. While it is arguable that the occurrence of these crises might have skewed the picture conveyed by our data set, it is not unlikely for such trends to continue as diseases and violent conflicts on a larger scale may become more frequent due to climate change [IPC14]. Accordingly, the strength of the discussed trends in modern-day data-driven storytelling also reflects the urgency of the associated issues.

Stories are becoming easier to consume. In parallel with what Tse announced in 2016 regarding visual storytelling at *The New York Times* [Tse16], also stories with spatio-temporal context are on average becoming easier to consume. This comes with a reduced use of controlled exploration, interaction (aside from scrolling), and cognitively demanding design choices like multiple coordinated views. Yet, this has not always been the case. Stolper et al. had still encountered a number of exploratory stories [Sto+16] and the trend for controlled exploration techniques in our data had started out much higher in 2018 with 50%. As Tse expressed, potential reasons for the decline could be that making interactivity work cross-platform is very time-consuming but also the overall preference that "readers just want to scroll" [Tse16]. A potential reason for this preference could also be that the audience for data-driven stories has become broader, including also less technophile readers who are not very keen on interactive exploration.

Techniques that have, in contrast, become more relevant are non-interactive means to link separate story elements, like color, and ways to control animations, e.g., through scrollytelling. In such scenarios, scrollytelling is often restricted to animation-based sections embedded in a larger story that is otherwise navigated via basic scrolling. Basic scrolling is even simpler to control and was by far the most-used technique for navigation. This is also reflected in the fact that barely any slideshows were created, which, due to their nature, use the only layout not prone to scrolling. However, exploratory techniques are still in use, e.g., for creating a familiar setting for the reader [Bac+18], and more research could help to clarify in which cases such interactivity is worth the cost of creation [Hoh+20].

Storytellers combine techniques creatively. The similarity analysis showed that storytelling techniques are combined in a variety of ways across the collected stories. We identified key differences between the themes but also provided theme-independent examples in Figure 32, where we identified four groups that relied on quite different and sometimes even opposing technique selections. This variety across the stories aligns with Stolper et al., who also emphasized how creatively different storytelling techniques were combined in their analyzed stories [Sto+16]. By allowing authors to explore this space of design decisions in our story notebook, we provide them with a source of inspiration for meaningful technique combinations. Even more patterns could be detected by deepening the analysis, e.g., by using clustering or a similarity measure that better accounts for the hierarchy introduced by the different categories of techniques.

Aside from representing a creativity input for authors and a potential foundation for storytelling guidelines, the diversity of technique combinations also implies an important consideration for the development of authoring tools. Together with the changing technique usage trends discussed before, it shows that, ideally, authoring tools should be flexible if they are designed for a broad application: Not only should they provide a diverse set of techniques to combine, but they should also be able to adapt to changing demands over time, like the inclusion of new techniques. This also highlights the relevance of developing frameworks that are not as high-level as automatic authoring tools, and, accordingly, allow an easier inclusion of new techniques, while still saving the authors time, as they do not have to "reinvent the wheel" with each new story they create.

Authors and tool developers could be supported even more by collecting information about further aspects. Particularly, knowing which precise visualization techniques are used to encode what kind of data types in space and time can provide strong guidance. We observed quite divergent and creative techniques in that regard.

Limitations. Follow-up studies are necessary to investigate whether a high usage frequency of certain techniques also means that they were actually the best options for the respective stories. Moreover, the theme *election* contained comparatively few stories (8/130). A possible explanation is the following: While a lot of stories reporting on elections exist, most of them do not compare the behavior along both spatial and temporal dimensions, which was a requirement to be included in our collection. In addition, the theme is more narrow than, e.g., *society & health*, and the events that *election* stories report on, i.e., elections, take place more rarely than events that lead to a story about *society & health*.

As the data set is not exhaustive, the detected trends and patterns should not be taken down to individual percentage values. Rather, the overall directions and their implications are relevant. This also alleviates the fact that some estimations yielded a relatively low R^2 when comparing the trends of the technique usages. Moreover, in the trend analysis, we provide the underlying year-wise developments in Figure 28 for reference.

An issue that remains with collections such as ours is that some sources put up pay walls for their stories. While we paid for them in the context of our research, other people trying to access them might not be able to afford the costs. Still, neglecting stories entirely that are not freely accessible would likely skew the analysis results.

Lastly, our design space is only able to represent a subset of the characteristics that spatio-temporal visual data stories can show. Considering other aspects like narrative design patterns [Bac+18], the interplay of visualizations and textual narratives [LCB21], or more detailed distinctions regarding the visual encoding of space-time [Bac+17] would certainly yield additional insights. Similarly, describing the aesthetics of the visualizations in the stories, i.e., their characteristic style, in a generalizable way might be fruitful. It would allow investigating to which extent different sources have a characteristic visual style that they follow throughout their stories and what impact a certain style has on the reading experience. This style might also differ between the themes, e.g., regarding what level of terrain details are displayed on maps.

5.7 CONCLUSION

We collected a set of 130 interactive visual data stories with a spatio-temporal context and characterized them based on a combined design space derived from three existing approaches. With this, we could get a deeper understanding of how such stories are commonly built: We analyzed how frequently the different visual storytelling techniques have been used over the past five years and in which direction the corresponding trends go.

Struggles with a global impact, like the COVID-19 pandemic, wars, and climate change, have a strong influence on which techniques are used for visual storytelling, even resulting in the development of new techniques tailored to the corresponding needs, like “real-time” extensions. It also became apparent that interactivity and exploration are decreasing in modern-day storytelling, resulting in stories that primarily revolve around the scrolling interaction. However, exceptions exist when advanced interaction can be used to provide the reader with a personalized perspective on the story. Taken together with the more fine-grained patterns that can be explored in our data set, e.g., via the accompanying story notebook, our results can support authors in crafting successful stories and allow developers of authoring tools to identify which techniques their tools might be expected to provide.

5.8 APPENDIX: FULL LIST OF STORIES

We provide the full list of stories across Figure 34, Figure 35, and Figure 36, including the title, source, and year for each story. As in Figure 29, we include the reference shorthand for the stories directly referred to in this chapter.

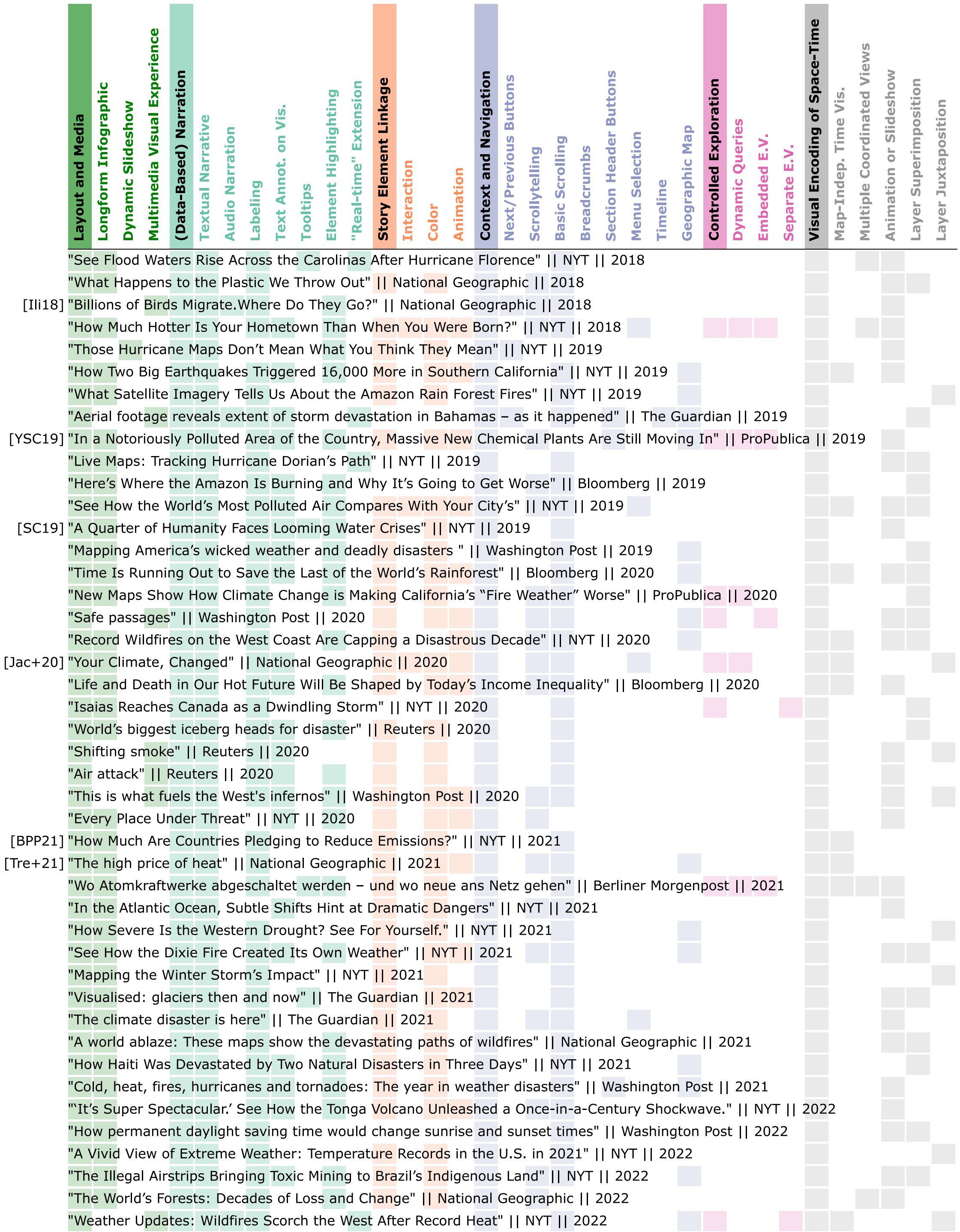


Figure 34: The full list of stories, part one of three. *(Figure adapted from Mayer et al. [May+23b], licensed under CC BY 4.0.)*

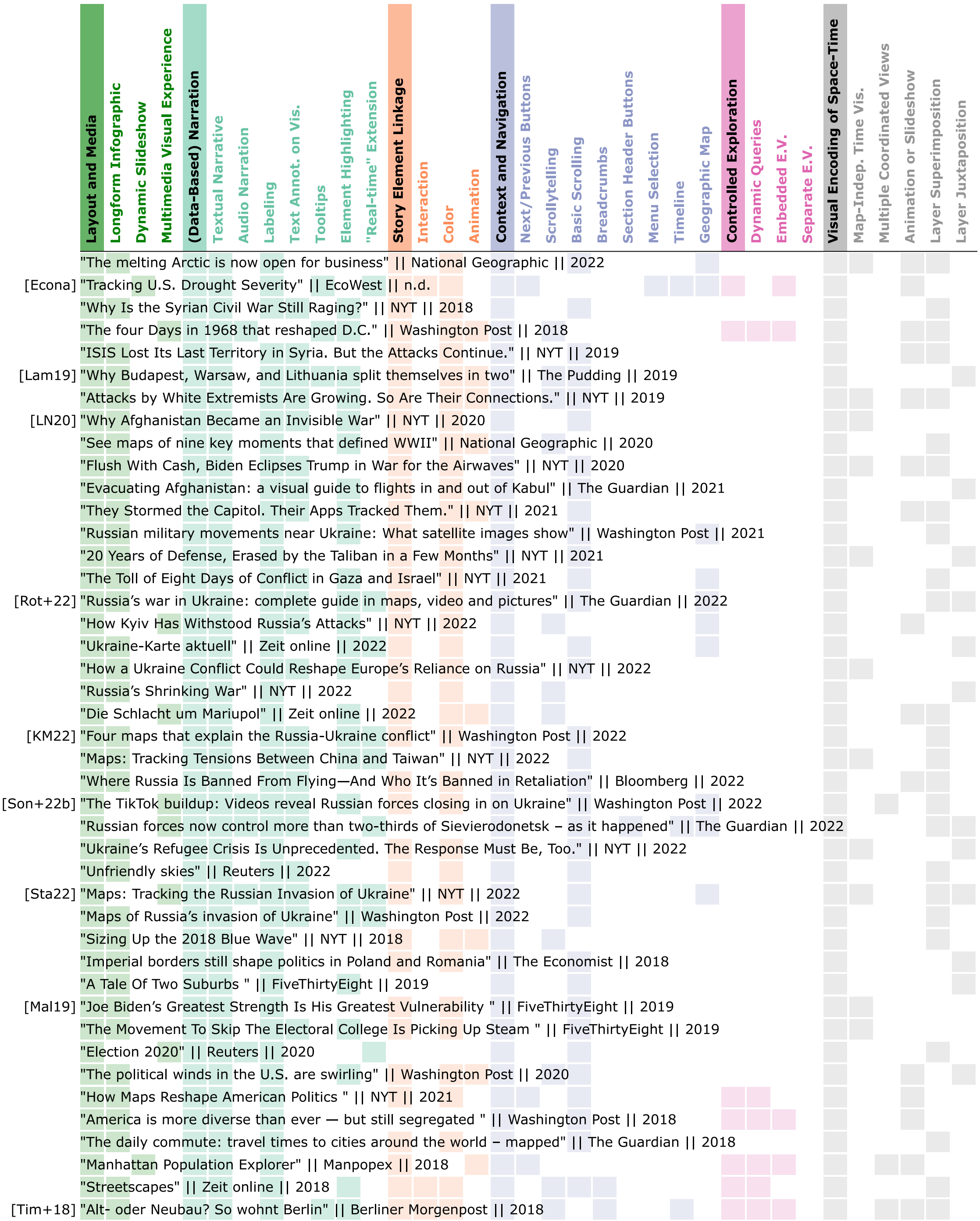


Figure 35: The full list of stories, part two of three. *(Figure adapted from Mayer et al. [May+23b], licensed under CC BY 4.0.)*

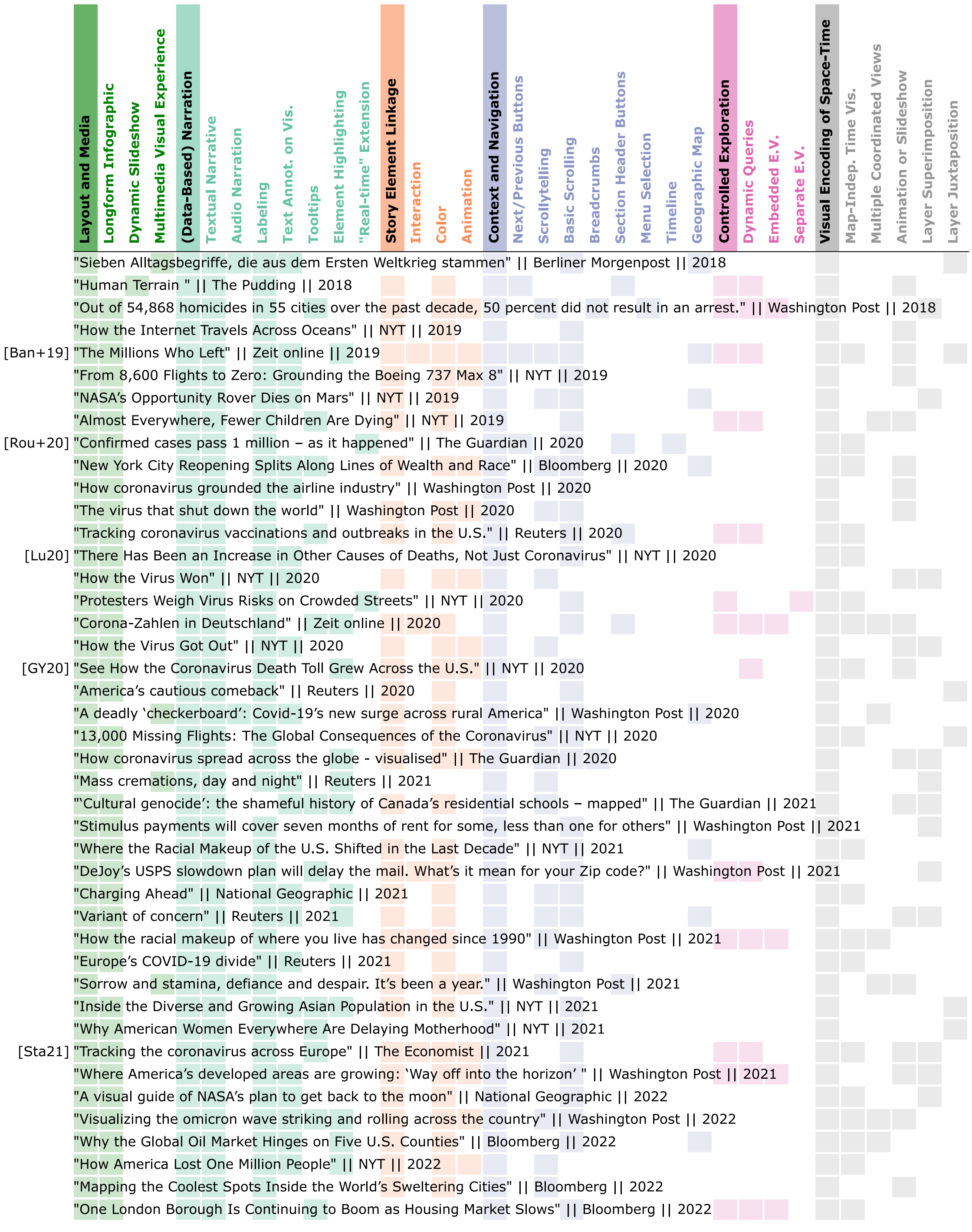


Figure 36: The full list of stories, part three of three. *(Figure adapted from Mayer et al. [May+23b], licensed under CC BY 4.0.)*

5.9 ADDENDUM

In the previous sections, we looked at spatio-temporal stories more broadly. While we also discussed the most relevant deviations that occurred for conflict stories, we still want to go into more detail for them in line with the overall scope of the dissertation.

Accordingly, this section is structured as follows. We first inspect more holistically how conflict stories differ from the other stories in the collection. Then, we focus the analysis only on the conflict stories. Lastly, we present a selection of more recent stories, reporting on the Israel-Hamas war, and discuss how they relate to our findings.

5.9.1 *Comparing Conflict Stories to the Remaining Collection*

In Section 5.5.1, we already presented the most remarkable deviations observed when comparing the distribution of techniques used in conflict stories to the overall average technique usage. In Figure 37, we provide a complete view of the deviations. We structure our remarks based on the six main technique categories, going from *layout and media* to *visual encoding of space-time*.

A considerably larger amount of **multimedia visual experiences** were produced for conflict stories. Potential reasons could be that the readers should be immersed more deeply in the stories, which often report on dramatic events, and that the readers' empathy with the victims should be raised in doing so.

The strongly increased use of **text annotations on visualizations** and **element highlighting** in conflict stories was already discussed, as the techniques can be used to convey information like territory gains and losses, or troop movements. In addition, **"real-time" extension** is more prominently used for advancing the *(data-based) narration*. This is reasonable, as several stories report on the long-lasting war against Ukraine, which is continuing to this day. In such cases, updating and extending a story over time can help to draw a more comprehensive image of the conflict and to clarify how individual events relate to it.

The *story element linkage* focused strongly on the use of **color**. **Interaction** was not used at all, and **animation** was also used less than average. These observations seem reasonable as, overall, means for interaction were barely present in conflict stories.

To provide *context and navigation*, conflict stories used only a selected subset of the available techniques. Accordingly, **next/previous buttons**, **breadcrumbs**, and **menu selection** were not used at all. In addition, **scrollytelling** was used substantially less than on average, keeping the focus on the straightforward technique of **basic scrolling**.

As already discussed, this straightforwardness also reflects in the fact that barely any techniques for *controlled exploration* were used. Particularly when reporting on sensitive and political topics such as conflicts, authors might be less in favor of leaving the conveyed information up to the reader's interpretation.

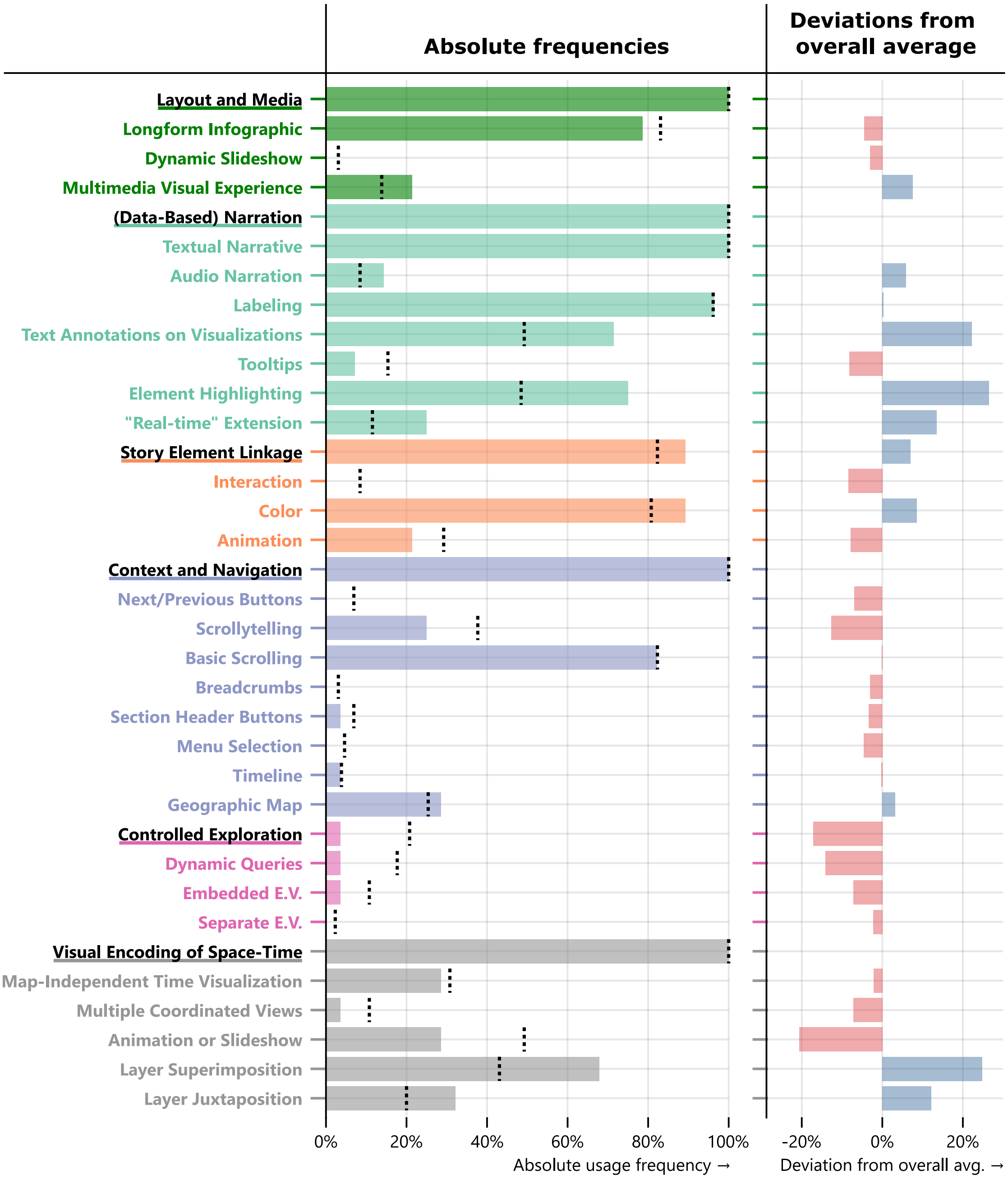


Figure 37: The bars in the center column display the conflict stories' average technique usage frequencies. As reference, the dotted vertical line in each row represents the overall average usage frequency. While this allows the deviations to be read off, we also provide the right column to depict them more clearly. In it, each bar represents the deviation of the conflict stories' average from the overall average.

The conflict stories used remarkably fewer **animations and slideshows** to *visually encode space-time*. The clear focus was on **layer superimposition** for providing comprehensive views of information like troop movements and territory gains, while also an increased amount of **layer juxtaposition** was used. **Multiple coordinated views**, which are more relevant in interactive and exploratory settings, were barely used.

5.9.2 *Further Analysis Focusing Only on Conflict Stories*

In this section, we focus only on the conflict stories. This also means that any average values represented via dotted lines in the figures only represent the average *among the conflict stories*, not among all stories. We first report on patterns found by analyzing the overall technique distribution, before presenting the results of agglomerative hierarchical clustering performed on the stories.

5.9.2.1 *Analyzing the Overall Technique Distribution*

We have split the techniques into three subsets, based on their usage frequency. We analyze each of the subsets in a different way, see also Figure 38:

- For the techniques which were used the most, it is not too insightful to investigate with which other techniques they were commonly combined. As almost all stories use these techniques, filtering out the stories which do not use them does not change the resulting distribution substantially. Therefore, we asked the inverse question: "If one of the highly frequent techniques is not used, what other techniques are used instead?"

- For each of the techniques which were less common but still relevant, we investigated how the technique distribution changes when considering only the stories which use the technique. This way, we answer the question: "How are the medium-frequent techniques combined?"

- For the techniques which were barely used and, therefore, not really relevant, we did not conduct any further analysis.

Below, we summarize our findings regarding the two listed questions. In the summary, we do not report on each individual technique, but only the most remarkable findings. This is also why we do not report on the usage of the overall technique categories, like *layout and media*, but only on the subcategories, like **longform infographics**, as they led to more interesting findings.

Note: Our interpretations of the findings and corresponding assumptions about the authors' design decisions behind them are only speculative and, accordingly, they are unlikely to represent the full spectrum of reasons that led to the design decisions.

"If one of the highly frequent techniques is not used, what other techniques are used instead?" To answer this question, for each of the most common techniques, we filtered out the stories using them. If stories do not use color to link story elements, they do not link them at all. Moreover, if basic scrolling is not used for navigation, scrollytelling is used instead, and stories that were not **longform infographics** were **multimedia visual experiences**.

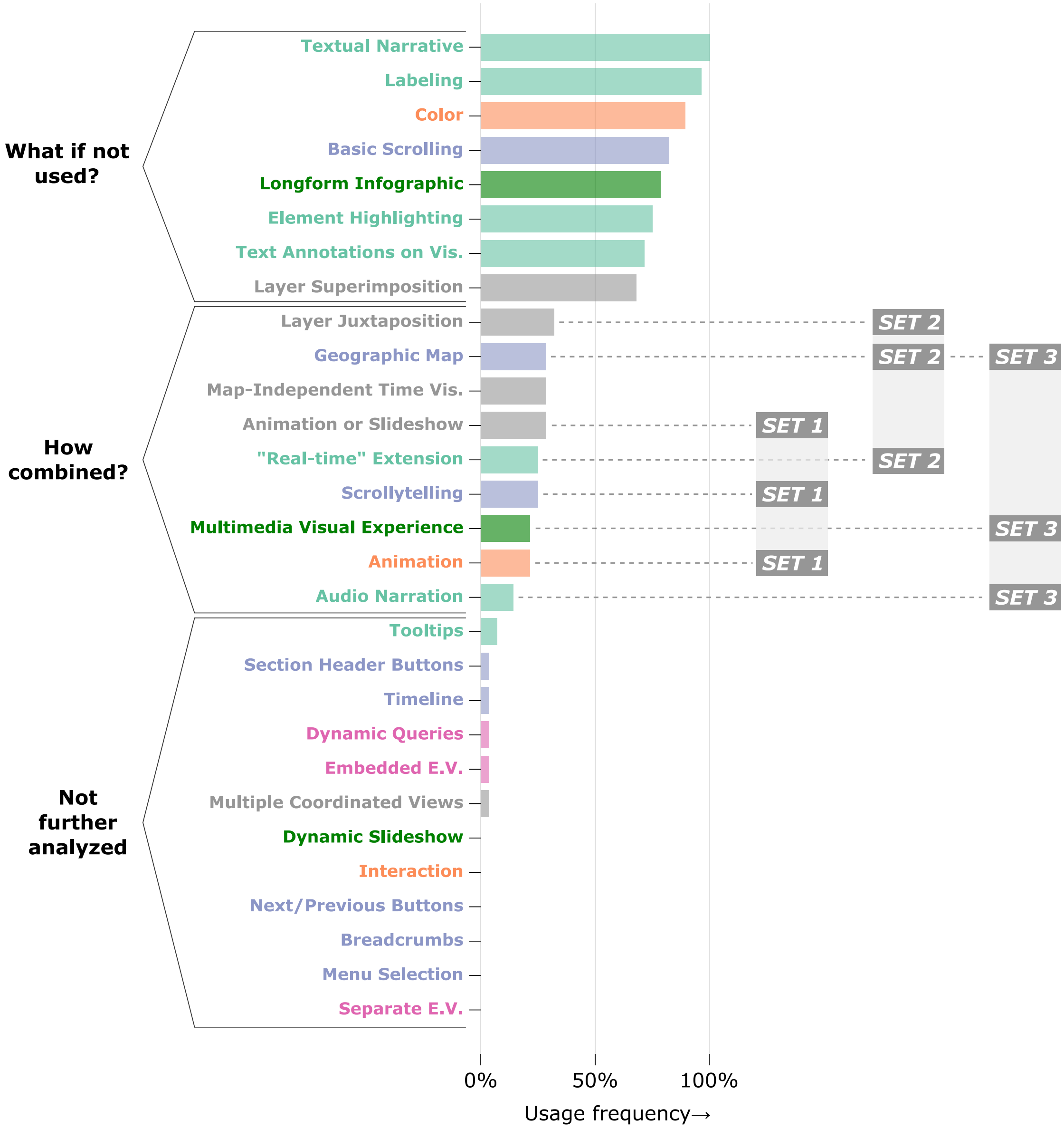
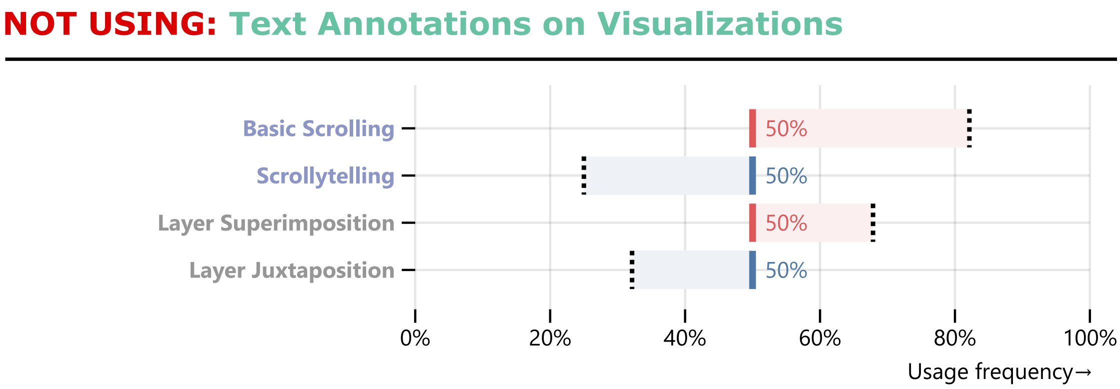


Figure 38: The techniques (excluding the main categories) sorted according to their usage frequency. The rectangles to the right represent three sets of techniques used to answer the question: "How are the medium-frequent techniques combined?"

When filtering for stories not using text annotations on visualizations, we found that the stories were more prone towards navigating via scrollytelling and using layer juxtaposition instead of layer superimposition for encoding space-time, see Figure 39. If multiple layers are juxtaposed, it is less important to annotate separate pieces of information on a single map, as it is the case in layer superimposition. Instead, different map views can be used for different pieces of information, reducing the need for text annotations in layer juxtaposition.

When filtering out stories using layer superimposition entirely, it also leads to a reduction of animations or slideshows for encoding space-time, see Figure 39. In line with the previous findings, layer juxtaposition and map-independent time visualization are used as replacement, causing text annotations on visualizations to become less relevant.



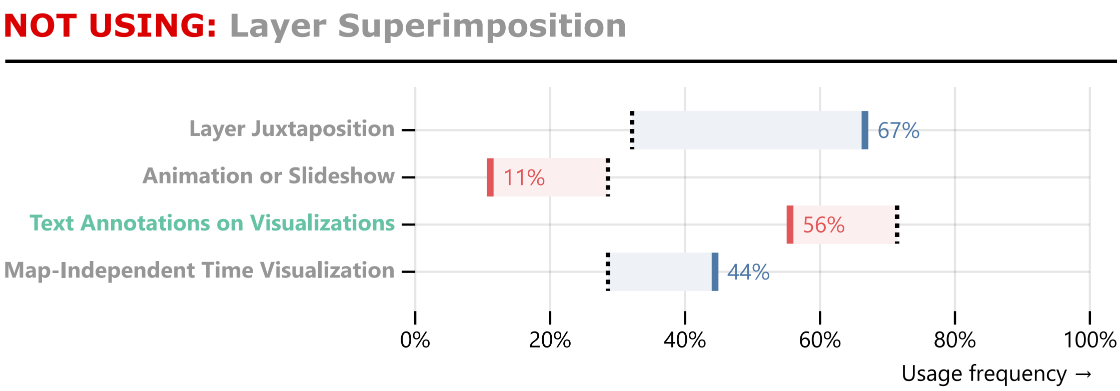


Figure 39: Remarkable deviations when filtering out stories using selected techniques. The deviations are in relation to the average across all conflict stories, which is represented as dotted lines. The rows are ordered based on the extent of the deviations, but the percentage annotations are the absolute values, not the deviations.

"How are the medium-frequent techniques combined?" Presenting visual summaries like in Figure 39 for each of the medium-frequent techniques would be quite extensive, and it would be demanding for the reader to extract the relevant key information from them. Instead, we aggregated our findings as follows.

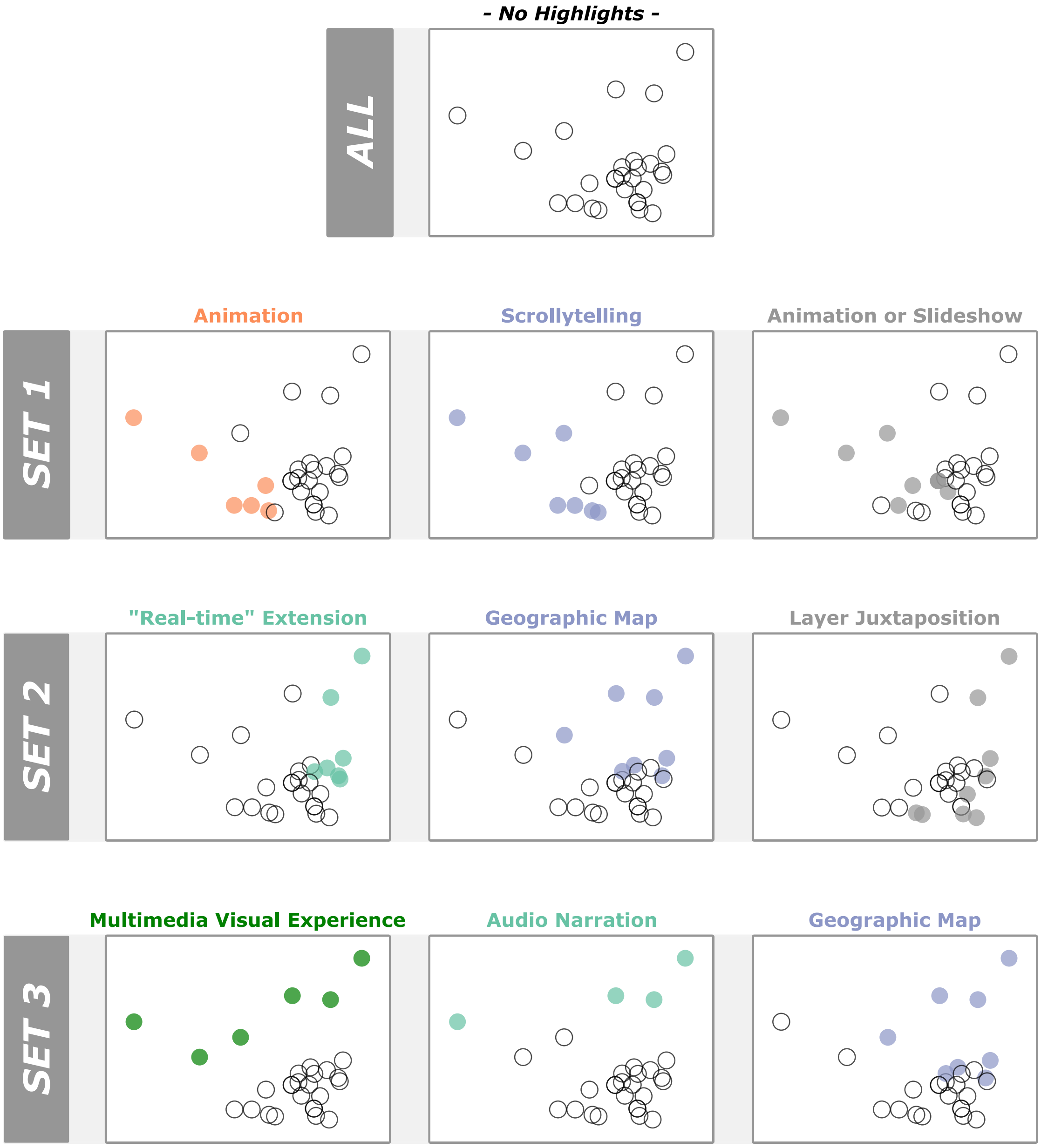


Figure 40: Three sets of techniques are visualized in a projected plane by highlighting which stories used the corresponding techniques. The closer two circles are, the more similar are the underlying stories. Circles with no fill color represent stories not using the corresponding technique.

We identified three sets, each consisting of three techniques that were commonly used together. In Figure 38, the composition of the three sets is indicated, and Figure 40 facilitates a more detailed visual comparison of their characteristics. To produce Figure 40, we first performed *multidimensional scaling* (MDS) [HPG13] only on the conflict stories, creating a 2D representation of the similarities between the stories, as in Section 5.5.2. The corresponding results are depicted in the topmost row in Figure 40, where each circle mark represents one story. Then, we created multiple versions of the projected view for each of the three sets. For each of the techniques used in a set, we highlighted all stories in the projected view that used the corresponding technique. This way, it becomes clear how different areas in the projection plane tend to correspond to different characteristics of the stories.

Set 1. The first set consists of animation-focused techniques, revolving around using **animations or slideshows** to visually encode space-time and **animation** to link separate story elements. Navigation is often supported via **scrollytelling**, which can be used to trigger animations.

Set 2. This set focuses on performing the narration via **"real-time" extensions**, providing context and navigation via **geographic maps**, and using **layer juxtaposition** to encode space-time. In the context of the more extensive "real-time" extensions reporting on multiple subsequent developments, geographic maps are rather used to provide context than to allow the user to navigate to specific parts of the story. In addition, layer juxtaposition tends not to be used to represent larger developments reported on across multiple updates of the story. Instead, it is rather used within a single update to display how events unfolded over the course of a few selected days.

Set 3. The third set revolves around techniques for making stories more immersive. It consists of **multimedia visual experiences** and **audio narration**. It slightly overlaps with *Set 2* as it also includes **geographic maps** for providing context, creating a deeper sense of place.

5.9.2.2 *Agglomerative Hierarchical Clustering of the Conflict Stories*

The projection performed using MDS inevitably introduces errors. Therefore, we also analyzed the similarity of the stories in the original space. To do so, we performed *agglomerative hierarchical clustering* (AHC) [DE84] using the *average linkage* criterion. We chose this approach based on the MDS results in Figure 40, which do not suggest a specific number of clusters, a certain density, or a particular cluster shape, like spherical or elongated clusters. Assuming such characteristics would have made other approaches more suitable.

The results of the clustering are displayed in Figure 41. We selected the threshold for building the clusters as an average distance between the stories of 0.14. We selected it with the goal of receiving a reasonable amount of clusters that have sufficient similarity between their elements while not being too asymmetrical in their size. In the lower

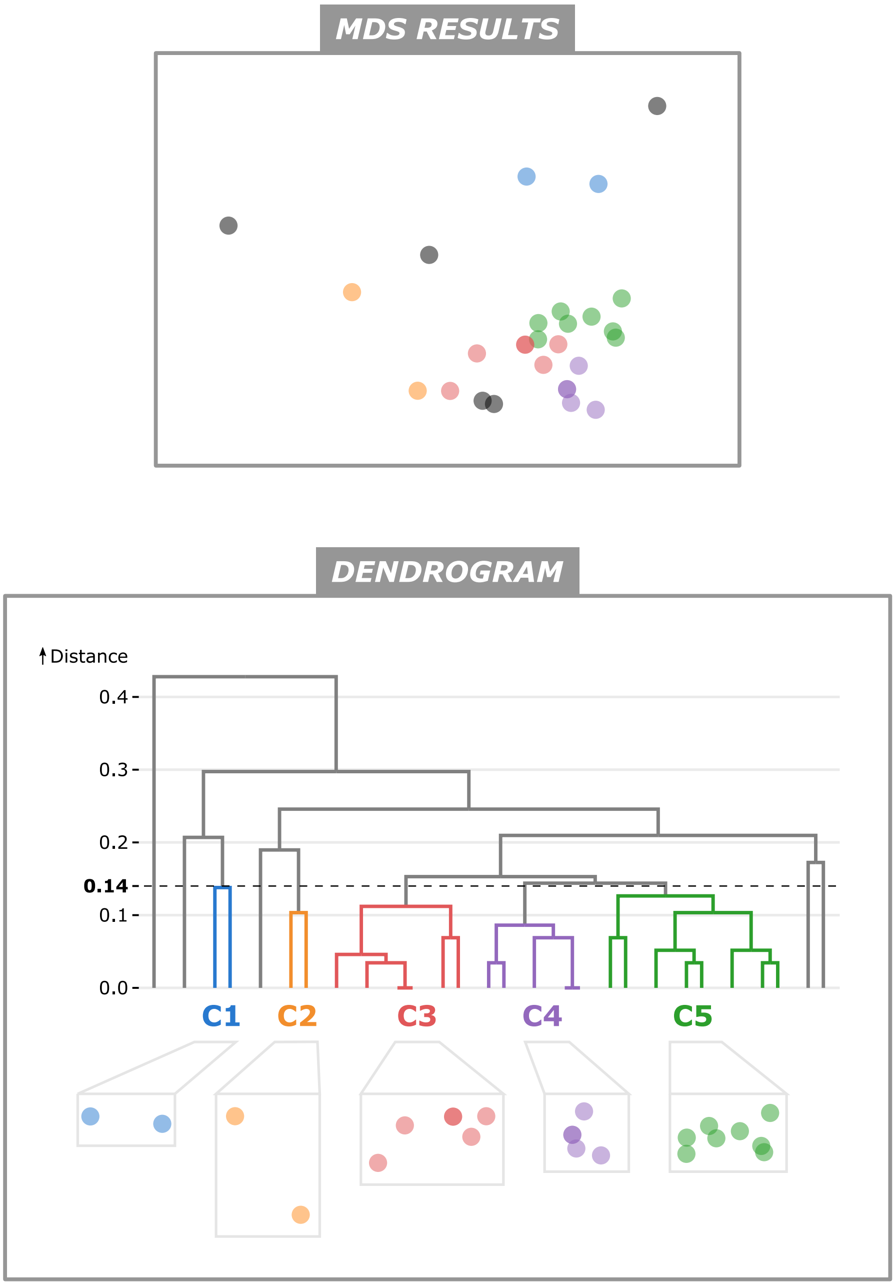


Figure 41: The AHC results with the identified clusters highlighted. The five edges of the dendrogram ending in gray lines correspond to outliers. In the MDS view, the corresponding marks have a dark gray fill color.

part of Figure 41, we highlight the five resulting clusters in a dendrogram view, and in the upper part, we display how the clusters are distributed across the MDS projection plane. The projected positions of the elements belonging to one cluster are repeated below the dendrogram to allow an easier identification of the clusters in the MDS view. With the selected threshold of 0.14, five outliers were identified that could not be allocated to any cluster. We characterize the clusters in more detail below, putting more focus on clusters **C3**-**C5** as they are considerably larger than clusters **C1** and **C2**.

Clusters C1 and C2. Both clusters consist of only two stories. The techniques used by the members of **C1** correspond closely to *Set 3* characterized in Figure 40, as they are both **multimedia visual experiences** using **audio narration** and **geographic maps** for providing context. In contrast, the stories from **C2** are related to *Set 1*. They both use **animation** to link separate story elements and **scrollytelling** instead of **basic scrolling** to provide navigation. However, only one of them uses **animations or slideshows** to visually encode space-time, while both make use of **layer superimposition**.

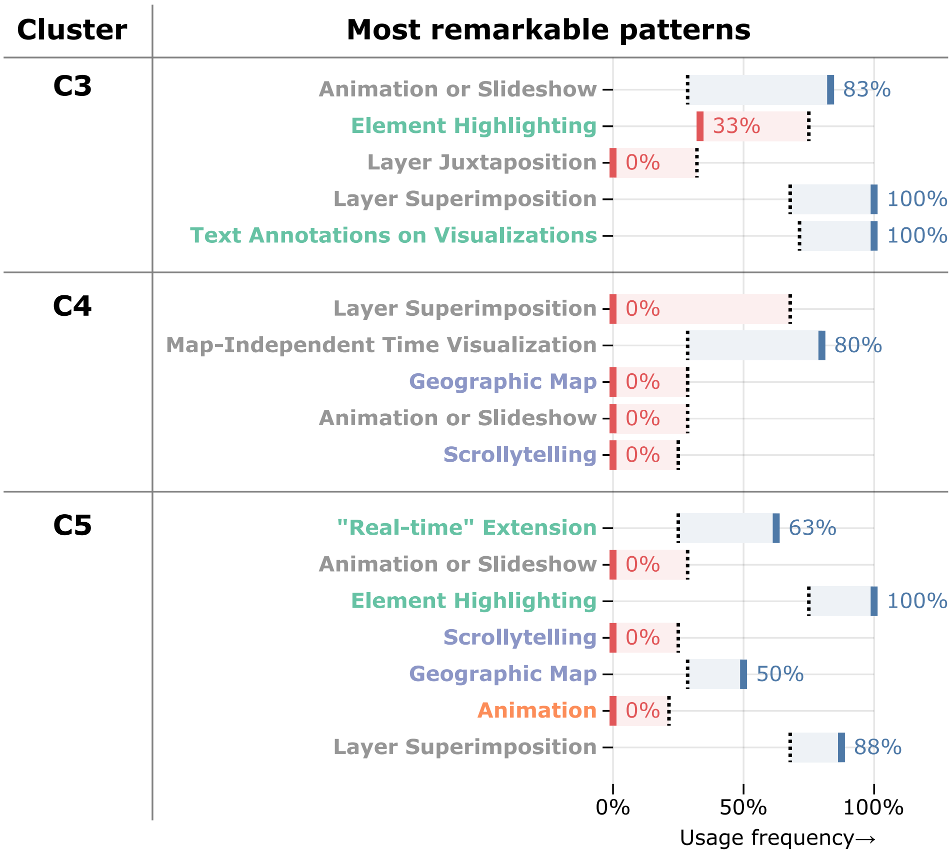


Figure 42: For each of the three larger clusters identified via AHC, the most remarkable deviations of technique usage are displayed, ordered by the extent of the deviation. As reference, the average across all conflict stories is displayed as dotted lines.

Cluster C3. This cluster consists of six stories (in Figure 41, two marks lie on top of each other). The most remarkable deviations of technique usage from the average across all conflict stories are depicted in Figure 42. Accordingly, the cluster is characterized by a frequent usage of **animations or slideshows** and **layer superimposition** to encode space-time. Techniques used a lot less by members of this cluster are **element highlighting** and **layer juxtaposition**. Particularly the strongly decreased use of **element highlighting** is surprising, as it is otherwise common for conflict stories and suitable for the two space-time encodings prominent in this cluster. However, Figure 42 suggests that it is compensated for by using **text annotations on visualizations** for communicating the narrative.

Cluster C4. This cluster consists of five stories (as for **C3**, two marks lie on top of each other in Figure 41). The stories focus strongly on using **map-independent time visualizations**, an otherwise rather unpopular technique. Other common options are not used at all, like **layer superimposition** or **animations or slideshows**. Moreover, the only technique the stories use for providing context and navigation is **basic scrolling**. Overall, these design decisions imply a more restricted story design. This is consistent with the fact that the stories are also rather short.

Cluster C5. The technique usage pattern of the eight stories contained in **C5** is quite complementary to *Set 1* characterized in Figure 40, relying on no means for animations or techniques supportive for it. Instead, the cluster is rather related to *Set 2*, as it contains stories using **"real-time" extensions** and **geographic maps** for providing context. However, in **C5**, **layer superimposition** is the strongly preferred option for encoding space-time, as compared to *Set 2*, where it is **layer juxtaposition**. The cluster also stands in contrast to **C3**, as all the stories in **C5** use **element highlighting** for communicating the narrative.

Outliers. Certain stories were too dissimilar from the others to be included in any cluster. The only conflict story providing *controlled exploration* is one such story, as well as a **multimedia visual experience** that uses several uncommon techniques for providing context and navigation, namely **section header buttons**, **timeline**, and **geographic map**.

To sum up the information from Section 5.9.2.1 and Section 5.9.2.2, several patterns can be identified in the collected conflict stories. In our investigations, analysis approaches coming from different directions led to similar insights, like identifying characteristics of animation-focused stories (*Set 1*, **C2**), of multimedia visual experiences (*Set 3*, **C1**), or of "real-time" extensions (*Set 2*, **C5**). They also revealed insights about more unexpected technique combinations (**C3**, **C4**).

However, it needs to be noted that different approaches may also lead to contradicting findings. For instance, in the MDS projection in Figure 41, there are two dark gray marks at the bottom of the projection plane representing outliers according to the AHC. However, in the projection, the two points are very close to each other, suggesting a high similarity between the corresponding stories. This contradicts the fact that they were identified as outliers in AHC, and it suggests that their proximity in the

MDS projection might be a projection error. This case is a good example of how one analysis technique alone may not suffice to provide a reliable picture of certain data.

5.9.3 *Characterizing More Recent Stories About the Israel-Hamas War*

To give an impression of what the most recent state in visual storytelling about conflicts is, we have collected four online stories reporting on the Israel-Hamas war to put them in relation to the findings from our analysis. We refer to the stories using abbreviations based on the source they were published by:

- "NYT" was published by *The New York Times* [Sta24],
- "REUT" was published by *Reuters* [Dut+23],
- "GUA1" was published by *The Guardian* [Swa+23], and
- "GUA2" was also published by *The Guardian* [Hoo+24].

We present our findings based on the overall technique categories. At that, we do not perform a detailed comparison of the technique usage ratios between the four newly collected stories and the 28 stories analyzed previously, as the new set of stories is too small for quantitative findings to be statistically significant. Rather, we compare the observed patterns qualitatively. To do so, it can still be beneficial to refer to the original technique usage distribution of the 28 stories, which is provided in Figure 37.

Regarding *layout and media*, the used techniques are similar to our previous observations. NYT and REUT are **longform infographics**, while GUA1 and GUA2 are **multimedia visual experiences**. Remarkably, both NYT and REUT also include videos, though they are not as central as in the other two stories.

To perform *(data-based) narration*, all stories use **textual narratives**, **text annotations on visualizations**, and **element highlighting**, similar to the 28 stories. The technique of **"real-time" extension** is strongly represented, as three of the four stories use it, namely NYT, REUT, and GUA1. They use it along with **labeling**, which is particularly helpful to not lose track of the narrative in longer stories, which stories using "real-time" extension tend to be. The fourth story, GUA2, does not use labeling. In general, GUA2 focuses on a more seamless experience, which we will touch on in the subsequent paragraphs. Techniques that are not used by any of the stories are **audio narration** and **tooltips**, which were also not too popular in our previous analysis.

To *link separate story elements*, all four stories use **color**. They do so primarily across multiple visualizations, using consistent coloring for annotations or area highlights corresponding to each other. Moreover, GUA2 uses **animation**. The story actually uses all techniques from the animation-focused *Set 1* described in Figure 40. As none of the stories uses **interaction** to link story elements, the technique usage in this category is also similar to that of the 28 stories.

To provide *context and navigation*, all stories use **basic scrolling** and **geographic maps**, the two techniques also most popular in the previous analysis. The similarity continues as none of the four stories uses **next/previous buttons**, **breadcrumbs**, or **menu selection**. In addition, none uses **timelines**, but Reut includes **section header buttons** at the beginning of the story to jump to specific sections. To facilitate its more seamless experience, Gua2 uses **scrollytelling** to navigate and trigger transitions.

Remarkably, two of the four stories use **dynamic queries** to provide *controlled exploration*. In Reut, it can be used to toggle which of three different event types is displayed on a map. In addition, Gua2 allows the interaction with satellite imagery, creating a striking effect. Upon hovering, the reader can compare what selected regions looked like before and after attacks destroyed vast parts of certain neighborhoods.

Lastly, the usage of techniques for *visually encoding space-time* also deviates slightly from the previous analysis. Accordingly, only two of the stories, Nyt and Reut, use **layer superimposition**. However, the three stories which use "real-time" extensions to communicate their narrative also use **layer juxtaposition** to encode space-time, which is in line with *Set 2* from the previous analysis, see Figure 40. As the three stories also use geographic maps for providing context, they actually align perfectly with *Set 2*. The corresponding stories are Nyt, Reut, and Gua1. In addition, **map-independent time visualizations** and **multiple coordinated views** are used by one story, Reut. Lastly, to round off Gua2's correspondence to *Set 1* from Figure 40, the story relies primarily on **animation or slideshow**, or, more precisely, only on animation.

The comparison performed above suggests that the techniques used for creating conflict stories have not changed substantially since 2022. Some patterns from the previous analysis could also be identified in the four new stories, particularly regarding the technique sets described in Figure 40. However, deviations from the previous analysis could also be observed. Most remarkably, it showed that exploratory techniques are still applied, despite their decreasing trend observed between 2018 and 2022. Another characteristic of the newer stories is that they make extensive use of satellite imagery augmented with annotations and area highlighting. It is mostly used to demonstrate the extent to which the infrastructure in certain areas was destroyed by fighting.

In the Addendum section, we provided a deeper look at the patterns that can be observed in the storytelling techniques used in conflict stories, and how the patterns compare to the complete collection of the 130 stories. This way, we also provided examples for directions that a more in-depth analysis of the complete collection can take, which we called for in the reflections in Section 5.6.

Overall, in this chapter, we identified patterns, trends, and inspirations for how visual storytelling can be performed in a spatio-temporal context, and how they are driven by large-scale struggles such as violent conflicts.

Part III

REFLECTIONS

In our work, we investigated how interactive data visualization can be used to support the exploration and explanation of spatio-temporal conflict data. In this part, we discuss our findings from the previous chapters, provide an outlook on potential future research directions, and conclude the work.

6 DISCUSSION AND FUTURE WORK

In this chapter, we address questions that may still be open, and questions that may have come up throughout this work, for instance, when different chapters led to seemingly contradictory findings. We do not repeat all the discussions from the individual chapters, but rather, we discuss the limitations and implications of this dissertation as a whole.

"Has the spectrum between exploration and explanation been fully explored?"

No. We moved along the spectrum between exploration and explanation to investigate how interactive visualization can support the work with conflict data. While we spread out the focal points of our studies across this spectrum, we could only investigate selected discrete points on it. Therefore, it would be interesting to also explore the remaining parts of the spectrum, e.g., by incorporating concepts from explanation in systems created for exploratory data analysis. While approaches like *guidance* [Cen+17] go in this direction, they do not cover how actual storytelling techniques could help to make such systems easier to understand and use for experts.

In addition, our project regarding exploratory explanation aimed at an expert audience. This direction seemed the most novel to us, because we found barely any investigations going in this direction even in application domains outside of conflict research. However, it would still be interesting to investigate how exploranation can help to communicate conflict research methods to broader audiences. While such approaches are better studied in other application domains, like medicine or astronomy, we did not find a lot of exploranation research targeted at broad audiences in the context of conflict research. Accordingly, one reviewer for our paper about expert exploranation expressed that non-expert audiences might also be interested in understanding methods like the one explained in the corresponding story, though, of course, on a higher level.

This is directly connected to the next point, the development of stories that can be used for audiences with differing background knowledge. Investigating how to make stories adaptable in this way could help to better justify the increased production costs for customized stories, as this adaptability could strongly broaden the range of potential audiences.

Another interesting aspect regarding explanation is to which extent data-driven stories about conflicts can and should include human actors, or even protagonists [Mit+23]. They can make conflict events and their consequences more relatable and raise the reader's empathy. However, such approaches could also be seen as too subjective or even manipulative. Yet, since these approaches are also common practice in journalism, they can be considered reasonable. To minimize the risk of a story being seen as manipulative, it is probably advisable to not artificially generate protagonists, which can be reasonable in other domains [Bud+23], but to represent actual human beings and their real-world stories.

In addition to these directions we have not explored, also the directions we have explored would benefit from further investigation. For instance, our projects targeted at experts in Chapter 3 and Chapter 4 were only individual application studies. Applying the concepts from these chapters also in other application-oriented projects would help to derive more generally applicable findings.

"How subjective are the findings?"

The evaluations conducted for the projects regarding EXPLORATION FOR EXPERTS and EXPLORATORY EXPLANATION FOR EXPERTS both only had a limited number of participants (five and eight). This makes the findings not as reliable as larger studies, since the opinions of individual participants can skew the overall outcome more easily. However, these shortcomings are difficult to avoid when creating applications specifically for domain experts, as the pool of potential evaluation participants is strongly limited, as well as the participants' availability.

Moreover, the classifications of the 130 online stories regarding the visual storytelling techniques in Chapter 5 are inherently subjective, even though we tried to limit the subjectivity by including three visualization researchers in the coding process and by reporting on the border cases. Even before the classification, the story selection introduced subjectivity regarding both the sources we considered and the actual decisions as to which stories to include. However, we also tried to provide transparency in this regard, listing the sources we considered and the search procedures we followed, along with explanations regarding the minimum requirements the stories had to fulfill to be included in our collection.

In addition, we primarily collaborated with one conflict researcher. While we also received feedback from other researchers in our evaluations, the main directions of the studies we conducted were developed together with our main collaboration partner. Accordingly, they were shaped based on his research directions and his perspective on the field. Collaborating with other experts can offer different perspectives, particularly in a somewhat divided field like conflict research, where some experts focus more on quantitative research while others focus more on qualitative research. This is an important step for future research.

"How does Visualization research benefit from the application-oriented studies?"

We address this (justified) question often also asked by reviewers as follows. The study in Chapter 3 and, to an extent, also the study in Chapter 4 focused on creating applications to solve problems from a domain other than visualization research. Yet, visualization research can still benefit from such application-oriented projects. Sedlmair et al. identified the contributions that such projects can have [SMM12]. We aimed to meet them by

- characterizing the domain, the problems, and the corresponding data,
- abstracting them into the vocabulary of data visualization,
- deriving requirements to judge a proposed design against,
- explaining and justifying a corresponding application design,
- evaluating the design with domain experts, and
- reflecting on the process and the findings.

This way, we generated transferable knowledge on how data visualization theories can be applied to solve actual domain problems.

"Should interaction be used for explanation?"

In Chapter 4, we used interactive means to support the explanation of a conflict research method to domain scientists, and most of the means were perceived and adapted well. However, in Chapter 5, we found that, over the recent years, less exploration has been used when communicating data-based insights to broad audiences. Instead, the trend went towards a simplified consumability of the stories. This raises the question whether interaction should be used in the context of explanation.

As expressed before, there still exist situations in which exploration can be beneficial, e.g., for creating a familiar setting for the reader [Bac+18]. However, such reasons do not explain why the interactivity was perceived so well by the domain experts in Chapter 4. Instead, we think that a key difference between expert audiences and broad audiences can explain this observation: Domain experts have a strong incentive to understand a method they are considering to base their own research on, and interactivity can help them deepen their understanding. In contrast, for the average reader of an online story, understanding the background of the communicated information in such detail is usually not the goal.

However, even the experts had a preference regarding the interaction techniques, namely, that they should provide new insights while not being too complex. This wish for new insights can be seen as contradictory to the guideline we refer to in Section 2.7, stating that no important information should be hidden behind interaction [Ais23]. To bring these pieces of advice in line, we refine the guideline to more precisely express that no information should be hidden behind interaction *that is necessary for a correct understanding of the core message of the story*. This is also more in line with Aisch's work providing the original guideline [Ais23].

"Should dynamic slideshows be used for expert exploranation?"

Another observation that seems contradictory is that, in Chapter 4, we used a dynamic slideshow layout for the exploranation story, including previous and next buttons for navigation as well as breadcrumbs. At the same time, in Chapter 5, we found that barely any online stories used this type of layout and navigation. This is a good example to show that patterns observed in one area should not automatically be assumed to hold for a related area. If there are reasons for diverging from the observed patterns, adjusting the design accordingly might still be reasonable.

For instance, one reason why so few online stories used slideshows layouts could be that slideshows are less similar to traditional newspaper layouts, including the fact that they cannot be easily skimmed. However, many of the stories were published by newspaper companies, so the longform infographic layout more suitable for skimming might be the preferred option.

In contrast, our exploranation story was created to explain the complex workings of a scientific method. In this situation, allowing readers easy skimming of the story was not our main interest. Instead, supporting understandability and an unambiguous correspondence between text and visualization was more important, justifying a slideshow layout. Nevertheless, this shows that more investigations are necessary to determine more precisely which storytelling design decisions are appropriate under what circumstances - at least, when aiming for the automation of the story creation process, which we discuss in the following.

"How to make the story creation process more efficient?"

A common goal in the field of computer science is the automation of processes. The same holds for the process of creating visual data stories. Corresponding story authoring tools [Che+23] are also important for democratizing the transfer of data and knowledge, allowing users with limited to no coding knowledge to create and share stories. More often than not, it is the people from backgrounds other than computer science who have the most interesting stories to tell. Supporting them in doing so is an important task in visual storytelling research.

Our work contributes to solving this task as, in Chapter 5, we have identified several techniques that are commonly used to create stories with a spatio-temporal context and would, accordingly, also be relevant to be included in corresponding authoring tools. Still, our results only scratch the surface of all the characteristics that visual data stories can encompass. Therefore, more research is necessary to determine which functionalities authoring tools should provide, and how to provide these functionalities in an accessible way.

In addition to supporting users with limited coding skills through automation, research on how to simplify the coding-based creation of stories would also be beneficial. Corresponding coding frameworks could open the door to a more economical creation of unique stories, as more customization can be provided inside frameworks than inside fully automated authoring tools. When consuming stories in our everyday life, e.g., in movies or books, usually, the stories that do things a little differently are the ones that get stuck in our head the most. Adding such peculiarities to a data-driven story is easier the more flexible the environment of creation is. Eventually, this boils down to the familiar trade-off between a tool's flexibility on the one hand, and its ease of use and learning effort on the other hand, with libraries like D3.js and chart creation platforms like Tableau being typical examples in the more general field of visualization design.

When creating our exploranation story for Chapter 4, we searched for suggestions or frameworks on how to economically implement the story, while facilitating the inclusion of D3.js to provide sufficient customization. However, besides some very basic scrollytelling frameworks that did not meet our needs, we did not find a lot of information. Accordingly, the creation of capable frameworks is another important direction for future research. We have already worked towards this goal by extracting the framework we had developed for the creation of our exploranation story, see Section 4.8. In addition, story templates [Meu+22] can be beneficial for facilitating the creation of stories, however, this approach also limits the customizability.

Despite the various suggestions made above, we want to highlight that we have not conducted any research focusing explicitly on authoring tools. Accordingly, the suggestions are merely based on our own experience, while other researchers certainly have more substantiated expertise regarding such automation [Che+23].

"Why do stories about COVID-19 provide exploration while stories about conflicts do not?"

In Chapter 5, we found that the conflict stories in our collection provided barely any exploration. As a possible explanation, we stated that authors of such political and sensitive topics might want to reduce the risk of readers with limited background knowledge misinterpreting the displayed information. A lot of the stories about COVID-19 show similar characteristics, covering political and sensitive topics. Yet, these stories provided a considerable amount of exploration, even though data about COVID-19 has also been misinterpreted on several occasions. Accordingly, other differences between the two kinds of stories must exist. We assume that a driving factor is how much desire for interaction the authors assume their readers to have. As mentioned, readers are more interested in additional exploration the more relevant a topic is for them, and a pandemic and its development have a very immediate impact on most readers, while violent conflicts often do not.

"Will conflict research return to focus more on interstate wars?"

In Section 2.1, we stated that, over the recent decades, conflict research has shifted its focus primarily on the analysis of intrastate wars. However, with the war against Ukraine and the Israel-Hamas war, the question can be raised whether these developments will cause conflict research to shift back some of its focus to interstate wars. To answer this question, it needs to be noted that the Hamas are not officially accepted representatives of a state. While the conflict is internationalized to a certain degree, as Lebanese and Yemeni militias participate in it, it is still an intrastate war of a state against a terrorist group. In addition, other large-scale intrastate wars continue to be studied, like the Tigray war [DPO23]. Moreover, the invasion of Ukraine has also not started as a typical interstate war, but it has developed to be considered and analyzed as such in conflict research literature [DPO23].

To summarize, new research is conducted on interstate wars, but as intrastate wars continue to exist, they remain relevant in conflict research. Data visualization can help to make such different kinds of conflicts understandable to both researchers and broad audiences. Therefore, the described developments are also important to be followed from a visualization perspective, as they can influence which forms of visual communication are most appropriate. In this context, meta-studies would also be beneficial, analyzing to which extent data visualization is used as an actual method in the field of conflict research, and not just as a means for presenting results in a publication.

7 CONCLUSION

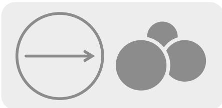
In this dissertation, we investigated how interactive data visualization can be used to support the work with spatio-temporal conflict data. To do so, we conducted three main studies along the spectrum between exploration and explanation.

In the study focusing on exploration, we developed an application for performing exploratory analysis of a conflict event data set. The data set was created using an integration method for merging separate data sets into one holistic set while sorting out duplicate recordings reporting on the same original incident. The goal of the application was to explore the integrated data set to validate the integration results. To design the application, we analyzed the underlying domain problems, abstracted them into analytical tasks, and presented a visualization design to solve the tasks. We evaluated the application with five domain experts. It showed that they could use it to get a deeper understanding of how the integration method's parameters influence the integration results. The application also helped the experts assess whether the number and structure of the identified duplicate event recordings was reasonable.

Overall, the study showed that conflict research can benefit from exploratory visualization, even though corresponding applications have barely been used in the field, so far. However, it also revealed that a considerable amount of explanation is necessary for allowing domain experts to properly use such applications.

We focused on combining explanation and controlled exploration in our second study. The goal was to explain the workings of a conflict research method to other researchers, using techniques from data-driven visual storytelling. We reported on the corresponding process, which we referred to as *expert exploranation* (ExEx), discussing its goals and challenges, how they deviate from exploranation targeted at broad audiences, and explaining the individual steps of the ExEx process. We evaluated the resulting interactive visual data story with eight conflict researchers. Half of the participants had already used the explained method before, while the other half had not, allowing us to collect feedback from different perspectives.

We reflected on the design process and the evaluation, deriving guidelines for performing expert exploranation. They include that, when explaining a method in the form of an interactive story, no knowledge from the paper should be assumed that originally introduced the method, and that it is beneficial to invent new visualizations as compared to the figures used in the original paper. Moreover, it showed that interactivity and clear textual interpretations help experts understand the explained method more deeply, and that providing a time-efficient version of the story can make it more attractive for users already familiar with some of the explained concepts. Notably, we found that even experts who had already used the explained method before can gain new insights from an ExEx story, as long as they are open for it.



To get a deeper understanding of how visual storytelling can be performed in a spatio-temporal context, in our third study, we collected and analyzed 130 online stories. To classify which storytelling techniques they used, we merged and adapted three existing design spaces. We analyzed the stories' distribution in the resulting design space, deriving common patterns of how the storytelling techniques were used, illustrating how creatively they were combined by the story authors. We also found that large-scale struggles, like violent conflicts, shape the landscape of storytelling techniques, even leading to the creation of new techniques. Furthermore, it showed that the stories in our collection had become more straightforward to consume over the years, indicated by a reduced use of interaction- and exploration-focused techniques.

In a separate analysis, focusing explicitly on the stories reporting on conflicts, we identified more specific groups of stories and patterns of technique combinations. They gave insights into the different approaches that exist for reporting on violent conflicts using data-driven storytelling.

Overall, our results provide several directions for how interactive visualization can support the exploration and explanation of spatio-temporal conflict data. However, additional investigations are necessary to reduce potential subjectivity of the results, and to further generalize the findings. In addition, more research along the spectrum between exploration and explanation should be performed.

Accordingly, it would be interesting to study how applications for exploratory analysis can be made more accessible by incorporating explanatory techniques from visual storytelling. In addition, it is crucial to investigate how the process of creating spatio-temporal data stories can be made more efficient. This could be achieved by using automation, dedicated frameworks, or by providing a way to create multiple version of the same story for different audiences. Moreover, the role of interaction in explanatory applications should be investigated more deeply.

Part IV

APPENDIX

We end the dissertation by providing the list of references and a list of abbreviations used across the work.

BIBLIOGRAPHY

[Acl] *ACLED Dashboard*. https://acleddata.com/dashboard/dashboard. Accessed: 2023-12-14. ongoing.

[Sig] *Afghanistan War (2002-2014) Significant Activities "SIGACTs"*. https://www.politicalviolencelab.com/sigacts-afghanistan. Accessed: 2024-01-31. 2014.

[Aig+23] W. Aigner, S. Miksch, H. Schumann, and C. Tominski. *Visualization of Time-Oriented Data*. Second Edition. Springer, 2023. DOI: 10.1007/978-1-4471-7527-8. URL: https://timeviz.net.

[Ais23] G. Aisch. *In Defense of Interactive Graphics*. https://www.vis4.net/blog/in-defense-of-interactive-graphics/. Accessed: 2023-10-23. 2023.

[All+20] J. Allen, S. Almukhtar, A. Aufrichtig, A. Barnard, M. Bloch, P. Bullock, et al. *Coronavirus in the U.S.: Latest Map and Case Count*. The New York Times. https://www.nytimes.com/interactive/2021/us/covid-cases.html. 2020.

[Ami+15] F. Amini, N. Henry Riche, B. Lee, C. Hurter, and P. Irani. "Understanding Data Videos: Looking at Narrative Visualization through the Cinematography Lens." In: *Proc. of ACM Conference on Human Factors in Computing Systems*. ACM, 2015, 1459–1468. DOI: 10.1145/2702123.2702431.

[Ami+18] F. Amini, N. H. Riche, B. Lee, J. Leboe-McGowan, and P. Irani. "Hooked on Data Videos: Assessing the Effect of Animation and Pictographs on Viewer Engagement." In: *Proc. of International Conference on Advanced Visual Interfaces*. ACM, 2018. DOI: 10.1145/3206505.3206552.

[And+07] P. André, M. L. Wilson, A. Russell, D. A. Smith, A. Owens, and M. Schraefel. "Continuum: designing timelines for hierarchies, relationships and scale." In: *Proc. of ACM Symposium on User Interface and Software Technology*. ACM, 2007, pp. 101–110.

[And+17] G. Andrienko, N. Andrienko, G. Fuchs, and J. Wood. "Revealing Patterns and Trends of Mass Mobility Through Spatial and Temporal Abstraction of Origin-Destination Movement Data." In: *IEEE Transactions on Visualization and Computer Graphics* 23.9 (2017), pp. 2120–2136.

[And+17] G. Andrienko, N. Andrienko, W. Chen, R. Maciejewski, and Y. Zhao. "Visual Analytics of Mobility and Transportation: State of the Art and Further Research Directions." In: *IEEE Transactions on Intelligent Transportation Systems* 18.8 (2017), pp. 2232–2249. DOI: 10.1109/TITS.2017.2683539.

[AA06] N. Andrienko and G. Andrienko. *Exploratory analysis of spatial and temporal data: a systematic approach*. Springer Science & Business Media, 2006.

[AA13] N. Andrienko and G. Andrienko. "Visual analytics of movement: An overview of methods, tools and procedures." In: *Information Visualization* 12.1 (2013), pp. 3–24.

[AAG03] N. Andrienko, G. Andrienko, and P. Gatalsky. "Exploratory spatio-temporal visualization: an analytical review." In: *Journal of Visual Languages & Computing* 14.6 (2003), pp. 503–541. DOI: 10.1016/S1045-926X(03)00046-6.

[AH+11] R. Arias-Hernandez, L. T. Kaastra, T. M. Green, and B. Fisher. "Pair Analytics: Capturing Reasoning Processes in Collaborative Visual Analytics." In: *Proc. of 44th Hawaii International Conference on System Sciences*. 2011, pp. 1–10. DOI: 10.1109/HICSS.2011.339.

[Bac+17] B. Bach, P. Dragicevic, D. Archambault, C. Hurter, and S. Carpendale. "A descriptive framework for temporal data visualizations based on generalized space-time cubes." In: *Computer Graphics Forum*. Vol. 36. 6. 2017, pp. 36–61. DOI: 10.1111/cgf.12804.

[Bac+18] B. Bach, M. Stefaner, J. Boy, S. Drucker, L. Bartram, J. Wood, et al. "Narrative design patterns for data-driven storytelling." In: *Data-driven storytelling*. AK Peters/CRC Press, 2018, pp. 107–133. DOI: 10.1201/9781315281575-5.

[BD05] M. Balzer and O. Deussen. "Voronoi treemaps." In: *IEEE Symposium on Information Visualization*. 2005, pp. 49–56.

[Ban+19] C. Bangel, P. Blickle, E. Erdmann, P. Faigle, A. Loos, J. Stahnke, J. Tröger, and S. Venohr. *The Millions Who Left*. Zeit Online. https://www.zeit.de/politik/deutschland/2019-05/east-west-exodus-migration-east-germany-demography. 2019.

[Bat+10] S. Bateman, R. L. Mandryk, C. Gutwin, A. Genest, D. McDine, and C. Brooks. "Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts." In: *Proc. of ACM Conference on Human Factors in Computing Systems*. ACM, 2010, 2573–2582. DOI: 10.1145/1753326.1753716.

[Bey+20] J. Beyer et al. "Case Studies for Working with Domain Experts." In: *Foundations of Data Visualization*. Ed. by M. Chen, H. Hauser, P. Rheingans, and G. Scheuermann. Springer International Publishing, 2020, pp. 255–278. DOI: 10.1007/978-3-030-34444-3_13.

[BBY03] V. C. Bhavsar, H. Boley, and L. Yang. "A Weighted-Tree Similarity Algorithm for Multi-Agent Systems in e-Business Environments." In: *Computational Intelligence*. 2003, pp. 53–72.

[Bil05] P. Bille. "A survey on tree edit distance and related problems." In: *Theoretical Computer Science* 337.1-3 (2005), pp. 217–239.

[Bir21] K. Biriukov. "Storytelling Maps Classification." In: *Culminating Projects in Geography and Planning* 10 (2021). https://repository.stcloudstate.edu/gp_etds/10/.

[Blo56] B. S. Bloom. *Taxonomy of educational objectives: The classification of educational goals*. Longman, 1956.

[Boc+18] A. Bock, E. Axelsson, C. Emmart, M. Kuznetsova, C. Hansen, and A. Ynnerman. "OpenSpace: Changing the Narrative of Public Dissemination in Astronomical Visualization from What to How." In: *IEEE Computer Graphics and Applications* 38.3 (2018), pp. 44–57. DOI: 10.1109/MCG.2018.032421653.

[BSJ19] T. P. Bojan Savric and B. Jenny. "The Equal Earth map projection." In: *International Journal of Geographical Information Science* 33.3 (2019), pp. 454–465. DOI: 10.1080/13658816.2018.1504949.

[Böt+20] M. Böttinger, H.-N. Kostis, M. Velez-Rojas, P. Rheingans, and A. Ynnerman. "Reflections on Visualization for Broad Audiences." In: *Foundations of Data Visualization*. Ed. by M. Chen, H. Hauser, P. Rheingans, and G. Scheuermann. Springer International Publishing, 2020, pp. 297–305. DOI: 10.1007/978-3-030-34444-3_16.

[BPP21] B. M. Brad Plumer and N. Popovich. *How Much Are Countries Pledging to Reduce Emissions?* The New York Times. https://www.nytimes.com/interactive/2021/11/01/climate/paris-pledges-tracker-cop-26.html. 2021.

[Bre+16] M. Brehmer, B. Lee, B. Bach, N. H. Riche, and T. Munzner. "Timelines revisited: A design space and considerations for expressive storytelling." In: *IEEE Transactions on Visualization and Computer Graphics* 23.9 (2016), pp. 2151–2164. DOI: 10.1109/TVCG.2016.2614803.

[Bro+23] M. Brossier, R. Skånberg, L. Besançon, M. Linares, T. Isenberg, A. Ynnerman, and A. Bock. "Moliverse: Contextually embedding the microcosm into the universe." In: *Computers & Graphics* 112 (2023), pp. 22–30. DOI: 10.1016/j.cag.2023.02.006.

[BB13] W. Bruine de Bruin and A. Bostrom. "Assessing what to address in science communication." In: *Proc. of National Academy of Sciences* 110.supplement_3 (2013), pp. 14062–14068. DOI: 10.1073/pnas.1212729110.

[BCH07] C. Brunsdon, J. Corcoran, and G. Higgs. "Visualising space and time in crime patterns: A comparison of methods." In: *Computers, Environment and Urban Systems* 31.1 (2007), pp. 52–75. DOI: 10.1016/j.compenvurbsys.2005.07.009.

[Bud+23] B. Budich, L. A. Garrison, B. Preim, and M. Meuschke. "Reflections on AI-Assisted Character Design for Data-Driven Medical Stories." In: *Eurographics Workshop on Visual Computing for Biology and Medicine*. The Eurographics Association, 2023. DOI: 10.2312/vcbm.20231216.

[BW14] M. Burch and D. Weiskopf. "On the Benefits and Drawbacks of Radial Diagrams." In: *Handbook of Human Centric Visualization*. Ed. by W. Huang. Springer, 2014, pp. 429–451.

[Bur+20] A. Burns, C. Xiong, S. Franconeri, A. Cairo, and N. Mahyar. "How to evaluate data visualizations across different levels of understanding." In: *IEEE Proc. of Workshop on Evaluation and Beyond - Methodological Approaches to Visualization*. 2020, pp. 19–28. DOI: 10.1109/BELIV51497.2020.00010.

[CCF95] M. S. T. Carpendale, D. J. Cowperthwaite, and F. D. Fracchia. "3-dimensional pliable surfaces: For the effective presentation of visual information." In: *Proc. of ACM Symposium on User Interface and Software Technology*. ACM, 1995, pp. 217–226.

[CM07] N. Cawthon and A. V. Moere. "The Effect of Aesthetic on the Usability of Data Visualization." In: *Proc. of International Conference Information Visualisation (IV)*. 2007, pp. 637–648. DOI: 10.1109/IV.2007.147.

[Cen+17] D. Ceneda, T. Gschwandtner, T. May, S. Miksch, H.-J. Schulz, M. Streit, and C. Tominski. "Characterizing Guidance in Visual Analytics." In: *IEEE Transactions on Visualization and Computer Graphics* 23.1 (2017), pp. 111–120. DOI: 10.1109/TVCG.2016.2598468.

[Che+23] Q. Chen, S. Cao, J. Wang, and N. Cao. "How Does Automation Shape the Process of Narrative Visualization: A Survey of Tools." In: *IEEE Transactions on Visualization and Computer Graphics* (2023), pp. 1–20. DOI: 10.1109/TVCG.2023.3261320.

[Obsa] *Collection of Observable notebooks*. https://observablehq.com/collection/@zykel/exex-prototypes. Accessed: 24 October 2023. 2023.

[Cop] *Copernicus Data Space Ecosystem*. https://dataspace.copernicus.eu/. Accessed: 2023-11-10.

[CA+20] V. J. Cortes 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." In: *Sustainability Science* 15 (2020), pp. 1013–1032. DOI: 10.1007/s11625-020-00793-y.

[Gsaa] *Creating an Anomaly Heatmap*. https://vinccenttt.github.io/anomaly-heatmap-aseq/. Accessed: 2024-02-15. 2023.

[Cui19] W. Cui. "Visual Analytics: A Comprehensive Overview." In: *IEEE Access* 7 (2019), pp. 81555–81573. DOI: 10.1109/ACCESS.2019.2923736.

[D3] *D3*. https://d3js.org/. Accessed: 2023-10-23. 2023.

[DPO23] S. Davies, T. Pettersson, and M. Öberg. "Organized violence 1989–2022, and the return of conflict between states." In: *Journal of Peace Research* 60.4 (2023), pp. 691–708. DOI: 10.1177/00223433231185169.

[DE84] W. H. E. Day and H. Edelsbrunner. "Efficient algorithms for agglomerative hierarchical clustering methods." In: *Journal of Classification* 1.1 (1984), 7–24. DOI: 10.1007/bf01890115.

[DP20] E. Dimara and C. Perin. "What is Interaction for Data Visualization?" In: *IEEE Transactions on Visualization and Computer Graphics* 26.1 (2020), pp. 119–129. DOI: 10.1109/TVCG.2019.2934283.

[Dis] *Distill*. https://distill.pub/. Accessed: 2023-10-23. 2023.

[DN21] S. Dodge and E. Noi. "Mapping trajectories and flows: facilitating a human-centered approach to movement data analytics." In: *Cartography and Geographic Information Science* 48.4 (2021), pp. 353–375. DOI: 10.1080/15230406.2021.1913763.

[Don+19] K. Donnay, E. T. Dunford, E. C. McGrath, D. Backer, and D. E. Cunningham. "Integrating Conflict Event Data." In: *Journal of Conflict Resolution* 63.5 (2019), pp. 1337–1364.

[DF14] K. Donnay and V. Filimonov. "Views to a war: Systematic differences in media and military reporting of the war in iraq." In: *EPJ Data Science* 3 (2014), p. 25.

[DGB14] K. Donnay, E. Gadjanova, and R. Bhavnani. "Disaggregating Conflict by Actors, Time, and Location." In: *Peace and Conflict*. Ed. by D. A. Backer, J. Wilkenfeld, and P. K. Huth. Paradigm, 2014, pp. 44–56.

[Dut+23] P. K. Dutta, J. McClure, M. Zafra, A. Bhandari, J. Saul, C. Trainor, M. Zafra, and D. Bankova. *Mapping the conflict in Israel and Gaza*. Reuters. https://www.reuters.com/graphics/ISRAEL-PALESTINIANS/MAPS/movajdladpa/. 2023.

[Dzw+05] W. Dzwinel, D. A. Yuen, K. Boryczko, Y. Ben-Zion, S. Yoshioka, and T. Ito. "Nonlinear multidimensional scaling and visualization of earthquake clusters over space, time and feature space." In: *Nonlinear Processes in Geophysics* 12.1 (2005), pp. 117–128. DOI: 10.5194/npg-12-117-2005.

[Econa] EcoWest. *Tracking U.S. Drought Severity*. http://vis.ecowest.org/interactive/drought.php. Accessed: 2022-09-20. n.a.

[ED07] G. Ellis and A. Dix. "A Taxonomy of Clutter Reduction for Information Visualisation." In: *IEEE Transactions on Visualization and Computer Graphics* 13.6 (2007), pp. 1216–1223.

[Eun22] Eunice Au & Connected Action. "Data Journalism Top 10: Protesting the War, Russian Sanctions, Mapping Refugees, and NICAR Tips." In: (2022). https://gijn.org/2022/03/11/data-journalism-top-10-protesting-the-war-russian-sanctions-mapping-refugees-and-nicar-tips/. Accessed: 2022-09-20.

[Exp] *Explorable Explanations*. https://explorabl.es/. Accessed: 2023-10-23. 2023.

[Gtda] *Exploring Global Terrorism Data: A Web-based Visualization of Temporal Data*. https://www.cs.umd.edu/hcil/gtd/gtd/intro.html. Accessed: 2023-12-14. 2008.

[Fel08] P. Felten. "Visual Literacy." In: *Change: The Magazine of Higher Learning* 40.6 (2008), pp. 60–64. DOI: 10.3200/CHNG.40.6.60-64.

[Fen+15] W. Feng, C. Zhang, W. Zhang, J. Han, J. Wang, C. Aggarwal, and J. Huang. "STREAMCUBE: Hierarchical spatio-temporal hashtag clustering for event exploration over the Twitter stream." In: *IEEE Proc. of 31st International Conference on Data Engineering*. 2015, pp. 1561–1572.

[FJL22] E. E. Firat, A. Joshi, and R. S. Laramee. "Interactive visualization literacy: The state-of-the-art." In: *Information Visualization* 21.3 (2022), pp. 285–310. DOI: 10.1177/14738716221081831.

[Flo] *FlowingData: Best Data Visualization Projects of...* https://flowingdata.com/tag/best-of/. Accessed: 2022-11-23.

[Fun18] J Fung. *Manhattan Population Explorer*. https://manpopex.us/. Accessed: 2022-09-20. 2018.

[Gsab] *GSAP-ASEQ*. https://github.com/vinccenttt/gsap-aseq. Accessed: 2024-02-15. 2023.

[Gsac] *GSAP*. https://gsap.com/. Accessed: 2024-02-15.

[GS62] D. Gale and L. S. Shapley. "College Admissions and the Stability of Marriage." In: *The American Mathematical Monthly* 69.1 (1962), pp. 9–15.

[GY20] L. Gamio and K. Yourish. *See How the Coronavirus Death Toll Grew Across the U.S.* The New York Times. https://www.nytimes.com/interactive/2020/04/06/us/coronavirus-deaths-united-states.html. 2020.

[GP01] N. Gershon and W. Page. "What Storytelling Can Do for Information Visualization." In: *Commun. ACM* 44.8 (2001), 31–37. DOI: 10.1145/381641.381653.

[GS03] F. Giunchiglia and P. Shvaiko. "Semantic matching." In: *The Knowledge Engineering Review* 18.3 (2003), 265–280.

[Gtdb] *Global Terrorism Database*. START (National Consortium for the Study of Terrorism and Responses to Terrorism). http://www.start.umd.edu/gtd. Accessed: 2024-03-15. 2013.

[God+08] A. Godwin, R. Chang, R. Kosara, and W. Ribarsky. "Visual analysis of entity relationships in the Global Terrorism Database." In: *Defense and Security*. Vol. 6983. International Society for Optics and Photonics. SPIE, 2008, 69830G. DOI: 10.1117/12.778084.

[GGCM18] S. J. Green, K. Grorud-Colvert, and H. Mannix. "Uniting science and stories: Perspectives on the value of storytelling for communicating science." In: *FACETS* 3.1 (2018), pp. 164–173. DOI: 10.1139/facets-2016-0079.

[GSS11] J. Górecki, K. Slaninová, and V. Snášel. "Visual investigation of similarities in Global Terrorism Database by means of synthetic social networks." In: *International Conference on Computational Aspects of Social Networks (CASoN)*. 2011, pp. 255–260. DOI: 10.1109/CASON.2011.6085954.

[Gö23] V. Göring. "Designing a Library to Create Animated Sequences Using D3.js." Bachelor's Thesis. Otto von Guericke University Magdeburg, 2023. URL: https://www.vismd.de/wp-content/uploads/2024/02/bachelor_thesis_vincent_goring.pdf.

[Har03] M. Harrower. "Tips for designing effective animated maps." In: *Cartographic Perspectives* 44 (2003), pp. 63–65. DOI: 10.14714/CP44.516.

[HR07] J. Heer and G. Robertson. "Animated transitions in statistical data graphics." In: *IEEE Transactions on Visualization and Computer Graphics* 13.6 (2007), pp. 1240–1247. DOI: 10.1109/TVCG.2007.70539.

[Hla+20] M. Hlawitschka, G. Scheuermann, C. Blecha, M. Streit, and A. Varshney. "Collaborating Successfully with Domain Experts." In: *Foundations of Data Visualization*. Ed. by M. Chen, H. Hauser, P. Rheingans, and G. Scheuermann. Springer International Publishing, 2020, pp. 285–293. DOI: 10.1007/978-3-030-34444-3_15.

[Hoh+20] F. Hohman, M. Conlen, J. Heer, and D. H. P. Chau. "Communicating with Interactive Articles." In: *Distill* (2020). DOI: 10.23915/distill.00028.

[Hoo+24] N. de Hoog, A. Voce, E. Morresi, M. Ganguly, and A. Kirk. *How war destroyed Gaza's neighbourhoods – visual investigation*. The Guardian. https://www.theguardian.com/world/ng-interactive/2024/jan/30/how-war-destroyed-gazas-neighbourhoods-visual-investigation. 2024.

[HPG13] M. C. Hout, M. H. Papesh, and S. D. Goldinger. "Multidimensional scaling." In: *Wiley Interdisciplinary Reviews: Cognitive Science* 4.1 (2013), pp. 93–103. DOI: 10.1002/wcs.1203.

[Hsu+18] Y.-C. Hsu, P. Dille, R. Sargent, C. Bartley, and I. Nourbakhsh. "A Web-based Large-scale Timelapse Editor for Creating and Sharing Guided Video Tours and Interactive Slideshows." In: *arXiv:1804.03307 (preprint)* (2018). https://arxiv.org/abs/1804.03307.

[HD11] J. Hullman and N. Diakopoulos. "Visualization Rhetoric: Framing Effects in Narrative Visualization." In: *IEEE Transactions on Visualization and Computer Graphics* 17.12 (2011), pp. 2231–2240. DOI: 10.1109/TVCG.2011.255.

[Hul+13] J. Hullman, S. Drucker, N. H. Riche, B. Lee, D. Fisher, and E. Adar. "A deeper understanding of sequence in narrative visualization." In: *IEEE Transactions on Visualization and Computer Graphics* 19.12 (2013), pp. 2406–2415. DOI: 10.1109/TVCG.2013.119.

[Hun19] Y.-H. Hung. "Affective Engagement in Information Visualization." PhD thesis. Purdue University Graduate School, 2019.

[HP17] Y.-H. Hung and P. Parsons. "Assessing User Engagement in Information Visualization." In: *Proc. of CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 2017, 1708–1717. DOI: 10.1145/3027063.3053113.

[Hun79] A. Hunsaker. "Enjoyment and Information Gain in Science Articles." In: *Journalism Quarterly* 56.3 (1979), pp. 617–619. DOI: 10.1177/107769907905600323.

[IPC14] IPCC (Adopted). *Climate change 2014 synthesis report*. IPCC: Geneva, CH. https://www.ipcc.ch/site/assets/uploads/2018/02/SYR_AR5_FINAL_full.pdf. 2014.

[Ive] *IVEE*. https://isgwww.cs.ovgu.de/~benedikt/endovis_workflow/. Accessed: 2024-02-12. 2023.

[IKP12] S. M. Iacus, G. King, and G. Porro. "Causal Inference without Balance Checking: Coarsened Exact Matching." In: *Political Analysis* 20.1 (2012), 1–24. DOI: 10.1093/pan/mpr013.

[ISB14] G. Ifrim, B. Shi, and I. Brigadir. "Event detection in twitter using aggressive filtering and hierarchical tweet clustering." In: *2nd Workshop on Social News on the Web (SNOW)*. ACM, 2014.

[Ili18] M. Iliff. *Billions of Birds Migrate. Where Do They Go?* National Geographic. https://www.nationalgeographic.com/magazine/graphics/bird-migration-interactive-maps. 2018.

[Ins14] Institute of Medicine. "Numeracy and the affordable care act: Opportunities and challenges." In: *Health literacy and numeracy: Workshop summary*. National Academies Press (US). 2014.

[Not] *Introduction: A Characterization of Interactive Visual Data Stories With a Spatio-Temporal Context*. https://youtu.be/35CLrxdC4Xk. Accessed: 2024-02-19. 2023.

[Jac+20] B. T. Jacobs, R. Champine, J. Treat, A. Borunda, and K. Berne. *Your Climate, Changed*. National Geographic. https://www.nationalgeographic.com/magazine/graphics/see-how-your-citys-climate-might-change-by-2070-feature. 2020.

[Jen+18] B. Jenny, D. M. Stephen, I. Muehlenhaus, B. E. Marston, R. Sharma, E. Zhang, and H. Jenny. "Design principles for origin-destination flow maps." In: *Cartography and Geographic Information Science* 45.1 (2018), pp. 62–75. DOI: 10.1080/15230406.2016.1262280.

[JS91] B. Johnson and B. Shneiderman. "Tree-maps: A space-filling approach to the visualization of hierarchical information structures." In: *IEEE Proc. of Visualization*. 1991, pp. 284–291.

[Jon+08] J. Jones, R. Chang, T. Butkiewicz, and W. Ribarsky. "Visualizing uncertainty for geographical information in the global terrorism database." In: *Defense and Security*. Vol. 6983. International Society for Optics and Photonics. SPIE, 2008, 69830E. DOI: 10.1117/12.777695.

[KS18] D. Karell and S. Schutte. "Aid, exclusion, and the local dynamics of insurgency in Afghanistan." In: *Journal of Peace Research* 55.6 (2018), pp. 711–725. DOI: 10.1177/0022343318777566. (Visited on 10/23/2023).

[KM22] L. Karklis and R. Mellen. *Four maps that explain the Russia-Ukraine conflict*. The Washington Post. https://www.washingtonpost.com/world/2022/01/21/ukraine-russia-explain-maps/. 2022.

[Kei02] D. Keim. "Information visualization and Visual Data Mining." In: *IEEE Transactions on Visualization and Computer Graphics* 8.1 (2002), pp. 1–8. DOI: 10.1109/2945.981847.

[Kel19] M. Kelly. *The standard errors of persistence*. CEPR Discussion Paper No. DP13783. Available at SSRN: https://ssrn.com/abstract=3401870. 2019.

[Kos16] R. Kosara. "Presentation-Oriented Visualization Techniques." In: *IEEE Computer Graphics and Applications* 36.1 (2016), pp. 80–85. DOI: 10.1109/MCG.2016.2.

[Kos+24] G. Kostiuchik, L. Sharan, B. Mayer, I. Wolf, B. Preim, and S. Engelhardt. "Surgical phase and instrument recognition: how to identify appropriate dataset splits." In: *International Journal of Computer Assisted Radiology and Surgery* -.- (2024). DOI: 10.1007/s11548-024-03063-9.

[Kra03] M.-J. Kraak. "The space-time cube revisited from a geovisualization perspective." In: *21st International Cartographic Conference*. 2003, pp. 1988–1996.

[KK17] M.-J. Kraak and I. Kveladze. "Narrative of the annotated Space–Time Cube–revisiting a historical event." In: *Journal of Maps* 13.1 (2017), pp. 56–61. DOI: 10.1080/17445647.2017.1323034.

[KL83] J. B. Kruskal and J. M. Landwehr. "Icicle plots: Better displays for hierarchical clustering." In: *The American Statistician* 37.2 (1983), pp. 162–168.

[LD07] G. LaFree and L. Dugan. "Introducing the Global Terrorism Database." In: *Terrorism and Political Violence* 19.2 (2007), pp. 181–204. DOI: 10.1080/09546550701246817.

[Lam19] M. Lambrechts. *Why Budapest, Warsaw, and Lithuania split themselves in two*. The Pudding. https://pudding.cool/2019/04/eu-regions/. 2019.

[Lan+21] F. Lan, M. Young, L. Anderson, A. Ynnerman, A. Bock, M. A. Borkin, A. G. Forbes, J. A. Kollmeier, and B. Wang. "Visualization in Astrophysics: Developing New Methods, Discovering Our Universe, and Educating the Earth." In: *Computer Graphics Forum* 40.3 (2021), pp. 635–663. DOI: 10.1111/cgf.14332.

[LCB21] S. Latif, S. Chen, and F. Beck. "A Deeper Understanding of Visualization-Text Interplay in Geographic Data-driven Stories." In: *Computer Graphics Forum*. Vol. 40. 3. 2021, pp. 311–322. DOI: 10.1111/cgf.14309.

[Lee+15] B. Lee, N. H. Riche, P. Isenberg, and S. Carpendale. "More than telling a story: Transforming data into visually shared stories." In: *IEEE Computer Graphics and Applications* 35.5 (2015), pp. 84–90. DOI: 10.1109/MCG.2015.99.

[Lee08] J. Lee. "Exploring global terrorism data: a web-based visualization of temporal data." In: *XRDS: Crossroads* 15.2 (2008), pp. 7–14.

[LKK17] S. Lee, S.-H. Kim, and B. C. Kwon. "VLAT: Development of a Visualization Literacy Assessment Test." In: *IEEE Transactions on Visualization and Computer Graphics* 23.1 (2017), pp. 551–560. DOI: 10.1109/TVCG.2016.2598920.

[Lee+19] S. Lee, B. C. Kwon, J. Yang, B. C. Lee, and S.-H. Kim. "The Correlation between Users' Cognitive Characteristics and Visualization Literacy." In: *Applied Sciences* 9.3 (2019). DOI: 10.3390/app9030488.

[LR20] I. Levy-Rubinett. *Data journalism during COVID-19: "this is the biggest story as a data journalist that I've ever encountered."* Medium. https://medium.com/nightingale/data-journalism-during-covid-19-this-is-the-biggest-story-as-a-data-journalist-that-ive-ever-e80e21297d4a. Accessed: 2022-09-20. 2020.

[Li+23] W. Li, Z. Wang, Y. Wang, D. Weng, L. Xie, S. Chen, H. Zhang, and H. Qu. "GeoCamera: Telling Stories in Geographic Visualizations with Camera Movements." In: *Proc. of CHI Conference on Human Factors in Computing Systems*. ACM, 2023. DOI: 10.1145/3544548.3581470.

[LD11] D. Lloyd and J. Dykes. "Human-Centered Approaches in Geovisualization Design: Investigating Multiple Methods Through a Long-Term Case Study." In: *IEEE Transactions on Visualization and Computer Graphics* 17.12 (2011), pp. 2498–2507.

[LN20] J. Love and R. Nordland. *Why Afghanistan Became an Invisible War*. The New York Times. https://www.nytimes.com/interactive/2020/03/01/world/asia/afghanistan-invisible-war.html. 2020.

[Lu20] D. Lu. *There Has Been an Increase in Other Causes of Deaths, Not Just Coronavirus*. The New York Times. https://www.nytimes.com/interactive/2020/06/01/us/coronavirus-deaths-new-york-new-jersey.html. 2020.

[Lu+16] Y. Lu, M. Steptoe, S. Burke, H. Wang, J. Tsai, H. Davulcu, D. Montgomery, S. R. Corman, and R. Maciejewski. "Exploring Evolving Media Discourse Through Event Cueing." In: *IEEE Transactions on Visualization and Computer Graphics* 22.1 (2016), pp. 220–229.

[LJ12] P. Lundblad and M. Jern. "Visual storytelling in education applied to spatial-temporal multivariate statistics data." In: *Expanding the Frontiers of Visual Analytics and Visualization*. Springer, 2012, pp. 175–193. DOI: 10.1007/978-1-4471-2804-5_11.

[Max] *MAXAR*. https://www.maxar.com/. Accessed: 2023-11-10.

[MH08] L. van der Maaten and G. Hinton. "Visualizing Data using t-SNE." In: *Journal of Machine Learning Research* 9.86 (2008). http://jmlr.org/papers/v9/vandermaaten08a.html, pp. 2579–2605.

[Mac04] A. M. MacEachren. *How maps work: representation, visualization, and design*. Guilford Press, 2004.

[Mah+22] S. Mahajan, B. Chen, A. Karduni, Y.-S. Kim, and E. Wall. "VIBE: A Design Space for VIsual Belief Elicitation in Data Journalism." In: *Computer Graphics Forum*. Vol. 41. 3. 2022, pp. 477–488. DOI: 10.1111/cgf.14556.

[Mal19] C. Malone. *Joe Biden's Greatest Strength Is His Greatest Vulnerability*. FiveThirtyEight. https://fivethirtyeight.com/features/the-front-runner/. 2019.

[Mar16] T. Marshall. *Prisoners of geography: ten maps that explain everything about the world*. Vol. 1. Simon and Schuster, 2016.

[Mat] *Matched Wake Analysis Explained*. https://matchedwake.com/. Accessed: 2023-10-23. 2023.

[Mwa] *Matched Wake Analysis: R package*. https://CRAN.R-project.org/package=mwa. Accessed: 2023-10-23. 2023.

[May+21] B. Mayer, K. Lawonn, K. Donnay, B. Preim, and M. Meuschke. "VEHICLE: Validation and Exploration of the Hierarchical Integration of Conflict Event Data." In: *Computer Graphics Forum* 40.3 (2021), pp. 1–12. DOI: https://doi.org/10.1111/cgf.14284.

[May+24] B. Mayer, K. Donnay, K. Lawonn, B. Preim, and M. Meuschke. "Expert exploranation for communicating scientific methods - A case study in conflict research." In: *Computers & Graphics* 120 (2024), p. 103937. DOI: https://doi.org/10.1016/j.cag.2024.103937.

[May+23a] B. Mayer, M. Meuschke, J. Chen, B. P. Müller-Stich, M. Wagner, B. Preim, and S. Engelhardt. "Interactive visual exploration of surgical process data." In: *International Journal of Computer Assisted Radiology and Surgery* 18.1 (2023), pp. 127–137. DOI: https://doi.org/10.1007/s11548-022-02758-1.

[May+23b] B. Mayer, N. Steinhauer, B. Preim, and M. Meuschke. "A Characterization of Interactive Visual Data Stories With a Spatio-Temporal Context." In: *Computer Graphics Forum* 42.6 (2023), e14922. DOI: https://doi.org/10.1111/cgf.14922.

[MW18] E. Mayr and F. Windhager. "Once upon a Spacetime: Visual Storytelling in Cognitive and Geotemporal Information Spaces." In: *ISPRS International Journal of Geo-Information* 7.3 (2018). DOI: 10.3390/ijgi7030096.

[MHM20] L. McInnes, J. Healy, and J. Melville. *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction*. 2020. arXiv: 1802.03426 [stat.ML].

[McK+17] S. McKenna, N. Henry Riche, B. Lee, J. Boy, and M. Meyer. "Visual narrative flow: Exploring factors shaping data visualization story reading experiences." In: *Computer Graphics Forum*. Vol. 36. 3. 2017, pp. 377–387. DOI: 10.1111/cgf.13195.

[Mer21] D. Meredith. *Explaining research: How to reach key audiences to advance your work*. Oxford University Press, 2021. DOI: 10.1080/0889311X.2013.769530.

[Meu+22] 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." In: *Computers & Graphics* 107 (2022), pp. 144–157. DOI: 10.1016/j.cag.2022.07.017.

[Mir] *Miro*. https://miro.com/. Accessed: 2023-10-23. 2023.

[Mit+23] S. Mittenentzwei, V. Weiß, S. Schreiber, L. A. Garrison, S. Bruckner, M. Pfister, B. Preim, and M. Meuschke. "Do Disease Stories Need a Hero? Effects of Human Protagonists on a Narrative Visualization about Cerebral Small Vessel Disease." In: *Computer Graphics Forum* 42.3 (2023), pp. 123–135. DOI: 10.1111/cgf.14817.

[Mun09] T. Munzner. "A Nested Model for Visualization Design and Validation." In: *IEEE Transactions on Visualization and Computer Graphics* 15.6 (2009), pp. 921–928.

[NPD17] T. Nagel, C. Pietsch, and M. Dork. "Staged analysis: From evocative to comparative visualizations of Urban mobility." In: *IEEE VIS Arts Program (VISAP)*. 2017, pp. 1–8. DOI: 10.1109/VISAP.2017.8282374.

[Nat17] National Academies of Sciences Engineering and Medicine. *Communicating Science Effectively: A Research Agenda*. The National Academies Press, 2017. DOI: 10.17226/23674.

[NL10] A. Negrete and C. Lartigue. "The science of telling stories: Evaluating science communication via narratives (RIRC method)." In: *Media and Communication Studies* 2.4 (2010), pp. 98–110.

[Nor13] D. Norman. *The design of everyday things*. 2nd ed. Basic Books, 2013.

[OT08] H. L. O'Brien and E. G. Toms. "What is user engagement? A conceptual framework for defining user engagement with technology." In: *Journal of the American Society for Information Science and Technology* 59.6 (2008), pp. 938–955. DOI: 10.1002/asi.20801.

[Obsb] *Observable Notebook: A Characterization of Interactive Visual Data Stories With a Spatio-Temporal Context*. https://observablehq.com/@zykel/characterization. Accessed: 2023-02-14.

[Obsc] *Observable*. https://observablehq.com/. Accessed: 2023-11-10.

[Pę17] P. Pędzich. "Image of the World on polyhedral maps and globes." In: *Polish Cartographical Review* 48.4 (2017), pp. 197–210. DOI: doi:10.1515/pcr-2016-0014.

[Ral+10] C. Raleigh, A. Linke, H. Hegre, and J. Karlsen. "Introducing ACLED: An Armed Conflict Location and Event Dataset: Special Data Feature." In: *Journal of Peace Research* 47.5 (2010), pp. 651–660.

[Rea] *React*. https://legacy.reactjs.org/. Accessed: 2023-10-23. 2023.

[RT81] E. M. Reingold and J. S. Tilford. "Tidier Drawings of Trees." In: *IEEE Transactions on Software Engineering* SE-7.2 (1981), pp. 223–228.

[Ren+17] D. Ren, M. Brehmer, B. Lee, T. Höllerer, and E. K. Choe. "Chartaccent: Annotation for data-driven storytelling." In: *Proc. of Pacific Visualization Symposium (PacificVis)*. 2017, pp. 230–239. DOI: 10.1109/PACIFICVIS.2017.8031599.

[Reu] *Reuters*. https://www.reuters.com/. Accessed: 2023-11-10.

[Ric48] L. F. Richardson. "Variation of the Frequency of Fatal Quarrels with Magnitude." In: *Journal of the American Statistical Association* 43.244 (1948), pp. 523–546.

[Ric+18] N. H. Riche, C. Hurter, N. Diakopoulos, and S. Carpendale. *Data-driven storytelling*. AK Peters/CRC Press, 2018. DOI: 10.1201/9781315281575.

[Rob+17] A. C. Robinson, D. J. Peuquet, S. Pezanowski, F. A. Hardisty, and B. Swedberg. "Design and evaluation of a geovisual analytics system for uncovering patterns in spatio-temporal event data." In: *Cartography and Geographic Information Science* 44.3 (2017), pp. 216–228.

[Roc] *Rochan Consulting*. https://rochan-consulting.com/. Accessed: 2023-11-10.

[Rot+22] A. Roth, D. Sabbagh, P. Scruton, H. Symons, F. Sheehy, G. Swann, and N. de Hoog. *Russia's war in Ukraine: complete guide in maps, video and pictures*. The Guardian. https://www.theguardian.com/world/2022/mar/17/russias-war-in-ukraine-complete-guide-in-maps-video-and-pictures. 2022.

[Rot13] R. E. Roth. "An Empirically-Derived Taxonomy of Interaction Primitives for Interactive Cartography and Geovisualization." In: *IEEE Transactions on Visualization and Computer Graphics* 19.12 (2013), pp. 2356–2365. DOI: 10.1109/TVCG.2013.130.

[Rot21] R. E. Roth. "Cartographic design as visual storytelling: synthesis and review of map-based narratives, genres, and tropes." In: *The Cartographic Journal* 58.1 (2021), pp. 83–114. DOI: 10.1080/00087041.2019.1633103.

[RRM15] R. E. Roth, K. S. Ross, and A. M. MacEachren. "User-Centered Design for Interactive Maps: A Case Study in Crime Analysis." In: *ISPRS International Journal of Geo-Information* 4.1 (2015), pp. 262–301. DOI: 10.3390/ijgi4010262.

[Rou+20] A. Rourke, K. Rawlinson, D. Gayle, A. Topping, A. Mohdin, and H. Sullivan. *Confirmed cases pass 1 million – as it happened*. The Guardian. https://www.theguardian.com/world/live/2020/apr/02/coronavirus-live-news-global-cases-latest-updates?page=with:block-5e85f1038f08532a0e666b02. 2020.

[Sac+16] D. Sacha, H. Senaratne, B. C. Kwon, G. Ellis, and D. A. Keim. "The Role of Uncertainty, Awareness, and Trust in Visual Analytics." In: *IEEE Transactions on Visualization and Computer Graphics* 22.1 (2016), pp. 240–249. DOI: 10.1109/TVCG.2015.2467591.

[SKB17] J. F. Saldarriaga, L. Kurgan, and D. Brawley. "Visualizing Conflict: Possibilities for Urban Research." In: *Urban Planning* 2.1 (2017), 100–107. DOI: 10.17645/up.v2i1.880.

[Sal+12] I. Salehyan, C. S. Hendrix, J. Hamner, C. Case, C. Linebarger, E. Stull, and J. Williams. "Social Conflict in Africa: A New Database." In: *International Interactions* 38.4 (2012), pp. 503–511.

[San22] R. B. Santos. *Crime analysis with crime mapping*. 5th ed. SAGE Publications, 2022.

[SH14] A. Satyanarayan and J. Heer. "Authoring Narrative Visualizations with Ellipsis." In: *Computer Graphics Forum* 33.3 (2014), pp. 361–370. DOI: 10.1111/cgf.12392.

[SS06] H. Schulz and H. Schumann. "Visualizing Graphs - A Generalized View." In: *Proc. of International Conference on Information Visualisation (IV)*. 2006, pp. 166–173.

[SHS10] H.-J. Schulz, S. Hadlak, and H. Schumann. "The design space of implicit hierarchy visualization: A survey." In: *IEEE Transactions on Visualization and Computer Graphics* 17.4 (2010), pp. 393–411.

[SD14] S. Schutte and K. Donnay. "Matched wake analysis: Finding causal relationships in spatiotemporal event data." In: *Political Geography* 41 (2014), pp. 1–10. DOI: 10.1016/j.polgeo.2014.03.001.

[SMM12] M. Sedlmair, M. Meyer, and T. Munzner. "Design Study Methodology: Reflections from the Trenches and the Stacks." In: *IEEE Transactions on Visualization and Computer Graphics* 18.12 (2012), pp. 2431–2440. DOI: 10.1109/TVCG.2012.213.

[SH10] E. Segel and J. Heer. "Narrative visualization: Telling stories with data." In: *IEEE Transactions on Visualization and Computer Graphics* 16.6 (2010), pp. 1139–1148. DOI: 10.1109/TVCG.2010.179.

[SC19] S. Sengupta and W. Cai. *A Quarter of Humanity Faces Looming Water Crises*. The New York Times. https://www.nytimes.com/interactive/2019/08/06/climate/world-water-stress.html. 2019.

[SZ18] D. Seyser and M. Zeiller. "Scrollytelling–an analysis of visual storytelling in online journalism." In: *Proc. of International Conference on Information Visualisation (IV)*. 2018, pp. 401–406. DOI: 10.1109/iV.2018.00075.

[Shn97] B. Shneiderman. "Direct Manipulation for Comprehensible, Predictable and Controllable User Interfaces." In: *Proc. of the 2nd International Conference on Intelligent User Interfaces*. ACM, 1997, 33–39.

[Shu+21] X. Shu, A. Wu, J. Tang, B. Bach, Y. Wu, and H. Qu. "What Makes a Data-GIF Understandable?" In: *IEEE Transactions on Visualization and Computer Graphics* 27.2 (2021), pp. 1492–1502. DOI: 10.1109/TVCG.2020.3030396.

[Sil+19] R. A. Silva, J. M. Pires, N. Datia, M. Y. Santos, B. Martins, and F. Birra. "Visual analytics for spatiotemporal events." In: *Multimedia Tools and Applications* 78.23 (2019), 32805–32847. DOI: 10.1007/s11042-019-08012-2.

[Slo+23] T. A. Slocum, R. B. McMaster, F. C. Kessler, and H. H. Howard. *Thematic Cartography and Geovisualization, Fourth Edition*. CRC Press, 2023.

[ST04] D. R. Sokoloff and R. K. Thornton. *Interactive Lecture Demonstrations*. John Wiley & Sons, 2004.

[Son+22a] Z. Song, R. E. Roth, L. Houtman, T. Prestby, A. Iverson, and S. Gao. "Visual Storytelling with Maps: An Empirical Study on Story Map Themes and Narrative Elements, Visual Storytelling Genres and Tropes, and Individual Audience Differences." In: *Cartographic Perspectives* 100 (2022). DOI: 10.14714/CP100.1759.

[Son+22b] P. Sonne, J. S. Lee, M. Ilyushina, R. Mellen, and A. Mirza. *The TikTok buildup: Videos reveal Russian forces closing in on Ukraine*. The Washington Post. https://www.washingtonpost.com/world/2022/02/11/russia-ukraine-military-videos-tiktok/. 2022.

[Sta21] Staff. *Tracking the coronavirus across Europe*. The Economist. https://www.economist.com/graphic-detail/tracking-coronavirus-across-europe. 2021.

[Sta22] T. N. Y. T. Staff. *Maps: Tracking the Russian Invasion of Ukraine*. The New York Times. https://www.nytimes.com/interactive/2022/world/europe/ukraine-maps.html. 2022.

[Sta24] T. N. Y. T. Staff. *Maps: Tracking the Attacks in Israel and Gaza*. The New York Times. https://www.nytimes.com/interactive/2023/10/07/world/middleeast/israel-gaza-maps.html. 2024.

[Sto+16] C. D. Stolper, B. Lee, N. Henry Riche, and J. Stasko. *Emerging and Recurring Data-Driven Storytelling Techniques: Analysis of a Curated Collection of Recent Stories*. Tech. rep. MSR-TR-2016-14. https://www.microsoft.com/en-us/research/publication/emerging-and-recurring-data-driven-storytelling-techniques-analysis-of-a-curated-collection-of-recent-stories/. 2016.

[Sun+23] M. Sun, L. Cai, W. Cui, Y. Wu, Y. Shi, and N. Cao. "Erato: Cooperative Data Story Editing via Fact Interpolation." In: *IEEE Transactions on Visualization and Computer Graphics* 29.1 (2023), pp. 983–993. DOI: 10.1109/TVCG.2022.3209428.

[SM13] R. Sundberg and E. Melander. "Introducing the UCDP Georeferenced Event Dataset." In: *Journal of Peace Research* 50.4 (2013), pp. 523–532.

[Swa+23] L. Swan, H. Symons, F. Ali, E. Morresi, and A. Olorenshaw. *Israel-Hamas war: a visual guide in maps, video and satellite images*. The Guardian. https://www.theguardian.com/world/2023/oct/13/israel-hamas-war-conflict-visual-guide-in-maps-satellite-images-and-video. 2023.

[Ctp] *The American Enterprise Institute's Critical Threats Project*. https://www.criticalthreats.org/. Accessed: 2023-11-10.

[Tim+18] M.-L. Timcke, A. Pätzold, D. Wendler, and C. Möller. *Alt- oder Neubau? So wohnt Berlin*. Berliner Morgenpost. https://interaktiv.morgenpost.de/so-alt-wohnt-berlin/. 2018.

[Tob73] W. R. Tobler. "Choropleth maps without class intervals." In: *Geographical Analysis* 5.3 (1973), pp. 262–265.

[TSWS05] C. Tominski, P. Schulze-Wollgast, and H. Schumann. "3D Information Visualization for Time Dependent Data on Maps." In: *Proc. of International Conference on Information Visualisation (IV)*. 2005, pp. 175–181. DOI: 10.1109/IV.2005.3.

[Tom+12] C. Tominski, H. Schumann, G. Andrienko, and N. Andrienko. "Stacking-Based Visualization of Trajectory Attribute Data." In: *IEEE Transactions on Visualization and Computer Graphics* 18.12 (2012), pp. 2565–2574.

[TAS04] C. Tominski, J. Abello, and H. Schumann. "Axes-Based Visualizations with Radial Layouts." In: *Proc. of ACM Symposium on Applied Computing*. ACM, 2004, 1242–1247.

[TS12] C. Tominski and H.-J. Schulz. "The Great Wall of Space-Time." In: *Vision, Modeling and Visualization*. The Eurographics Association, 2012. DOI: 10.2312/PE/VMV/VMV12/199-206.

[TS20] C. Tominski and H. Schumann. *Interactive Visual Data Analysis*. AK Peters/CRC Press, 2020.

[Ton+18] C. Tong, R. Roberts, R. Borgo, S. Walton, R. S. Laramee, K. Wegba, et al. "Storytelling and visualization: An extended survey." In: *Information* 9.3 (2018), p. 65. DOI: 10.3390/info9030065.

[Tre+21] J. Treat, C. Fellenz, E. Conant, and K. Elliott. *The High Price of Heat*. National Geographic. https://www.nationalgeographic.com/magazine/graphics/the-places-where-heat-will-have-the-greatest-cost-in-money-and-lives-feature. 2021.

[Tse16] A. Tse. *Why we are doing fewer interactives*. Malofiej Infographics World Summit. https://github.com/archietse/malofiej-2016/blob/master/tse-malofiej-2016-slides.pdf. 2016.

[Unh] *UNHCR: Forced displacement continues to grow as conflicts escalate*. https://www.unhcr.org/news/unhcr-forced-displacement-continues-grow-conflicts-escalate. Accessed: 2024-02-05. 2023.

[UKW17] C. Udovicich, V. Kasivisvanathan, and C. L. Winchester. "Communicating your research (part 1) – to the scientific community." In: *Journal of Clinical Urology* 10.4 (2017), pp. 396–399. DOI: 10.1177/2051415816668941.

[Veha] *VEHICLE: Overview Video*. https://youtu.be/35CLrxdC4Xk. Accessed: 2024-02-12. 2021.

[Vehb] *VEHICLE*. https://www.meltt.net/. Accessed: 2024-02-12. 2021.

[VV99] J. J. Van Wijk and H. Van de Wetering. "Cushion treemaps: visualization of hierarchical information." In: *IEEE Symposium on Information Visualization*. 1999, pp. 73–78.

[Vis] *VisualisingData: Best of the Visualization Web...* https://www.visualisingdata.com/2022/02/best-of-the-visualisation-web-october-2021/. Accessed: 2022-11-23.

[WBWK00] M. Q. Wang Baldonado, A. Woodruff, and A. Kuchinsky. "Guidelines for using multiple views in information visualization." In: *Proc. of the Working Conference on Advanced Visual Interfaces*. ACM, 2000, 110–119. DOI: 10.1145/345513.345271.

[Wan+08] X. Wang, E. Miller, K. Smarick, W. Ribarsky, and R. Chang. "Investigative Visual Analysis of Global Terrorism." In: *Computer Graphics Forum* 27.3 (2008), pp. 919–926. DOI: 10.1111/j.1467-8659.2008.01225.x.

[WGK10] M. O. Ward, G. Grinstein, and D. Keim. *Interactive Data Visualization*. AK Peters/CRC Press, 2010.

[Wei15] N. B. Weidmann. "On the Accuracy of Media-based Conflict Event Data." In: *Journal of Conflict Resolution* 59.6 (2015), pp. 1129–1149.

[Wei16] N. B. Weidmann. "A Closer Look at Reporting Bias in Conflict Event Data." In: *American Journal of Political Science* 60.1 (2016), pp. 206–218.

[WK09] N. B. Weidmann and D. Kuse. "WarViews: Visualizing and Animating Geographic Data on Civil War." In: *International Studies Perspectives* 10.1 (2009), pp. 36–48. DOI: 10.1111/j.1528-3585.2008.00356.x.

[WSM19] F. Windhager, G. Schreder, and E. Mayr. "On Inconvenient Images: Exploring the Design Space of Engaging Climate Change Visualizations for Public Audiences." In: *Workshop on Visualisation in Environmental Sciences (EnvirVis)*. The Eurographics Association, 2019. DOI: 10.2312/envirvis.20191098.

[Yan+21] 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." In: *IEEE Transactions on Visualization and Computer Graphics* 28.1 (2021), pp. 922–932. DOI: 10.1109/TVCG.2021.3114774.

[YKT05] R. Yang, P. Kalnis, and A. K. H. Tung. "Similarity Evaluation on Tree-Structured Data." In: *Proc. of ACM SIGMOD International Conference on Management of Data*. ACM, 2005, 754–765.

[YLT18] A. Ynnerman, J. Löwgren, and L. Tibell. "Exploranation: A New Science Communication Paradigm." In: *IEEE Computer Graphics and Applications* 38.3 (2018), pp. 13–20. DOI: 10.1109/MCG.2018.032421649.

[YSC19] L. Younes, A. Shaw, and P. Claire. *In a Notoriously Polluted Area of the Country, Massive New Chemical Plants Are Still Moving In*. ProPublica. https://projects.propublica.org/louisiana-toxic-air/. 2019.

[Zha+22] Y. Zhang, M. Reynolds, A. Lugmayr, K. Damjanov, and G. M. Hassan. "A Visual Data Storytelling Framework." In: *Informatics* 9.4 (2022). DOI: 10.3390/informatics9040073.

[ZFH08] J. Zhao, P. Forer, and A. S. Harvey. "Activities, ringmaps and geovisualization of large human movement fields." In: *Information Visualization* 7.3-4 (2008), pp. 198–209.

[ZOM19] Q. Zhi, A. Ottley, and R. Metoyer. "Linking and Layout: Exploring the Integration of Text and Visualization in Storytelling." In: *Computer Graphics Forum* 38.3 (2019), pp. 675–685. DOI: 10.1111/cgf.13719.

ABBREVIATIONS

ACLED Armed Conflict Location and Event Data [Ral+10]

AHC Agglomerative Hierarchical Clustering [DE84]

EA Example Application [KS18]

E.V. Exploratory Visualization

ExEx Expert Exploranation [May+24]

GED Uppsala Conflict Data Project – Georeferenced Event Dataset [SM13]

GTD Global Terrorism Database [Gtdb]

MDS Multidimensional Scaling [HPG13]

MELTT Matching Event Data by Location, Time, and Type [Don+19]

MWA Matched Wake Analysis [SD14]

SCAD Social Conflict Analysis Database [Sal+12]

VEHICLE Validation and Exploration of the Hierarchical Integration of Conflict Event Data [May+21]

colophon

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/