DOI: 10.1111/cgf.14284
Eurographics Conference on Visualization (EuroVis) 2021
R. Borgo, G. E. Marai, and T. von Landesberger
(Guest Editors)

Volume 40 (2021), Number 3

VEHICLE: Validation and Exploration of the Hierarchical Integration of Conflict Event Data

B. Mayer[1], K. Lawonn[2], K. Donnay[3], B. Preim[1], and M. Meuschke[1,2]

[1]University of Magdeburg, Department of Simulation and Graphics, Germany [2]University of Jena, Department of Computer Science, Germany
[3]University of Zurich, Department of Political Science, Switzerland

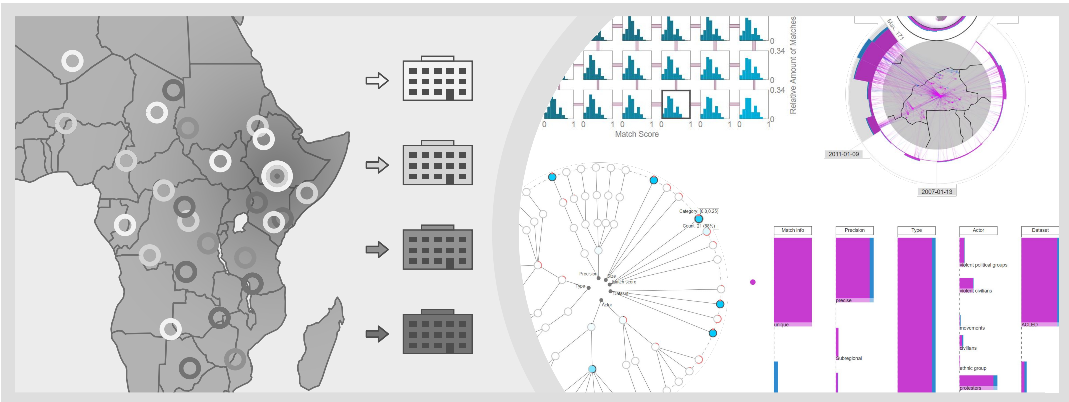

Figure 1: VEHICLE *is a web-based tool to validate and explore the integration of conflict event data that was recorded by different institutes. It allows analyzing parameter influence as well as spatial, temporal, and hierarchical distributions.*

Abstract

The exploration of large-scale conflicts, as well as their causes and effects, is an important aspect of socio-political analysis. Since event data related to major conflicts are usually obtained from different sources, researchers developed a semi-automatic matching algorithm to integrate event data of different origins into one comprehensive dataset using hierarchical taxonomies. The validity of the corresponding integration results is not easy to assess since the results depend on user-defined input parameters and the relationships between the original data sources. However, only rudimentary visualization techniques have been used so far to analyze the results, allowing no trustworthy validation or exploration of how the final dataset is composed.

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

CCS Concepts

• ***Human-centered computing*** → ***Information visualization;*** *Geographic visualization; Visualization toolkits;*

1. Introduction

The study of violent and non-violent conflicts is an active research area in the socio-political context. Over the last decades, an increasing number of datasets 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 dataset. 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 dataset. The different datasets, 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 datasets. Therefore, it is possible for recordings from different datasets to represent the same original incident while containing different information, as depicted on the left in Figure 1.

To solve this issue, experts recently started to integrate the information from different datasets to receive one holistic set. The most prominent method to do so is the semi-automatic method MELTT [DDM*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 dataset is the foundation for all further analysis and inference. Despite that, only basic techniques have been employed in the validation process so far.

Therefore, we present *VEHICLE*, a web-based tool to analyze the results of hierarchically integrated conflict event data. We designed it according to Munzner's Nested Design Model [Mun09] in collaboration with a domain expert. It allows the inspection of the influence of input parameters as well as the characteristics and similarities among the integrated data. For this purpose, we use multiple linked interactive visualizations, such as map-based depictions and radial glyph-based layouts, see Figure 1, right. We demonstrate the effectiveness of the tool by presenting two case studies and a qualitative and quantitative 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 visualization.
- A tool design and implementation following the derived workflow to solve the identified tasks.

2. Background on Conflict Data

In this section, we give a brief overview of how conflict data has been analyzed in the field of social sciences thus far.

Analysis of Conflict Data. 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 have been a number of large-scale institutional data collection initiatives that systematically collect data on constituent conflict events, e.g., individual attacks. These conflict events are geocoded, time-stamped, and annotated with contextual variables that describe the actors involved and the event type. The leading datasets in this area are the Armed Conflict Location and Event Data (ACLED) [RLHK10], the Uppsala Conflict Data Project–Georeferenced Event Data (GED) [SM13], the Global Terrorism Database (GTD) [GTD13], and the Social Conflict Analysis Database (SCAD) [SHH*12]. These data have been used by social science researchers 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 so far has been on the statistical analysis of conflict patterns and deriving causal relationships between contextual or behavioral covariates and the observed conflict patterns.

Analysis of Integration Results. The analysis of integrated data is of critical relevance for the quantitative study of conflict, given the incomplete and often complementary coverage of individual datasets [DDM*19]. The hierarchical integration of event data from different sources according to Donnay et al. [DDM*19] requires at least two hierarchical taxonomies describing different aspects of the data. The assumption is then 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 datasets. This automatic strategy is more efficient and replicable than attempting manual integration. Yet, it cannot be precluded that potential biases from a single dataset might carry over into the integrated dataset [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 example, prior work has considered mainly time series visualizations and corresponding statistical analyses, or static maps aggregated over fixed time windows [DF14, Wei16, DDM*19].

3. Related Work

In this work, we analyze geospatial event data with point spatial footprint and instant duration. Moreover, the events are subject to a hierarchical structure. We structured the related work accordingly. For the development of our tool, the fundamental book of Andrienko and Andrienko on the exploratory analysis of spatio-temporal data [AA06] provided us with structured guidance to identify principles, tasks, and techniques to apply in our design.

3.1. Visualization and Analysis of Event-based Data with Geospatial and Temporal Context

For data with a spatial context, map displays are suitable to allow the preservation of the spatial structure [AA06]. Maps can be used to perform exploratory data analysis, generate hypotheses, and construct knowledge [FG05]. They also support the identification and comparison of specific patterns as well as the estimation of values [Mac82, Mac95]. In general, three types of **spatial** encoding can be used to display data on a map: point, line, and area information [SMKH99, KPS03].

Point. To represent phenomena at a specific point, glyphs can be used, where their properties, such as shape, size, and color, are varied [DBBCM04, AAA*16]. Such maps facilitate qualitative analysis. However, depending on the number of events, this variant can cause visual clutter. To cope with this, one can change the appearance of the glyphs, distort them spatially, or animate them over time [KH98, ED07]. When combining spatial point information with a temporal dimension, the 3D *Space-Time Cube* is a common example of visualization [Kra03, GAA04]. Aside from constructing a 3D view, there is a variety of operations to derive visualizations from the Space-Time Cube model, which Bach et al. summarized in a descriptive framework [BDA*17]. However, the number of points and different event categories that can be displayed meaningfully in such views is limited and does not scale well with the number of events in applications such as ours. When, instead, depicting events as 2D point marks, the temporal information can be displayed using a separate view [AAB*13, LSB*16], *Small Multiples* [Tuf01] like by Li et al. [LCZ*19], or animation [AA06]. However, this usually hinders the perception of the spatio-temporal context as a whole.

Line. Encoding events as lines is only applicable if they are associated with network or movement data [Guo09, AA13, AAFW17]. However, conflict event data have static locations, so this encoding could only be used to display the relation between matched events.

Area. To reduce visual complexity, information can be mapped to area. To do so, events need to be aggregated over space using administrative boundaries such as cities or counties [Tob73] or artificial boundaries like grids [AA06, CSC*18]. The aggregates can also be determined using techniques like clustering [Yu06, IP14, RPP*17]. Appropriate visualizations for spatially aggregated data include the use of *Choropleth Maps* [Tob73], special glyphs [AAB*13, AAFW17, CSC*18], or, more general, spatially embedded charts [AA06]. The appropriateness of aggregation depends on the spatial resolution.

Aggregation may also be performed along the **temporal** dimension to receive time series. Those can be analyzed w.r.t. the occurrence of anomalies [CTB*12, IP14, CLZ*18] or characteristic patterns [AAM*10, AAB*13, LSB*16, RPP*17, LCZ*19], again, depending on the resolution. Two works [RPP*17, LSB*16] stand out as some of the few approaches to analyze conflict data with advanced visualization support. While Robinson et al. [RPP*17] only extract patterns of interest from event data streams belonging to a single dataset, the work by Lu et al. [LSB*16] is more closely related to our approach. Like us, they make use of multiple datasets to understand certain behaviors. However, they do not create a holistic dataset but, instead, use a secondary dataset to understand the potential causes for a change in behaviors in a primary dataset. Here, only specific points in time are of high relevance, identified by an automated model. Overall, the techniques in this paragraph are rather suitable for investigations with specific goals than to create a general understanding of the integration process.

3.2. 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 entities of subordinate levels can be arranged side by side to their parent entity or inside of it. One example of the side-by-side arrangement are *Icicle Plots* [KL83], where the hierarchy of rectangular shapes is expressed by stacking them. Extending that idea, the *Sunburst Visualization* [SZ00] allows to better convey structure and hierarchy by using radial arc segments instead of rectangles. In contrast, to display child entities inside of their parents, a self-containing arrangement of the visual entities is required. The most well-known representatives of this class are *Treemaps* [JS91]. They express the hierarchy through recursively 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 in our hierarchical visualizations.

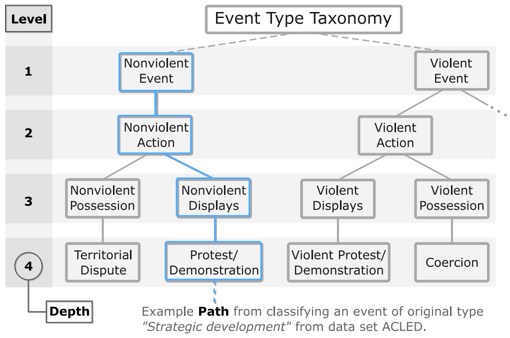

Figure 2: *An extract from the taxonomy for the attribute "event type" and corresponding terminology [DDM*19] is depicted.*

4. Characterization of the Data

The integrated datasets that we analyze in *VEHICLE* come from the projects ACLED [RLHK10], GED [SM13], GTD [GTD13], and SCAD [SHH*12]. We consider the events which took place in Africa between 1997 and 2016, a total of 197,502 events. Out of these, 140,738 belong to the dataset ACLED, 25,788 to UDP-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. We plan to release the integrated data on the same website as *VEHICLE* (www.meltt.net).

4.1. Definitions

In the following, the actual occurrence of an event in the real world is referred to as an *incident*, while the recording of an incident in a dataset is referred to as an *event*. This differentiation is made since 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 datasets prior to the integration or in the context of the integrated data. In the original datasets, events have *continuous attributes*, namely the location, i.e., longitude and latitude, and 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 datasets, hierarchical *taxonomies* were introduced by Donnay et al. [DDM*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 datasets into overarching *categories*. An example of how an event can be encoded in a taxonomy is given in Figure 2. The actor taxonomy consists of three levels and 21 categories, the type taxonomy of four levels and 28 categories, and the precision taxonomy consists of four levels and 19 categories. When an attribute value of an event from one of the original datasets 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. The higher the level is in the tree, the lower is the corresponding level number. 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 datasets, as explained in more detail in Section 4.2. As a result, events are either identified as *matches* or *uniques*. A *match* is a tuple of events from different datasets that were identified to represent the same 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".

4.2. Matching Procedure

For the application in this paper, we adopted the matching procedure and the underlying taxonomies from the MELTT algorithm [DDM*19]. It has two primary input parameters: Δs and Δt. These parameters control how far two different events may occur apart from each other w.r.t. 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 datasets to a uniform shape, we used this algorithm to add the matching information to the data ($140MB$ in total). It took about three hours on a standard desktop computer.

4.3. Classification of the Data

For the *data abstraction* (according to Munzner [Mun09]), we adopt the taxonomy for time-oriented data proposed by Aigner et al. [AMST11]. 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 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 *non-temporal*, as for each event only one precise timestamp is given and the external time is *dynamic*.

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.

5. Task Abstraction

In the following, we present our *domain problem characterization* in line with Munzner's Nested Design Model [Mun09]. Together with a domain 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 formulated which, in our case, reach from $\Delta s = 0km$ to $\Delta s = 250km$ and $\Delta t = 0d$ to $\Delta t = 2d$. 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. 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 spatio-temporal 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 helpful 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 datasets overlap and for which categories this is not the case.

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 of interest. 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 tool. This would substantially increase the complexity of the tool while it showed that not all users would want to analyze the results that deeply. For instance, the original datasets 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 dataset such as identifying causal effects. The techniques applied for such tasks vary strongly in the field of conflict research and would cause the tool to become too complex.

6. The Design of VEHICLE

The visual interface of *VEHICLE* comprises multiple linked components to handle the multi-faceted integration data. We followed Munzner's Nested Design Model [Mun09] when developing our tool in an iterative process using prototypes [LD11]. In this section, we describe our contributions w.r.t. *Operation and Data Type Abstraction* and *Visual Encoding and Interaction Design* [Mun09]. The different components of *VEHICLE* are embedded in an abstract workflow which we present before discussing their design in detail.

6.1. Workflow

Based on the analytical tasks from Section 5, we designed a workflow, including applicable *subtasks* 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 6.2.1). 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. During this 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 w.r.t. space, time (Section 6.2.2), and categorical attributes (Section 6.2.3). 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 datasets 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 6.2.2) as well as how they are distributed across the taxonomies (Section 6.2.4) 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 determining factors for certain kinds of matches are, e.g., for those with a good matching score.

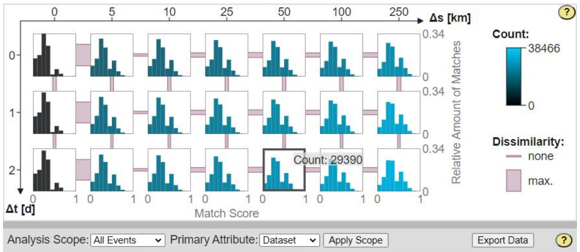

Figure 3: *The ParaMultiples summarize the matching outcomes in histograms. They depend on spatial (horizontal axis) and temporal (vertical axis) constraints. In addition, the dissimilarities between match distributions of adjacent outcomes are depicted as bars.*

6.2. 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 3. 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 3 (**T3**). Moreover, all components provide a help button to explain their key functionality. We allowed direct user interaction [Shn97] wherever possible.

6.2.1. 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 3. The analysts can compare how the matches are distributed for pre-defined parameter selections in Small Multiple [Tuf01] 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 sub-trees can be considered similar even if their weights are not distributed exactly across the same nodes.

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.

With this score, two kinds of similarity are 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 2, 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.

6.2.2. TempMap

In this view, the spatio-temporal distribution of the events is displayed in a comprehensive way (**T2a**, **T3**), see Figure 4. 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 [AWR*07]. If a primary attribute is selected in the "All Events" mode, the bars of each stack correspond to the different values of the primary attribute. If

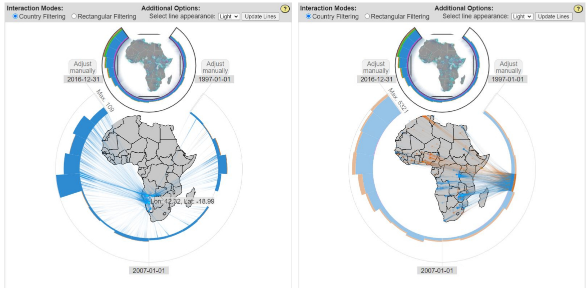

Figure 4: *The TempMap shows the spatio-temporal distribution of events. Hovering over a country (left) or a temporal bin (right) highlights the contained events. Hovering over the map also displays the longitude and latitude of the mouse cursor position.*

"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. [TSAA12].

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 glyph. To avoid this, we adjust the opacity of the glyphs 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 $\alpha(x) = c + m(x+b)^{-k}$. 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 glyph visibility for small sets and reducing visual clutter for large sets. In addition, we imposed a lower and upper limit for $\alpha(x)$. For the exact parameter values, please refer to the supplemental material. To further reduce clutter, the lines 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, see Figure 4.

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 glyphs. 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

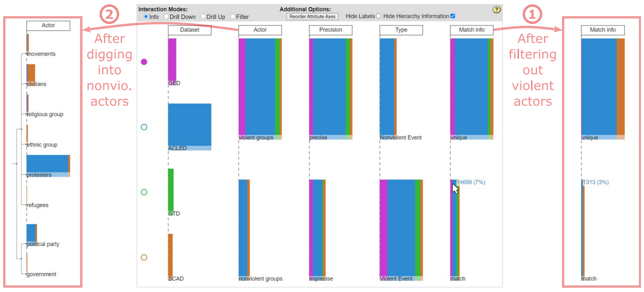

Figure 5: *The EventCharts display the distribution of the events across the categorical attributes. The results of filtering (1) and subsequent drilling into the actor taxonomy (2) are shown. In addition, the hierarchical information is depicted in (2).*

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 1, 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.

6.2.3. EventCharts

To analyze how the events are distributed across the categorical attributes, the *EventCharts* use hierarchical stacked barcharts, see Figure 5, as barcharts are effective for lookup and comparison tasks [AA06]. 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 dataset 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.

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*. 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". 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. The *Drill Up* mode allows the analysts to hover over bar stacks and 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.

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, a binary classification is assumed for each of the categories. That 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* [JS91], *Icicle Plots* [KL83], or *Parallel Sets* [BKH05] 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.

6.2.4. MatchTree

The *MatchTree* uses a radial tree layout to visualize the distribution of the matches across the categorical attributes, see Figure 6. 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 (*score*) discretized into four equally-sized bins, and the datasets that participate in each match (*dataset*).

Glyphs. Each category is represented by a circular glyph indi-

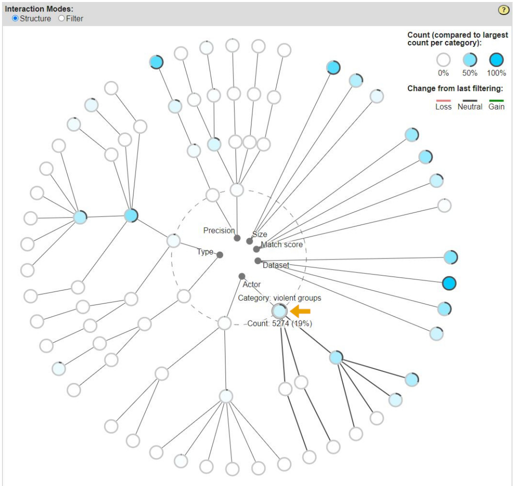

Figure 6: *The MatchTree displays the distribution of the matches across the categories. As the mouse hovers over a category, a level indicator ring is displayed and the induced sub-tree is highlighted.*

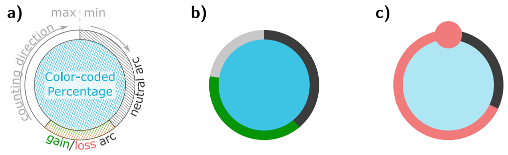

Figure 7: *The sketch (a) depicts the components of a node glyph. The green/red arc (b/c) indicates the number of matches gained/lost over the last filtering state. A dot (c) shows if the scale is exceeded.*

cating how many matches were identified in that category, see Figure 7. 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 6. Moreover, the sub-tree induced by the hovered node is highlighted.

Filtering. The filtering, designed to be consistent with the EventCharts, 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 7. This supports the user to track the changes across all nodes (**T3**). For instance, this is interesting when filtering out the matches with low matching scores to see which datasets primarily participated in these rather bad matches, or for what types of events they were identified. If the count of a node increases after filtering, the corresponding arc section (*gain arc*) is colored in dark green. If the count decreases, the lost section of the arc (*loss arc*) is colored in bright red. The remaining arc section (*neutral arc*) is colored in dark grey. Since the gain arc is part of the current count while the loss arc is not, the luminosity of the neutral arc was selected to resemble the gain arc more than the loss arc. If the length of the loss arc would exceed the current scale (it refers to the count from the previous step), it is shown by a red dot on top.

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

7. Evaluation

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

7.1. 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 seem quite low, except for the jump from 0 to 5km (**T1**). This shows both in the small dissimilarity bars and the relatively similar histogram distributions, see Figure 3. To investigate this behavior, the analyst inspects the MatchTree for different parameter combinations. This shows 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 those with high geographic coding precision, see Figure 6.

To examine the categories which barely have any matches, e.g., those with "non-violent 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 5. 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 dataset, see Figure 5. 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 datasets varies quite a lot, see Figure 4. 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 datasets 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. This is the case because the stated reasons hold for most parameter combinations.

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

7.2. 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 3. 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 changes between adjacent histograms are compared using the dissimilarity bars and the histogram fill colors. The analyst identifies a clear jump between $\Delta s = 0km$ and $\Delta s = 5km$ and confirms its appropriateness. Requiring the matched events to be recorded in exactly the same location is the constraint with the largest impact by far. As the overall number of found matches seems reasonable, too, 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 dataset. 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 $\Delta s = 50km$ and $\Delta t = 2d$, the analyst continues with "All Events" as the scope and "Match info" as the primary attribute. They inspect the TempMap and filter for the country Burkina Faso, see Figure 1. They gain an impression of how the matches are distributed by using temporal highlighting (**T2a**). To inspect clusters around the main capital and 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 1. Using the TempMap, they also find that for the most part the matches were identified around the capital, increasingly since 2011, see Figure 1. Finally, to see how the matches of the selected subset are distributed, they open the MatchTree (**T2b**). They find that the matches are of high quality according to the match scores, see Figure 1. 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.

7.3. Pair Analytics Sessions

The sessions were inspired by the evaluation methodology from Kaastra and Fisher [KF14] and Cakmak et al. [CSB*20]. In individual sessions that lasted between 50 and 90 minutes, in total five conflict researchers excluding our collaboration partner used *VEHICLE*. With our remote assistance, they analyzed the data described in this paper. We video-recorded the sessions and analyzed them afterward. Each session started by collecting some background information about the analysts. They had been working in conflict research for 3/6/9/13/19 years. They had at least some programming experience with languages like *R* and a background in statistics. They had experience with static data visualizations but not much with interactive ones. Two of them were female and three male.

After that, we explained the different components of the tool to them during which they already interacted with it and we discussed the first insights. Afterward, they explored the tool freely, if necessary with our assistance, while exchanging insights and impressions with us. Eventually, they were asked to fill out a short questionnaire. We structure the results according to the components.

ParaMultiples. The analysts 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 analyst would have liked to extend the range of Δt and to have hover information to better compare the dissimilarity counts.

TempMap. The analysts found the view "useful", and "straightforward". They liked "that you can manually type the date" but one would have wished to also adjust the temporal bin sizes and to have a less abstract map, e.g., with terrain or cities. Another would have liked to see the count of individual temporal bars in the histograms when hovering over them. In addition, some had trouble with using the temporal brush at the beginning but adapted it after a while.

EventCharts. They 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 analyst 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 initially struggled with processing what happened when we changed the primary attribute. The quantitative feedback reflects these difficulties, see Figure 8.

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

Overall. In the free investigation, the analysts showed excitement and curiosity for the various ways to explore the data. With one analyst, we even lost track of time to which he stated "it shows how much there is you can do". The analysts found the tool "super useful" but said that it was not for casual users. One analyst stated it "provides a much deeper look into [the integrated data]" than possible before. This also shows in the quantitative feedback, see

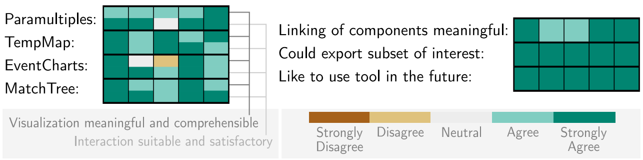

Figure 8: *Each column of the two tables corresponds to the quantitative feedback of one analyst after the pair analytics session.*

Figure 8. Still, one of the analysts would have liked to "go into the original event text to feel psychologically closer to what the system does". At the start of the session, the analysts required quite some assistance, but towards the end, they could do almost all investigations on their own. They expressed confidence to use the tool on their own and were excited to do so in the future.

8. Discussion and Future Work

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. We want to provide those minor improvements in the future and are excited to get even more such input after the release. This expectation seems justified considering the eager feedback of the analysts from the evaluation. A more difficult drawback to deal with turned out to be the tool's complexity, especially regarding the EventCharts. To make this view more comprehensible, 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.

To extend the approach to other application areas, the corresponding data needs to have the following properties. The event recordings, which come from different sources, represent their original incidents with a loss of information. Ideally, they are geotemporal and the data 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 glyph size could be decreased.

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, the ranges would differ between areas like Paris or Syberia. 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 [DDM*19] needs to be run on the given data. This should take at most a few hours. As indicated in Case Study 1, 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 dataset 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 in this paper. An example of such an approach could be imagined in social media analysis when trying to collect a set of original incidents from Twitter postings. The different data sources would then be the users and the taxonomies could be either manually created by domain experts or the results of different hierarchical clusterings applied on the data [FZZ*15, ISB14].

9. Conclusion

We introduced *VEHICLE*, a web-based tool to validate and explore the hierarchical integration of conflict event data. These investigations were only possible in a basic 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 tool in an iterative process. *VEHICLE* consists of multiple linked views that allow switching between the analysis of all events from the different integrated datasets and only those events that were identified to represent the same original incident. For the identification, we adopted the matching algorithm MELTT [DDM*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.

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, explore subsets of events and export them for further analysis. At the same time, the results showed that due to its complexity, *VEHICLE* can be considered an expert tool. In addition, the evaluation provided us with directions on how to extend the tool in the future. We plan to follow them when releasing it to the broad audience.

10. Acknowledgements

This research was supported in part by the Carl-Zeiss Foundation. Open access funding enabled and organized by Projekt DEAL. [Correction added on 14 Oct 2021, after first online publication: Projekt Deal funding statement has been added.]

References

[AA06] ANDRIENKO N., ANDRIENKO G.: *Exploratory analysis of spatial and temporal data: a systematic approach*. Springer Science & Business Media, 2006. 2, 3, 5, 6, 7

[AA13] ANDRIENKO N., ANDRIENKO G.: Visual analytics of movement: An overview of methods, tools and procedures. *Information Visualization 12*, 1 (2013), 3–24. 3

[AAA*16] ALI M., AHSAN Z., AMIN M., LATIF S., AYYAZ A., AYYAZ M.: Id-viewer: a visual analytics architecture for infectious diseases surveillance and response management in pakistan. *Public health 134* (2016), 72–85. 2

[AAB*13] ANDRIENKO G., ANDRIENKO N., BOSCH H., ERTL T., FUCHS G., JANKOWSKI P., THOM D.: Thematic patterns in georeferenced tweets through space-time visual analytics. *Computing in Science Engineering 15*, 3 (2013), 72–82. 3

[AAFW17] ANDRIENKO G., ANDRIENKO N., FUCHS G., WOOD J.: Revealing patterns and trends of mass mobility through spatial and temporal abstraction of origin-destination movement data. *IEEE Transactions on Visualization and Computer Graphics 23*, 9 (2017), 2120–2136. 3

[AAM*10] ANDRIENKO G., ANDRIENKO N., MLADENOV M., MOCK M., POELITZ C.: Extracting events from spatial time series. In *14th International Conference Information Visualisation* (2010), pp. 48–53. 3

[AMST11] AIGNER W., MIKSCH S., SCHUMANN H., TOMINSKI C.: *Visualization of Time-Oriented Data*, 1st ed. Springer, 2011. 4

[AWR*07] ANDRÉ P., WILSON M. L., RUSSELL A., SMITH D. A., OWENS A., SCHRAEFEL M.: Continuum: designing timelines for hierarchies, relationships and scale. In *Proc. of ACM symposium on User interface and software technology* (2007), pp. 101–110. 6

[BBY03] BHAVSAR V. C., BOLEY H., YANG L.: A weighted-tree similarity algorithm for multi-agent systems in e-business environments. In *Computational Intelligence* (2003), pp. 53–72. 6

[BD05] BALZER M., DEUSSEN O.: Voronoi treemaps. In *IEEE Symp. Information Visualization* (2005), pp. 49–56. 3

[BDA*17] BACH B., DRAGICEVIC P., ARCHAMBAULT D., HURTER C., CARPENDALE S.: A descriptive framework for temporal data visualizations based on generalized space-time cubes. *Computer Graphics Forum 36*, 6 (2017), 36–61. 2

[Bil05] BILLE P.: A survey on tree edit distance and related problems. *Theoretical computer science 337*, 1-3 (2005), 217–239. 6

[BKH05] BENDIX F., KOSARA R., HAUSER H.: Parallel sets: visual analysis of categorical data. In *IEEE Symp. Information Visualization* (2005), pp. 133–140. 7

[BW14] BURCH M., WEISKOPF D.: *On the Benefits and Drawbacks of Radial Diagrams*. Springer, 2014, pp. 429–451. 8

[CCF95] CARPENDALE M. S. T., COWPERTHWAITE D. J., FRACCHIA F. D.: 3-dimensional pliable surfaces: For the effective presentation of visual information. In *Proc. of ACM symposium on User interface and software technology* (1995), pp. 217–226. 6

[CLZ*18] CAO N., LIN C., ZHU Q., LIN Y., TENG X., WEN X.: Voila: Visual anomaly detection and monitoring with streaming spatiotemporal data. *IEEE Transactions on Visualization and Computer Graphics 24*, 1 (2018), 23–33. 3

[CSB*20] CAKMAK E., SCHÄFER H., BUCHMÜLLER J., FUCHS J., SCHRECK T., JORDAN A., KEIM D. A.: Motionglyphs : Visual abstraction of spatio-temporal networks in collective animal behavior. *Computer Graphics Forum 39*, 3 (2020). 9

[CSC*18] CHOI M., SHIN S., CHOI J., LANGEVIN S., BETHUNE C., HORNE P., KRONENFELD N., KANNAN R., DRAKE B., PARK H., CHOO J.: Topicontiles: Tile-based spatio-temporal event analytics via exclusive topic modeling on social media. In *Proc. of the 2018 CHI Conference on Human Factors in Computing Systems* (2018), p. 1–11. 3

[CTB*12] CHAE J., THOM D., BOSCH H., JANG Y., MACIEJEWSKI R., EBERT D. S., ERTL T.: Spatiotemporal social media analytics for abnormal event detection and examination using seasonal-trend decomposition. In *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)* (2012), pp. 143–152. 3

[DBBCM04] DE BERG M., BOSE P., CHEONG O., MORIN P.: On simplifying dot maps. *Computational Geometry 27*, 1 (2004), 43–62. 2

[DDM*19] DONNAY K., DUNFORD E. T., MCGRATH E. C., BACKER D., CUNNINGHAM D. E.: Integrating conflict event data. *Journal of Conflict Resolution 63*, 5 (2019), 1337–1364. 2, 3, 4, 10

[DF14] DONNAY K., FILIMONOV V.: Views to a war: Systematic differences in media and military reporting of the war in iraq. *EPJ Data Science 3* (2014), 25. 2

[DGB14] DONNAY K., GADJANOVA E., BHAVNANI R.: *Disaggregating Conflict by Actors, Time, and Location*. Paradigm, 2014, pp. 44–56. 2

[ED07] ELLIS G., DIX A.: A taxonomy of clutter reduction for information visualisation. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (2007), 1216–1223. 2, 6

[FG05] FABRIKANT S. I., GOLDSBERRY K.: Thematic relevance and perceptual salience of dynamic geovisualization displays. In *Proc. of International Cartographic Conference* (2005). 2

[FZZ*15] FENG W., ZHANG C., ZHANG W., HAN J., WANG J., AGGARWAL C., HUANG J.: 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. 10

[GAA04] GATALSKY P., ANDRIENKO N., ANDRIENKO G.: Interactive analysis of event data using space-time cube. In *Proc. of Eighth International Conference on Information Visualisation (IV'04)* (2004), pp. 145–152. 2

[GS62] GALE D., SHAPLEY L. S.: College admissions and the stability of marriage. *The American Mathematical Monthly 69*, 1 (1962), 9–15. 4

[GS03] GIUNCHIGLIA F., SHVAIKO P.: Semantic matching. *The Knowledge Engineering Review 18*, 3 (2003), 265–280. 10

[GTD13] Global terrorism database. START (National Consortium for the Study of Terrorism and Responses to Terrorism), 2013. (Accessed April 20, 2018.) [Online]. Available: http://www.start.umd.edu/gtd. URL: `http://www.start.umd.edu/gtd`. 2, 3

[Guo09] GUO D.: Flow mapping and multivariate visualization of large spatial interaction data. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (2009), 1041–1048. 3

[IP14] IZAKIAN H., PEDRYCZ W.: Anomaly detection and characterization in spatial time series data: A cluster-centric approach. *IEEE Transactions on Fuzzy Systems 22*, 6 (2014), 1612–1624. 3

[ISB14] IFRIM G., SHI B., BRIGADIR I.: Event detection in twitter using aggressive filtering and hierarchical tweet clustering. In *Second Workshop on Social News on the Web (SNOW)* (2014), ACM. 10

[JS91] JOHNSON B., SHNEIDERMAN B.: Tree-maps: A space-filling approach to the visualization of hierarchical information structures. In *IEEE Proc. of Visualization* (1991), IEEE Computer Society Press, pp. 284–291. 3, 7

[KF14] KAASTRA L. T., FISHER B.: Field experiment methodology for pair analytics. In *Proc. of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization* (2014), p. 152–159. 9

[KH98] KEIM D. A., HERRMANN A.: The gridfit algorithm : An efficient and effective approach to visualizing large amounts of spatial data. In *IEEE Proc. of Visualization* (1998), pp. 181–188. 2

[KL83] KRUSKAL J. B., LANDWEHR J. M.: Icicle plots: Better displays for hierarchical clustering. *The American Statistician 37*, 2 (1983), 162–168. 3, 7

[KPS03] KEIM D. A., PANSE C., SIPS M.: Visual data mining of large spatial data sets. In *International Workshop on Databases in Networked Information Systems* (2003), pp. 201–215. 2

[Kra03] KRAAK M.-J.: The space-time cube revisited from a geovisualization perspective. In *21st International Cartographic Conference* (2003), pp. 1988–1996. 2

[LCZ*19] LI J., CHEN S., ZHANG K., ANDRIENKO G., ANDRIENKO N.: Cope: Interactive exploration of co-occurrence patterns in spatial time series. *IEEE Transactions on Visualization and Computer Graphics 25*, 8 (2019), 2554–2567. 3

[LD11] LLOYD D., DYKES J.: Human-centered approaches in geovisualization design: Investigating multiple methods through a long-term case study. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (2011), 2498–2507. 5

[LSB*16] LU Y., STEPTOE M., BURKE S., WANG H., TSAI J., DAVULCU H., MONTGOMERY D., CORMAN S. R., MACIEJEWSKI R.: Exploring evolving media discourse through event cueing. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 220–229. 3

[Mac82] MACEACHREN A. M.: Map complexity: Comparison and measurement. *The American Cartographer 9*, 1 (1982), 31–46. 2

[Mac95] MACEACHREN A.: *How maps work: representation, visualization and design*. Guilford Press, 1995. 2

[Mun09] MUNZNER T.: A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (2009), 921–928. 2, 4, 5

[Ric48] RICHARDSON L. F.: Variation of the frequency of fatal quarrels with magnitude. *Journal of the American Statistical Association 43*, 244 (1948), 523–546. 2

[RLHK10] RALEIGH C., LINKE A., HEGRE H., KARLSEN J.: Introducing ACLED: An armed conflict location and event dataset: Special data feature. *Journal of Peace Research 47*, 5 (2010), 651–660. 2, 3

[RPP*17] ROBINSON A. C., PEUQUET D. J., PEZANOWSKI S., HARDISTY F. A., SWEDBERG B.: Design and evaluation of a geovisual analytics system for uncovering patterns in spatio-temporal event data. *Cartography and Geographic Information Science 44*, 3 (2017), 216–228. 3

[RT81] REINGOLD E. M., TILFORD J. S.: Tidier drawings of trees. *IEEE Transactions on Software Engineering SE-7*, 2 (1981), 223–228. 3

[SHH*12] SALEHYAN I., HENDRIX C. S., HAMNER J., CASE C., LINEBARGER C., STULL E., WILLIAMS J.: Social conflict in africa: A new database. *International Interactions 38*, 4 (2012), 503–511. 2, 3

[Shn97] SHNEIDERMAN B.: Direct manipulation for comprehensible, predictable and controllable user interfaces. In *Proc. of the 2nd International Conference on Intelligent User Interfaces* (1997), Association for Computing Machinery, p. 33–39. 5

[SHS10] SCHULZ H.-J., HADLAK S., SCHUMANN H.: The design space of implicit hierarchy visualization: A survey. *IEEE Transactions on Visualization and Computer Graphics 17*, 4 (2010), 393–411. 3

[SM13] SUNDBERG R., MELANDER E.: Introducing the UCDP georeferenced event dataset. *Journal of Peace Research 50*, 4 (2013), 523–532. 2, 3

[SMKH99] SLOCUM T. A., MCMASTER R., KESSLER F., HOWARD H.: *Thematic cartography and visualization*. Prentice hall Upper Saddle River, NJ, 1999. 2

[SS06] SCHULZ H. ., SCHUMANN H.: Visualizing graphs - a generalized view. In *Proc. of Tenth International Conference on Information Visualisation (IV'06)* (2006), pp. 166–173. 3

[SZ00] STASKO J., ZHANG E.: Focus+context display and navigation techniques for enhancing radial, space-filling hierarchy visualizations. In *IEEE Symp. Information Visualization* (2000), pp. 57–65. 3

[TAS04] TOMINSKI C., ABELLO J., SCHUMANN H.: Axes-based visualizations with radial layouts. In *Proc. of ACM Symposium on Applied Computing* (2004), p. 1242–1247. 6

[Tob73] TOBLER W. R.: Choropleth maps without class intervals. *Geographical analysis 5*, 3 (1973), 262–265. 3

[TS20] TOMINSKI C., SCHUMANN H.: *Interactive Visual Data Analysis*. CRC Press, 2020. 6

[TSAA12] TOMINSKI C., SCHUMANN H., ANDRIENKO G., ANDRIENKO N.: Stacking-based visualization of trajectory attribute data. *IEEE Transactions on Visualization and Computer Graphics 18*, 12 (2012), 2565–2574. 6

[Tuf01] TUFTE E. R.: *The visual display of quantitative information*, vol. 2. Graphics press Cheshire, CT, 2001. 3, 5

[VV99] VAN WIJK J. J., VAN DE WETERING H.: Cushion treemaps: visualization of hierarchical information. In *IEEE Symp. Information Visualization* (1999), pp. 73–78. 3

[Wei15] WEIDMANN N. B.: On the accuracy of media-based conflict event data. *Journal of Conflict Resolution 59*, 6 (2015), 1129–1149. 2

[Wei16] WEIDMANN N. B.: A closer look at reporting bias in conflict event data. *American Journal of Political Science 60*, 1 (2016), 206–218. 2

[YKT05] YANG R., KALNIS P., TUNG A. K. H.: Similarity evaluation on tree-structured data. In *Proc. of ACM SIGMOD International Conference on Management of Data* (2005), p. 754–765. 6

[Yu06] YU H.: Spatio-temporal gis design for exploring interactions of human activities. *Cartography and Geographic Information Science 33*, 1 (2006), 3–19. 3

[ZFH08] ZHAO J., FORER P., HARVEY A. S.: Activities, ringmaps and geovisualization of large human movement fields. *Information Visualization 7*, 3-4 (2008), 198–209. 6