# Narrative Maps: An Algorithmic Approach to Represent and Extract Information Narratives

BRIAN KEITH NORAMBUENA*, Virginia Tech, USA

TANUSHREE MITRA, University of Washington, USA

Narratives are fundamental to our perception of the world and are pervasive in all activities that involve the representation of events in time. Yet, modern online information systems do not incorporate narratives in their representation of events occurring over time. This article aims to bridge this gap, combining the theory of narrative representations with the data from modern online systems. We make three key contributions: a theory-driven *computational representation* of narratives, a novel *extraction algorithm* to obtain these representations from data, and an *evaluation* of our approach. In particular, given the effectiveness of visual metaphors, we employ a route map metaphor to design a narrative map representation. The narrative map representation illustrates the events and stories in the narrative as a series of landmarks and routes on the map. Each element of our representation is backed by a corresponding element from formal narrative theory, thus providing a solid theoretical background to our method. Our approach extracts the underlying graph structure of the narrative map using a novel optimization technique focused on maximizing coherence while respecting structural and coverage constraints. We showcase the effectiveness of our approach by performing a user evaluation to assess the quality of the representation, metaphor, and visualization. Evaluation results indicate that the Narrative Map representation is a powerful method to communicate complex narratives to individuals. Our findings have implications for intelligence analysts, computational journalists, and misinformation researchers.

## 1 INTRODUCTION

> *People don't see the world before their eyes until it's put in a narrative mode.*
>
> - Filmmaker Brian De Palma [1]

Narratives are fundamental to our understanding of the world [1] and are central to human relations [38]. They are the frameworks that enable humans to associate otherwise unconnected events [7] and play a key role in collaborative sensemaking in society [4, 67]. For decades, communication science scholars have emphasized the importance of narratives in human communication

---

*Also with Universidad Católica del Norte, Department of Computing & Systems Engineering.

---

Authors' addresses: Brian Keith Norambuena, Virginia Tech, Department of Computer Science, 620 Drillfield Drive, Blacksburg, VA, USA, 24061, briankeithn@vt.edu; Tanushree Mitra, University of Washington, The Information School, Seattle, WA, USA, tmitra@uw.edu.

---

[1]. Narrative theorists have gone so far to claim that "narrative is the principled way in which our species organizes its understanding of time."

In this line of research, narratives are defined as a *coherent system of interrelated stories* [25] and are considered to be omnipresent in all activities that involve the representation of events in time. Through these systems of stories, humans produce a shared understanding of the world [62].

Despite the compelling case of narratives in human communication, modern-day online information systems do not incorporate narratives in their representation of events happening over time. Instead, most representations are still chronological order of events from the earliest to the latest; for example, consider Google News or Twitter's news feed.

This paper aims to bridge this gap, merging the theory behind narrative representations with data left in modern online systems (specifically news systems) while leveraging graph-driven approaches to identify such narrative structures. With theory as a guide, our central question is: Can we computationally represent and extract narratives from large-scale online data?

Our contribution is not just methodological; narrative representation has the potential to significantly impact online information consumers. Consider the unrelenting barrage of online news articles that users get exposed to or the ever-increasing problem of information overload in online systems [27, 53]. Not just regular users, even expert news investigators find it daunting to keep track of the evolving storylines, often losing track of the big picture. Given narratives are the primary mode in which humans apprehend time-based information [1], our paper offers just such a representation and the algorithmic approach to surface them directly from data.

In particular, we use a route map metaphor to provide an intuitive basis for our representation. This metaphor helps connect the elements of narrative theory and graph theory. Building upon this representation (RQ1), we propose a novel extraction algorithm to generate a narrative map representation automatically from data (RQ2). Our algorithmic approach leverages optimization techniques and graph-theoretical concepts to find all the elements in our narrative representation.

Specifically, our work addresses three research questions. First, we formalize the computational representation of narratives and ask:

**RQ1 Representation**: How can we computationally represent a narrative?

Based on a route map metaphor, we define a computational representation of a narrative grounded in formal narrative theory. Our representation maps concepts from narrative theory into graph theory elements. In particular, our proposal considers the use of a directed acyclic graph since this structure naturally reflects the temporal structure of stories. We also consider edge weights that allow us to evaluate the coherence of the storylines in the narrative map. Note that a narrative is composed of at least one storyline. After identifying the elements of our representation, we focus on how to extract these elements from data.

**RQ2 Extraction**: How can we extract a narrative representation from data?

To extract the narrative representation, we design an optimization approach that seeks to maximize the coherence of the narrative map subject to structural and topic coverage constraints. After generating the basic structure of the narrative map, we extract the main route (the most coherent path in the narrative) and the representative landmarks (sets of representative events from each alternative storyline). Using this information we can infer the themes—the recurrent or prominent ideas—of the narrative. Finally, we evaluate how well our proposed approach works.

**RQ3 Evaluation**: How well does our approach work to represent narratives?

We present a detailed evaluation of our proposed approach. In particular, we study the news narrative of the Coronavirus outbreak with three maps from the first three months of 2020, each generated after processing hundreds of news articles. We find that our representation provides a

coherent narrative and that our metaphors are appropriate to model the narrative. We also show how our representation provides an easy way for users to answer questions such as: What is the main storyline—the main route—of this narrative? Apart from the main storyline, what other alternative storylines are there in the narrative? What are the major themes in the narrative?

Thus, by answering these research questions, our key contributions are:

- A **computational representation** for narratives, based on a mapping between narrative theory concepts and computational elements grounded in the route map metaphor. Our representation also includes a visualization based on the route map metaphor.
- An **extraction algorithm** to construct a narrative representation from data while using an optimization method and graph-theoretical concepts.
- An **evaluation** of our approach through a case study and a user study. In particular, we test our methods on the news narrative of the first three months of the Coronavirus outbreak.

Furthermore, we note that our work builds upon the foundation created by existing techniques. However, our proposed representation extends them by providing a theoretical framework based on formal narrative theory and the route map metaphor. The inclusion of this framework distinguishes our work from similar methods.

We believe that our approach has widespread implications. At its core, our narrative representation and our extraction method can help users understand the big picture of a narrative. Moreover, expert users, such as journalists, fact-checkers, and intelligence analysts can leverage our extraction method to explore the narrative landscape in the context of their domain of work. Furthermore, our narrative representation could be enhanced by considering additional elements, such as contextual information, credibility and political bias annotations, or more detailed event descriptors. In particular, adding credibility and political bias information can help understand how misinformation spreads in news narratives.

In the rest of this section, we introduce the route map metaphor and provide a motivating example, explaining the reasoning behind the metaphor and detailing how it can be applied to analyze a narrative. In Section 2, we introduce important background concepts and describe related works. In Section 3, we show our algorithmic approach, including both the computational representation of narratives and the extraction algorithm itself. In Section 4, we present a case study of the Coronavirus narrative. Section 5 presents a user evaluation of our approach followed by a discussion of our results in Section 6. Finally, we end with the main conclusions of our work.

## 1.1 Route Map Metaphor

Humanity has relied on maps and cartography to understand its physical surroundings since antiquity. In particular, route maps and similar cartographic representations have been used to describe routes, that is, how to get from one physical place to another. These routes are usually accompanied by information of places and important locations along the way, as well as additional information about relatively less important elements of its surroundings. There can also be many routes joining two places, with differing characteristics and importance.

In general, the idea of using maps to represent stories is called information cartography [58]. In this paper, we propose a way to represent a narrative (i.e., a system of stories) using a narrative route map (or simply narrative map). Instead of joining two physical places, we join two narrative places (a starting and ending event). The locations shown along the routes in our map should correspond to narrative elements.

Consider a route map that shows routes between two physical places as shown in Figure 1(a). The routes connect the "Start Stadium" and the "Classroom Hall" landmarks. Suppose a student needs to go to his engineering class at Classroom Hall after watching a practice session at Start
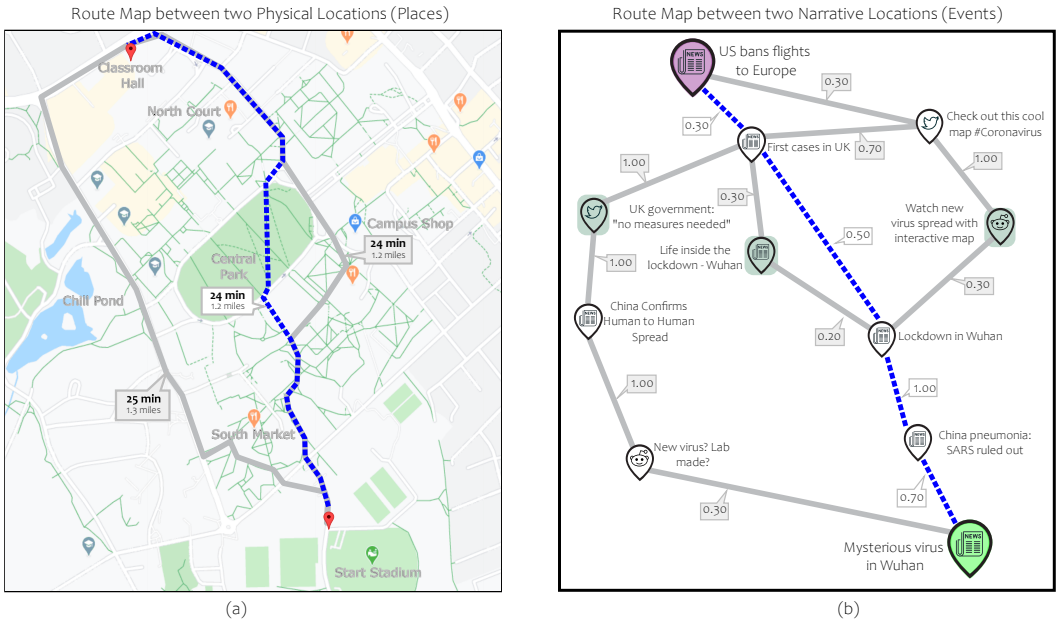
Fig. 1. Illustration of the Route Map Metaphor. (a) A route map showing how to get from Start Stadium to Classroom Hall. (b) A narrative route map showing how to get from the Mysterious Virus in Wuhan event (starting event ⬦ ) to the US banning flights to Europe (ending event ⬦ ). We highlight some representative landmarks ⬦ for each route in the narrative map.

Stadium. However, they are not sure which route to take, so they would take out their phone, enter their current location, and where they want to go into the route map application. The application then provides them with a route map as shown in Figure 1. We can identify multiple elements of interest in this map. In particular, it contains multiple interconnected *routes* between the *starting place* (Start Stadium) and the *ending place* (Classroom Hall).

One of these routes is highlighted as the *main route* (in blue) because it is the shortest path. All routes contain information about its distance and the time it takes to go through them, this allows users to decide which one to take (i.e., it is an *evaluation metric*), although in this particular example there is not much difference between the three routes.

We can also see information about different kinds of *landmarks* throughout each route. We note that there are several types of landmarks (e.g., restaurants, shops, parks, and educational buildings). For example, the main route passes close to the South Market and goes through the Central Park, and then North Court before getting to the Classroom Hall.

In fact, we could identify each route by mentioning a *representative landmark*, such as the Chill Pond, the Central Park, and the Campus Shop for each route in our case. By identifying these representative landmarks, we can differentiate among these routes. Someone using the route map could make choices informed by these representative landmarks. For example, someone seeking to walk through a more scenic view might choose to go through the Chill Pond route, while someone interested in buying a snack or some materials could go through the Campus Shop route. Even though neither of these routes is the main one they still offer a potentially useful alternative.

From the previous analysis we identify the following key elements of a route map: a starting place, an ending place, routes joining these places, landmarks along those routes, representative

landmarks for parallel routes, an evaluation metric (e.g., distance or time) for those routes, and the main route obtained using that metric. Using these elements, we construct a metaphorical representation of narratives with a route map. Just as physical landscapes can be represented using these route maps, we use the same elements to construct a route map representation of the narrative landscape surrounding a certain event.

According to some authors [1], narratives must contain at least two events: a starting event 🟢 and ending event 🟣. These events work similarly to the starting and ending place on the map and we need to choose them before creating the narrative map. We are interested in understanding how these two events are joined together by a story (i.e., a series of events and discourse elements). Thus, events would be akin to the landmarks of each route ⬭, and discourse elements would be the additional information associated with them. Furthermore, narratives can be seen as a system of stories, which in our metaphor correspond to the routes on the map ╱. By doing this we gather new insights about the original events and the surrounding themes of the underlying stories. In addition, we can identify and characterize these storylines by finding a representative landmark for them 🟢. Finally, we can evaluate the stories 🔲 (e.g., by looking at its coherence) to find the main storyline of the narrative ⋰ (i.e., the main route).

See for example Figure 1(b), which shows a metaphoric rendition of a narrative route map representing the first month of the Coronavirus outbreak timeline. This map contains several sources of information (news and social media) and provides an overview of the evolving narrative landscape during this period. The dashed line ⋰ represents the main route (i.e., storyline). It reflects the development of the outbreak from its first cases as a mysterious virus to the US banning flights.

Various other elements are intertwined with this route, representing alternative paths towards the ending event. Representative landmarks 🟢, in this case, could be the tweets about the UK government indicating that no measures are needed, the news article about the life inside the lockdown in Wuhan, and the submission on Reddit about watching how the virus spreads.

As in the physical route map, these landmarks can inform the user of the map on where to go. For example, a data analyst could be interested in reading about visualizations related to the virus as examples of data visualizations. This analyst would presumably want to go through the interactive map route. A social science researcher might be interested in looking at how news sources presented life during lockdown at the start of the crisis. Therefore, they would likely be interested in reading the lockdown route. Finally, a political activist from the UK could be interested in reading about the UK response to the virus to build a case against the government's handling of the outbreak. Thus, they would probably find the route on the UK response more interesting.

In one of the paths we see conspiracy talks about the virus being artificial, then news about human to human spreading, leading into the spread into other countries. In the continuation of these paths, we see how the lack of appropriate government measures leads to the spread of the virus. Some paths are simply smaller deviations with interesting tidbits of information, such as the article on life during lockdown and people sharing an interactive map in social media.

This route map metaphor forms the intuitive basis of our computational representation and corresponding visualization. However, to give our intuition a stronger foundation, we associate each of the elements of our representation with a formal concept from narrative theory.

## 2 BACKGROUND AND RELATED WORK

In this section, we introduce important background concepts from narrative theory and present research related to our main contributions. We start with online narratives, then we move onto

narrative theory definitions, and finally, we describe existing approaches to extract and represent narratives and stories.

## 2.1 Online Narratives

The exponential growth of the World Wide Web and the improved accessibility of information has exacerbated the problem of information overload in our daily lives [27]. In particular, this becomes an issue when trying to process and understand the barrage of news events [53]. Furthermore, the rise of social media and its subsequent place as a staging ground for information operations [50, 65] has given place to the creation of alternative narratives [44] and the spread of fake news [11].

Moreover, narratives play a key role as one of the basic building blocks of societal sensemaking [67] and political actors attempt to use strategic narratives to create a shared understanding of the world [39]. Thus, given the importance of narratives in society, it is critical to understand how they emerge and evolve over time in this environment of information overload and fake news.

Existing work has focused on algorithms to extract timelines [53], trees [3], or other variants of graphs to represent stories [71]. However, they have mostly neglected the theoretical foundations of narratives on their models. The present work is an attempt to bridge this theoretical gap by creating a computational representation model for narratives grounded in formal narrative theory. In particular, the proposed model is based on a route map metaphor and it links elements from narrative theory and graph theory in an intuitive way.

## 2.2 Narrative Theory

The design of our narrative representation is informed by concepts from general narrative theory and strategic narrative theory. Both approaches contain elements that help us define and analyze our narrative representation.

General narrative theory focuses explicitly on understanding the general rules of narrative and its different arrangements that make it meaningful [1, 48]. The key intuition in formal narrative theory is that there is a distinction between the story itself and its representation. Narrative theory tries to understand the relationships between stories and their many possible representations [48].

In contrast, instead of focusing on the narrative itself, strategic narrative theory studies the construction and application of these narratives by political actors. In particular, strategic narrative theory focuses on studying how political actors communicate narratives in a strategic way to construct a shared meaning of events with the purpose of influencing the behavior of domestic and international actors [38, 51].

Another approach to narratives by Halverson et al. [25] defines narratives not just as one story, but rather as a system of stories. That is, narratives are a systematic collection of interrelated stories with coherent themes. We rely on this definition of narrative as a system of interrelated stories to build our narrative map extraction method. In particular, our method also ensures a certain degree of thematic coherence between the extracted storylines.

## 2.3 Existing approaches to extract stories and narratives

To inform the design of our extraction method, we refer to the existing story and narrative representations found in the literature. First, we discuss related works based on the level of detail (*representation resolution*), for example, events or topics in the narrative. Afterward, we analyze related works based on the internal structure of the representation (*representation complexity*), for example, a tree or a timeline.

*2.3.1 Representation Resolution.* Multiple resolution levels could be studied in the context of story and narrative visualization. In particular, most of the story visualization research has focused

on either the event level [34, 53] or the topic level [29, 42, 73]. Event-level approaches focus on concrete sequences of events. In contrast, topic-level approaches are more abstract. Thus, instead of focusing on the specific events that happen throughout the story, they focus on the overarching topics and how they relate to each other.

In particular, previous studies have used graph representations based on event clustering to represent topics [42]. Our method builds upon the idea of using clusters of events to represent topics, in particular, we use them to extract coherent storylines aligned with these clusters.

Some authors have also proposed more fine-grained resolution levels to represent stories, such as individual named entities [16], claims and attributions made in an article [61], or mixed resolution approaches that allow zooming between the different levels, such as the information cartography approach from Shahaf et al. [58]. In contrast to these fine-grained approaches, we focus on the event level to build our narrative map representation. By choosing this resolution level, we are able to easily develop a direct mapping between narrative and computational elements.

*2.3.2 Representation Complexity.* There exist many approaches to visualize the evolution of news events. The simplest cases correspond to a linear representation of the events (i.e., simple sequences of events), such as timelines [64, 70] and coherent story chains [53, 54]. All these representations consist of a single starting and ending event with a sequence of events joining them.

Previous research on timeline and story extraction relies on optimizing different criteria to select events. Furthermore, these criteria could depend on the type of data used by the extraction method. For example, in the socially informed timelines approach by Wang et al. [64], their optimization method seeks to balance topical cohesion between articles and comments in social media. In contrast, in the news timeline summarization method of Yan et al. [70], the optimization algorithm tries to balance relevance, coverage, coherence, and cross-date diversity. Another approach corresponds to the Connect-the-Dots algorithm by Shahaf and Guestrin [53], which creates story chains using a linear programming approach to maximize the coherence of the timeline. An improved version of their approach [54] used integer programming to generate the story chain. Building upon the intuition of these previous works, our approach also relies on an optimization problem to select the events that appear in our narrative representation.

A logical extension to the single timeline approach is to consider more than one timeline. These timelines can be shown in parallel in the story map. Such as in the work done by Xu et al. [69], where they extract multiple storylines surrounding a news event. Due to the complexity of generating such maps by brute force, they develop a near-optimal algorithm based on locality-sensitive hashing to perform dimensionality reduction. They display all the storylines in the same story map as parallel timelines. In contrast to this approach, our method allows for the different storylines in the narrative to interact with one another, allowing them to merge or separate into new storylines. This allows a more appropriate representation of narratives as a system of interrelated stories, rather than disjointed timelines.

More advanced cases use directed acyclic graph structures, such as event evolution graphs [71, 74], and metro maps [55, 58]. In these cases, the maps can have multiple starting and ending events. Furthermore, these structures allow for more advanced local substructures, such as convergence (merging stories) and divergence (separating stories). Scholars have studied different models and features to represent and extract the relationships between events in these graphs. For example, the event evolution graph of Yang et al. [71] uses event timestamps, event content similarity, temporal proximity, and document distributional proximity to model relationships. In contrast, Zhou et al. [74] use a model based on term and event frequency, alongside event sequences, content similarity, and temporal distance costs. Another approach corresponds to optimizing the connectivity of the graph, subject to coherence and coverage constraints [55]. Using these We model our event

| Elements | | | Description |
|---|---|---|---|
| Narratological | Computational | Metaphorical | |
| Events | Nodes | Landmarks | Events are actions or happenings that are part of the narrative. |
| Discourse Elements | Node Attributes | Landmark Attributes | Discourse Elements represent additional or complementary information about the events that are not an event of the story (e.g., the news source that published the event). |
| Starting Event | Source | Starting Place | The first event of a story. All stories must begin at one of these events and there must be at least one such event on the map. |
| Ending Event | Sink | Ending Place | The last event of a story. All storylines terminate in an end event and there must be at least one such event on the map. |
| Story | Chain (Path) | Route | A story is a sequence of events and discourse elements. Under the definition of narrative as a system of stories, there can be multiple stories in the same narrative. |
| Coherence | Edge Weights | Evaluation Metric | Coherence represents how much sense it makes to join two events and serves as an evaluation metric for the route. |
| Constituent Events | Maximum Likelihood Chain | Main Route | Constituent events form the core events of the narrative and are part of the main storyline. |
| Supplementary Events | Maximum Antichain | Representative Landmarks | Supplementary events correspond to events that are not the core part of the narrative but that provide further information and are part of alternative storylines. |
| Narrative | st-graph | Route Map | The narrative is a system of interrelated stories, we represent the narrative as an st-graph. |

**Visual Representation**



Table 1. Mapping of narrative theory elements and computational elements alongside their route map metaphor equivalent. The concepts shown here are used interchangeably throughout the paper to describe our algorithmic approach. Narratives have at least one storyline, and a storyline has at least two events. Thus, we have the hierarchy: event < storyline ≤ narrative.

relationships using a similarity-based approach in conjunction with clustering information to ensure topical coherence and a certain degree of coverage.

Finally, we also have tree-based approaches to represent stories. For example, the tree-based representation of Ansah et al. [3] leverages information propagation among users in a community, temporal proximity, and semantic context to construct coherent paths and build a timeline summary. In addition, the news content organization system of Liu et al. [34] build a forest representation using keyword extraction, event clustering, story clustering, coherence, and time-based penalties.

## 3 AN ALGORITHMIC APPROACH TO REPRESENT AND EXTRACT NARRATIVES

We offer a graph-driven approach to represent and find narratives from large troves of unstructured data corresponding to real-world events. We start by showing our computational representation of narratives (RQ1). Next, we present our algorithmic approach to extract these narratives automatically from data (RQ2).

### 3.1 RQ1: Narrative Representation

We offer a conceptual representation of narratives while synthesizing elements from the narrative theory literature and following the route map metaphor. Table 1 lays out the key elements in a narrative, their metaphoric equivalents, and their respective computational representations.

*3.1.1 What are the key elements in a narrative?* Synthesizing across theories of narratives, we compile eight elements fundamental toward forming a narrative. Most of these elements are based on or related to the concept of events. For example, stories are sequences of events, coherence represents how much sense it makes to join events, and constituent events are a special kind of event that is fundamental to the progress of the narrative. All these elements rely on the basic concept of event for their definition, but they represent a specific narratological concept [1]. The narrative itself is computationally represented as a graph, with each narrative element corresponding to a building block of this graph.

*Story, Events, and Discourse Elements.* According to Abbott's general narrative theory [1], narratives are composed of two primary elements: *story* and *discourse*. The story is a sequence of *events* involving entities and follows a chronological order. Thus *events* are one of the fundamental building blocks of a *story*. Meanwhile, the narrative *discourse* is the story as narrated, that is, the rendering of the story in a particular narrative (in other words, a representation of the story). While *events* correspond to actions or happenings in the story, *discourse* elements are part of the representation of the story. Discourse elements provide additional framing or complementary information about the events that influence the way the narrative is told but are not events themselves. Examples include public reactions to an event on social media or attributes of an event, such as the credibility of the source reporting the event or bias in reporting.

Thus, in formalizing the computational equivalent of these elements, we represent events as nodes in a graph, discourse elements as node attributes (e.g., an attribute containing the event source), and the graph itself corresponds to the narrative map. We unpack the computational representation of a narrative by introducing two other elements—starting and ending events.

*Starting and Ending Events.* General narrative theory [1] further states that a narratological story is defined by a single starting and ending event. Some authors argue that only one event is needed to have a story [1, 21, 60], but many concur that at least two connected events are essential [5, 49, 63, 68]. We align with the popular view and require our narrative map to contain at least two events: *starting event* and *ending event*. We represent the starting event as a source in the graph—a node with no incoming edges. Likewise, we represent the ending event as a sink—a node with no outgoing edges. As per the route map metaphor, these two events correspond to the starting and ending places in a physical map, respectively. As in the physical map, the starting and ending events are selected by the user.

*Story Coherence and System of Stories.* What makes a story coherent? A narrative story comprising of events and discourse elements needs to flow together, i.e., it needs to be coherent [1]. For example, a narrative story that focuses on a certain topic should not drastically change to a different topic without appropriate events joining them. In other words, *coherence* represents how much sense it makes to join two events. Thus, we can computationally represent story coherence through the weight of the edge joining the two events, where events are represented as nodes in a graph; higher edge weight implies higher coherence. Starting with this representation, next in RQ2, we offer computational techniques to automatically find the edge weight and determine whether to join the two events. Story coherence, thus, corresponds to the evaluation metric of our route map metaphor.

Moreover, prior theoretical work posits "narrative as a system of stories" [25]. Following this definition, our narrative map should not only contain a single story but should represent a narrative as a system of stories. These *stories* must be interrelated and form coherent sequences of events.

*Constituent and Supplementary Events.* An important part of analyzing a narrative is distinguishing which events are the core and which ones simply provide complementary information to the narrative [1]. According to theories of narrative, *constituent events* represent the core. These are the

main events that drive the story forward and cannot be omitted without fundamentally altering the story. In contrast, *supplementary events* do not drive the main story forward and if removed, the main story would still remain intact. However, this does not imply that supplementary incidents are not necessary to understand the context and general impact of the narrative. Hence, we allow both constituent and supplementary events as key elements for our narrative representation.

In computationally formalizing these two narratological elements, we want the overall sequence of constituent events to make as much sense as possible. Thus, we represent constituent events as nodes along the most coherent path (i.e., the maximum likelihood path) in the narrative map that joins the starting event with the ending event. The most coherent path connecting the sequence of constituent events also provides us with the main storyline. In terms of the route map metaphor, the sequence of constituent events corresponds to the "main route" that has the best value for our evaluation metric (i.e., coherence) out of all the routes. For example, recall the blue route in Figure 1(b), which reflects the development of the outbreak from its first case as a mysterious virus to the US banning flights to Europe. Note in particular, that his route has a likelihood of $0.7 \times 1.0 \times 0.5 \times 0.3 = 0.105$, the highest out of all routes in the map.

For supplementary events, we focus on finding a representative subset of these events. To do this, we focus on all potential alternative storylines that join the starting event with the ending event. However, enumerating all potential alternative storylines is computationally impractical, except for the smallest of the narrative maps. Instead, we extract meaningful representative events from each alternative storyline in the narrative map. In terms of the route map metaphor, each alternative route can be associated with a representative landmark (i.e., a supplementary event). While these routes do not have the highest likelihood, they can still be viable alternatives and can offer different perspectives. For instance, recall the example from our route map metaphor (Figure 1(a)) where we elaborated alternative routes, such as a longer scenic route through the Chill Pond versus the preferred main route through the South Market.

*3.1.2 How can we define a structured way to navigate across the narrative elements?* Now that we have formalized the basic building blocks of the narrative representation, we define the underlying computational structure that relates them. Recall that a narrative map is computationally represented as a graph. However, to ensure that our representation aligns with the theoretical structure of narratives, we must specify additional constraints to this graph structure. According to narrative theory, narratives can arrange events in many ways [21], but the underlying stories are made up of *sequences* of events that follow *chronological order* [1]. To capture these properties, we use a directed acyclic graph (DAG) because the event nodes must be joined in sequences (directed) and they must follow chronological order (acyclic). Moreover, this structure also fits the requirements of the route map metaphor, which does not contain cycles and provides a sequential order for the landmarks in a route.

However, there is one caveat of having a general DAG with no additional constraints. A DAG can have multiple sources (i.e., nodes with no incoming edges) and sinks (i.e., nodes with no outgoing edges). Yet, our route map metaphor only contains one source and one sink (the starting and ending place). Thus, we need to find a graph structure that respects these constraints. Specifically, we impose two additional restrictions on our DAG representation: the graph must have exactly one source node and one sink node. This is a special kind of graph called an *st-graph* [28]—a directed acyclic graph with a single source and a single target (sink). See for example the graph of Figure 1(b), where the direction of the edges is implicitly given by the chronological order of the events. This representation allows us to encode a single system of stories connecting the ending and starting events (i.e., a single narrative map). By using an *st-graph* rather than a general directed acyclic graph, the representation allows us to understand how the events are connected without

considering unnecessary information about other starting and ending events. Furthermore, we incorporate the coherence measure as edge weights in our graph. Finally, this structure allows us to navigate throughout the elements of the narrative map, offering us a chronological view of all the events and storylines contained in it.

*3.1.3 Formalizing the narrative map definition.* Given our previous discussion, we propose the following formal definition for a narrative map:

DEFINITION 1. **Narrative map**: *We formally define a narrative map as a weighted directed acyclic graph $G = (D, E)$ with a single source $s \in D$ (the starting event) and a single sink $t \in D$ (the ending event), such that weights represent the probability of going from one node to another based on coherence and the paths from $s$ to $t$ represent the different storylines in the narrative.*

Figure 1(b) from the motivational example section shows an example of a weighted *st-graph* for a sample narrative, where the direction of the edges is implicitly given by the chronological order of the events. It should be noted that our representation is general enough, such that, if we modify some of the constraints, we can obtain other representations. For instance, a simple timeline (one that only allows sequential patterns, such as a single story chain [53]) or an event tree (one that allows at most one parent for each node) can be constructed from our generic representation. As a case in point, to convert our narrative map in Figure 1(b) to a narrative tree, we could require that nodes have at most one parent, relax the restriction of only one ending event, and then drop the less important edges. More complex representations, such as parallel timelines [69] or story forests [34], can also be obtained by combining the simpler representations.

## 3.2 RQ2: Narrative Map Extraction

Here we introduce our narrative map extraction method. We first define the narrative map problem formally before presenting the narrative extraction process. Figure 2 gives an overview.

*3.2.1 Formal Problem Definition.* Based on our previous formalization of a narrative map as a directed acyclic graph with a single source (the starting event) and a single sink (the ending event), we provide a formal definition of the narrative map problem.

DEFINITION 2. **Narrative map problem**: *Given a set of time-stamped documents $D$ representing events for a certain issue of interest (e.g., news articles for the Coronavirus outbreak), a starting event $s \in D$ and an ending event $t \in D$, our objective is to find a graph in the form of a narrative map $G = (D, E)$, where source $s$ and sink $t$ are connected (by potentially many paths).*

Since the narrative map problem has multiple potential solutions, we need to define the criteria to find an appropriate solution—a good narrative map.

*3.2.2 What makes a good narrative map?* To answer, we rely on the notion of narrative coherence [53, 55, 70]—how much sense it makes to join two events—and coverage [55, 58, 70]—how well the map covers the different topics in the narrative.

Inspired by previous works [34, 58, 64, 74], we base our measure of coherence on the following intuition: coherence should be high if the events are similar and if they share common topics. That is, in a narrative map when a user follows a chain of events along edges joining them, they should get a clear understanding of the underlying theme in the storyline. Therefore, to determine whether it makes sense to join two events with an edge on the narrative map graph, we need to compute both event similarity [74] and topic similarity [34, 58, 64].

Based on previous works [55, 56], we base our measure of coverage on the following intuition: coverage should be high if the map covers as many topics of the narrative as possible. Prior work has

relied on simple word frequency-based models and pre-defined keyword lists to compute coverage and give priority to important words [55]. However, a pre-defined keyword-based approach makes assumptions about how events will look a priori, potentially increasing the risk of selection bias [40]. Thus, to avoid this issue we use a more abstract approach by defining coverage with respect to clusters of events (representing topics) that are automatically extracted from data.

In general, we want narrative maps to be as coherent as possible while also covering the topics in an appropriate way. Thus, we regard a narrative map as good if it is maximally coherent for a given coverage requirement. This leads naturally to an optimization approach, which we cover in more detail in the next subsection.

*3.2.3 Narrative Extraction Process.* Following our previous discussion on what makes a good narrative map, we seek to maximize the coherence of the narrative map, subject to coverage and structural constraints. First, we give a high-level overview of our optimization method and the extraction process. Then, we provide additional details for each phase of our algorithmic approach starting from data extraction and ending with the final visualization.

*Extraction Process Overview.* We develop an optimization-based approach informed by prior scholarly work [53]. In particular, we employ a linear programming technique which automatically generates a narrative map by selecting a subset of events and connections between them, such that it maximizes coherence subject to two constraints: 1) coverage constraints to ensure that our solution has at least a certain level of coverage and 2) structural constraints to ensure that our solution has an st-graph structure.

Based on this optimization approach, our extraction process consists of 6 phases. The input corresponds to a data table with headlines, publication dates and URLs. The output corresponds to a visualization representing the narrative through the route map metaphor. In phase 1 we extract the data and perform basic preprocessing. In phase 2, we compute coherence information for each pair of events. Coherence values are then used as input to the linear program. In phase 3, we solve the linear program to obtain the final st-graph representation. Note that during this phase we compute coverage and ensure that our solution meets the required threshold. In phase 4, we extract the main route from the st-graph representation by finding the route with the highest coherence. In phase 5, we extract the representative landmarks from the st-graph representation using graph antichains—sets of pairwise incomparable nodes that can be used to represent the storylines. Finally, in phase 6, we generate the final narrative map and its visualization based on the route map metaphor.

*Phase 1: Extracting Data and Preprocessing.* Our data set of choice is a collection of news articles. News articles capture real-world information and are structured to provide the key event details at the very beginning in the news headline [45]. Thus, we retain headlines of the news articles as our unit of analysis. Each headline corresponds to an event in the narrative. Next, we perform basic data preprocessing steps, such as fixing character encoding issues.

*Phase 2: Computing Coherence.* Coherence comprises of two components—event similarity and topic similarity. First, to compute event similarity, we represent the events (corresponding news headlines in our case) using neural embeddings of the headlines. In particular, for each event, we used the pre-trained models from the Universal Sentence Encoder v4.0 from TensorFlow [9] to extract the document embeddings from their headlines. This representation is appropriate for our purpose because it has been trained on news data and is designed for short text [9]. After mapping news headlines on the document embedding space, we find semantically similar headlines by comparing the similarities between pairs of embeddings. Instead of traditional raw cosine similarity, we use angular distance-based similarity measure which distinguishes near parallel vectors better
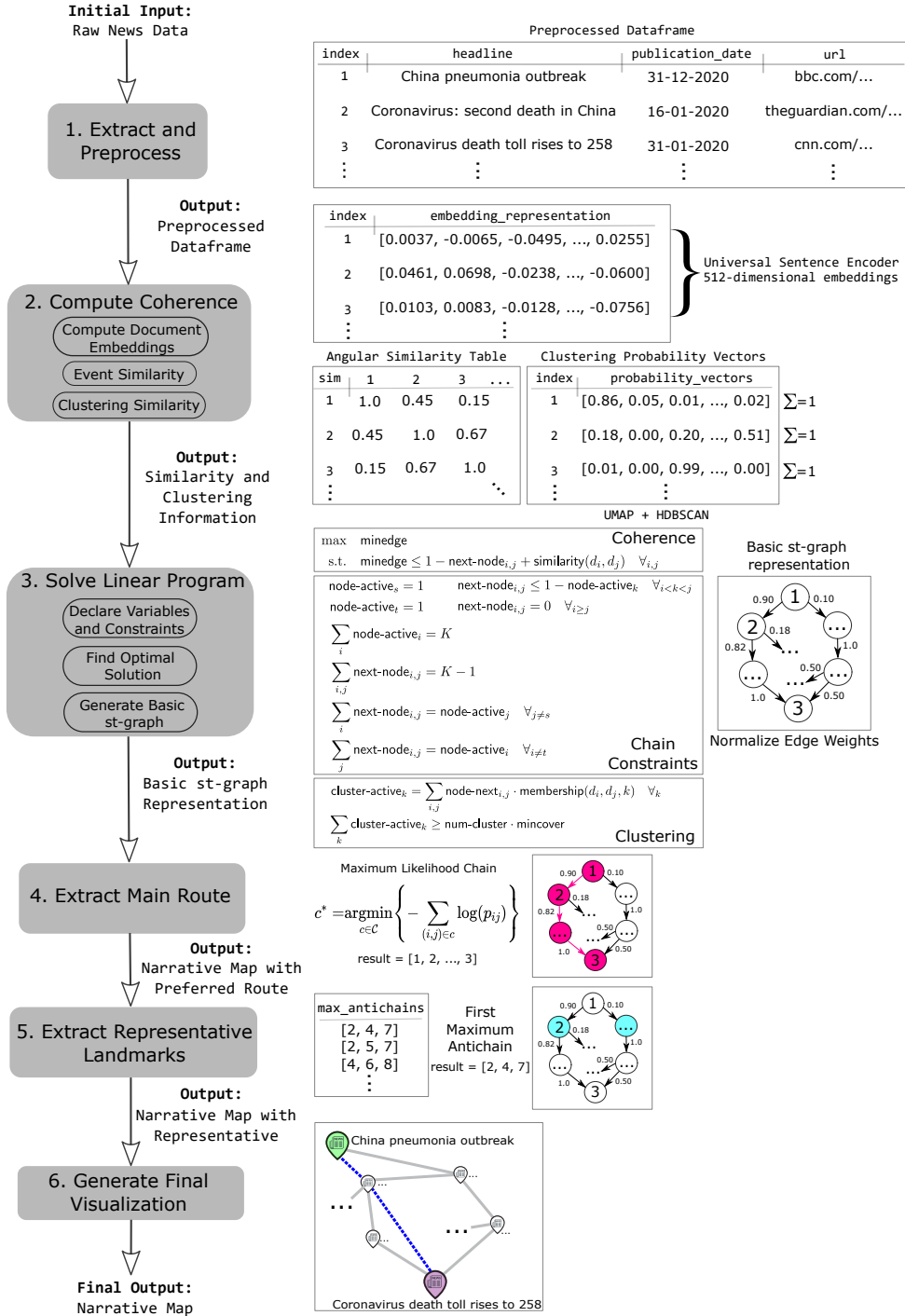
Fig. 2. Figure illustrating the steps in our narrative extraction algorithm that generates a narrative map from data (RQ2). Each phase is annotated with their respective outputs.

and is well suited when used with the Universal Sentence Encoder [9]. Angular similarity ranges between 0 and 1; values close to 1 denote similar events.

Next, to compute the topic similarity component, we obtain potential topics through clustering, where each cluster corresponds to an underlying topic. In particular, we cluster the documents using HDBSCAN [35] with UMAP [36]. HDBSCAN is an advanced density-based technique, but it scales poorly with the dimensionality (number of features) of the data. Thus, to improve the speed and performance of HDBSCAN, we use UMAP on the document embeddings for dimensionality reduction as a preprocessing step. In addition, an important advantage of HDBSCAN is that it does not require specifying the number of clusters (or topics) a priori, allowing us to automatically surface the most important topics directly from data. The step results in cluster probability vectors, containing the probability of whether a certain document belongs to each cluster.

Next, we need to compare the topic similarity of two events using their cluster probability distributions. We compare the cluster distribution vectors using the Jensen-Shannon similarity (JS-score)—a preferred method to compare probability distributions [23]. If the events have perfectly equal topic or cluster distribution vectors, we obtain a JS similarity score of 1 (e.g., both events have $[0.6, 0.3, 0.1]$ as its cluster distribution vector). If the events focus on completely different storylines or topics, the JS similarity score will be 0 (e.g. $[0.4, 0.6, 0]$ vs. $[0.0, 0.0, 1.0]$).

How can we combine the two components—event similarity and topic similarity—to obtain the final coherence score? A simple approach would be to compute the arithmetic mean. However, this approach is not strict enough for our purposes. We want a scoring mechanism that has strong penalties if any of the components are low and forces a strict balance them. Thus, we rely on the geometric mean (square root of their product), which imposes stricter penalties when a score is low. In particular, by formulating it through a geometric mean rather than a simple arithmetic mean, we ensure that the coherence will be zero if and only if any of the components are zero.

*Phase 3: Solving the Linear Program.* Using our coherence information from the previous step, we can now implement the linear program that maximizes coherence. The solution that we generate from this linear program is an *st-graph* that represents interrelated stories, thus fitting with the definition of narrative as a system of stories from Halverson et al. [25]. The mathematical formulation of the linear program is shown in the corresponding phase of Figure 2. Here we explain each component of this formulation in an intuitive way. The linear program seeks to optimize the strength of the weakest link in the narrative map in terms of coherence. Intuitively, this is based on the idea that a story is as coherent as its weakest link [53]. Furthermore, recall that narratives have a single starting event, a single ending event, and are formed by sequences of events ordered chronologically, represented as an *st-graph*. Thus, our optimization problem needs to include a series of structural constraints to ensure that our solution is an *st-graph*. Namely, it needs to set up a single source (starting event) and a single sink (ending event) and then force the events to be joined chronologically, These are all characteristics of a directed acyclic graph (DAG for short).

Finally, the last constraint—coverage constraints—ensures that a certain percent of the extracted topic clusters are covered. We found empirically that a minimum coverage of 80% worked well for our data in preliminary qualitative evaluations. Using this parametrization, we found that in general, the extracted storylines made sense and the covered topics were relevant. Furthermore, we note that coverage, unlike coherence, can not be pre-computed because it can change with each iteration of the linear program optimization process (different potential solutions of the linear program can yield different coverage values).

To compute coverage, we rely on the following intuition: a topic is covered if it appears on coherent sequences of events. Following this intuition, nodes on their own are not enough to consider a topic covered, because they might be isolated events sprinkled throughout the map.

**Node-Membership:**
Economic Impacts: 0.05
Scientific Information: 0.95

**Edge-Membership:**
Economic Impacts: $\sqrt{0.05 * 0.95} = 0.22$
Scientific Information: $\sqrt{0.95 * 0.05} = 0.22$

01/26/2020 - China says coronavirus can spread before symptoms show -- calling into question US containment strategy

**Node-Membership:**
Economic Impacts: 0.95
Scientific Information: 0.05

**Edge-Membership:**
Economic Impacts: $\sqrt{0.95 * 0.90} = 0.92$
Scientific Information: $\sqrt{0.05 * 0.10} = 0.07$

01/27/2020 - Oil prices have fallen 10% as coronavirus raises fears for global growth

**Node-Membership:**
Economic Impacts: 0.90
Scientific Information: 0.10

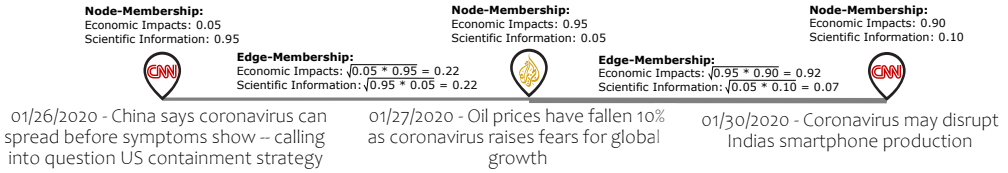01/30/2020 - Coronavirus may disrupt Indias smartphone production

Fig. 3. Overview of the edge membership computation used in the coverage constraint. The example shows three events from the Coronavirus narrative and two clusters. The first edge does not contribute much to the coverage of either cluster, because it is joining nodes from different clusters. In contrast, the second edge joins nodes from the same cluster, and thus it strongly contributes to the economic impacts cluster.

Instead, we rely on edges to determine if a topic is properly covered. Consider for example the storyline in Figure 3 comprised of three events. This storyline focuses mostly on the economic impact of the virus. However, the first event *[China says Coronavirus can spread before symptoms show (...)]* is mostly unrelated to economic impacts, instead, it talks about scientific information. Since this event is isolated in a storyline about a completely different topic (economic impacts) and does not connect with other events talking about its own topic (scientific information), it would not meaningfully contribute to the scientific information topic coverage.

Computationally, we define the coverage of the topic cluster $k$ as the sum of all edge weights that belong to that cluster weighted by edge membership to cluster $k$ (see for example Figure 3). The membership of an edge to a cluster is defined as the geometric mean of the membership of each endpoint. We use the geometric average because its properties work intuitively well with the concept of membership. Similar to what we did with the coherence computation, we want a scoring mechanism that has strong penalties if any of the endpoints does not belong to the cluster.

For example, if both endpoints have a non-zero probability of being in that cluster, their edge gets a positive membership. Otherwise, if any of them has a zero probability, the edge membership is also zero. Consider the storyline in Figure 3 based on the Coronavirus narrative, with only two clusters we can clearly see the impact of joining two nodes from distinct clusters to the values of edge membership. Furthermore, while the first edge does not contribute much [0.22] to the coverage of either of the clusters, the second edge strongly contributes [0.92] to the economic impacts cluster. We normalize the outgoing edge weights for each node to ensure that they sum up to 1. This turns our coherence weights into probabilities, which helps to interpret them.

*Phase 4: Extracting the Main Storyline.* Thus far, our algorithmic approach has constructed the basic narrative map structure. Now we use this structure to extract the main route or storyline (i.e., the constituent events). To do this, we find the most coherent storyline from the starting event to the ending event in the narrative map. In other words, we find the path with the highest probability from the source to the sink, or equivalently, the maximum likelihood path. In general, we can find the probability of each path by multiplying the edge weights. However, instead of maximizing the multiplication of probabilities, in practice, it is more convenient to minimize the sum of negative log-probabilities [17]. By turning this into an additive minimization problem, we can use shortest path algorithms on our graph with log-probability weights to find the maximum likelihood path.

In practice, there is usually only one such route, but it is theoretically possible to have more. If there are multiple routes, this would imply that the algorithm has produced two potential sets of constituent events to describe the narrative and is incapable of choosing the best one.
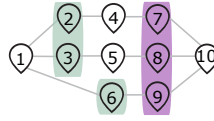
Fig. 4. Antichain examples. The nodes highlighted in green form the first maximum antichain of the graph with 3 elements. The nodes highlighted in purple also form a maximum antichain, but it is not the earliest antichain. The nodes 2 and 3 on their own also form an antichain, but it is not a maximum antichain. Note how each element of the antichain is associated with a chain (left, center, and right)

*Phase 5: Finding Representative Landmarks with Antichains.* Now that we have found the main route, we next find the representative landmarks in the narrative map. Before delving into the method, we introduce some important concepts from graph theory that informed our technique.

To extract the representative landmarks, we leverage the concept of maximum antichains in directed acyclic graphs. Formally, antichains are sets of pairwise incomparable nodes in the DAG (i.e., nodes that are not connected by a path) [47]. For example nodes 2 and 3 form an antichain in Figure 4. In particular, an antichain is called *maximum* antichain if no other antichain has more nodes. For example, nodes 2, 3, and 6 form one possible maximum antichain, no other antichain that we select can have more nodes.

Previous works have shown that maximum antichains formalize the concepts of abstract graph views, slices [18], or abridgments (i.e., a small part of the graph that is of interest) [14]. Furthermore, researchers have used them to define a partitioning of the nodes of the graph [2]. In particular, according to Dilworth's theorem [13], the size of the maximum antichain—called the width of the DAG—is equal to the minimum number of chains (paths) into which we can partition the DAG. Intuitively, we can partition the graph into chains (i.e., storylines) and interpret each node of the maximum antichain as a representative element for one of the chains. This intuitive interpretation is backed by Dilworth's formal theorem. Figure 4 demonstrates this with an example. Take the chains of Figure 4, $A : 1 \rightarrow 2 \rightarrow 4 \rightarrow 7 \rightarrow 10$, $B : 1 \rightarrow 3 \rightarrow 5 \rightarrow 8 \rightarrow 9$, and $C : 1 \rightarrow 6 \rightarrow 9 \rightarrow 10$, we can associate one element from an antichain to each of these chains, such as $(A, 2)$, $(B, 3)$, $(C, 6)$.

To find the maximum antichains, we use the antichain extraction method implemented in the NetworkX library [24] in Python. This gives us a list with all the antichains (i.e., list of node indices). We filter the list, ensuring that it only contains maximum antichains. Note that there are potentially multiple maximum length antichains. In this case, our tiebreaker strategy is to select the earliest antichain according to their timestamps, as this provides us with representative landmarks close to the starting points of each storyline. For example, consider the two maximum antichains in Figure 4. In that case, we would choose the green antichain for our representative landmarks.

*3.2.4 Phase 6: Generating the Final Visualization.* Finally, we turn our computational representation into a visual representation, which is used for the case study and user evaluation. In particular, we use GraphViz [15] and the DOT language [19] to generate the layout of our visualization in an SVG file. Then, we embed the SVG file into an HTML page and add basic interactive elements using CSS.

Our visual representation follows the route map metaphor that we described in the introduction. The metaphor elements in Table 1 guide our design and we try to faithfully follow our original concept in the illustrations of Figure 1. The visualization shows the date and the headline accompanied by a landmark icon with the news source logo. We note that our visualization omits the edge weight information. We removed it because preliminary evaluations proved that adding explicit weights confused the users. Instead, we make edges wider to represent higher weights.
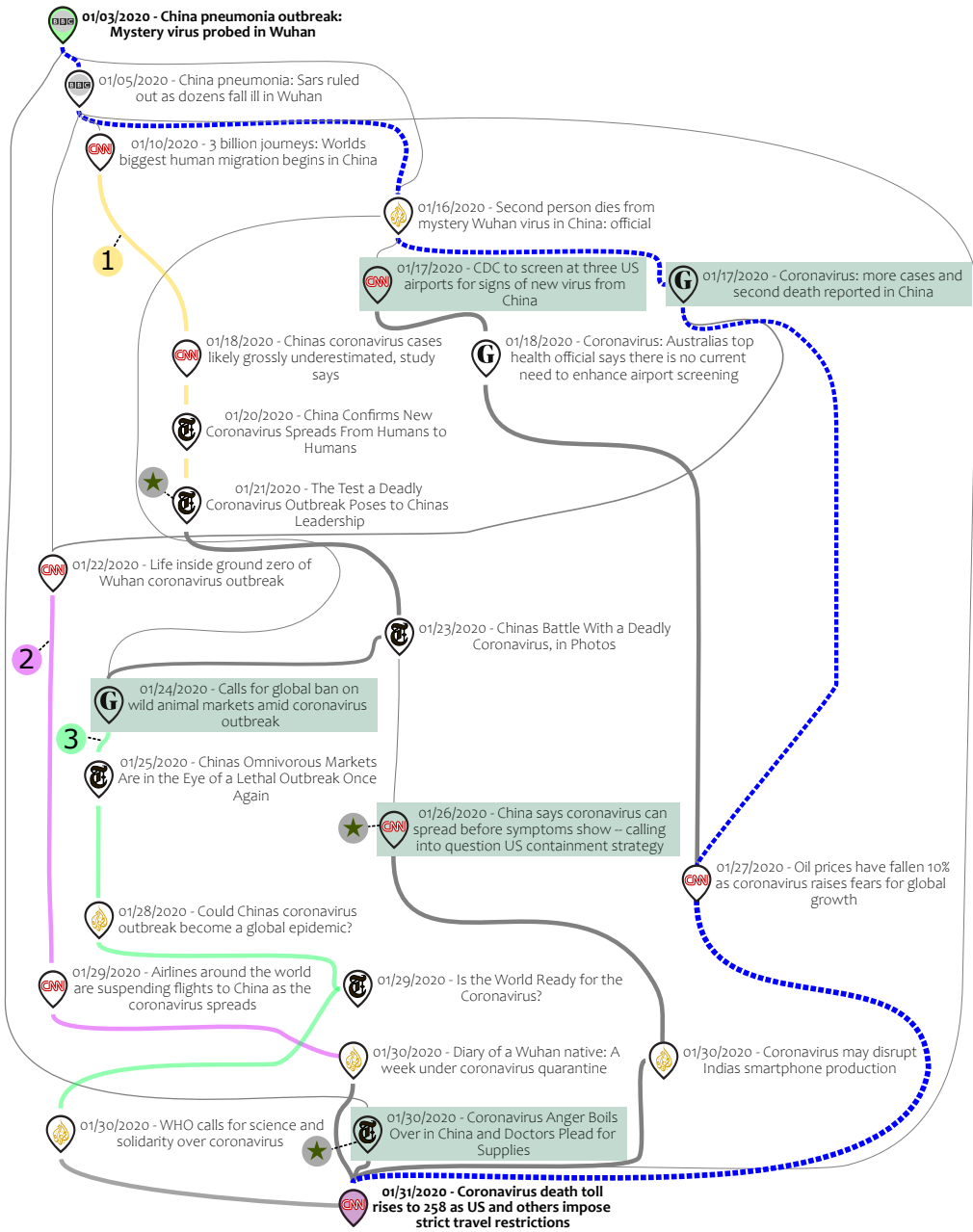
Fig. 5. Narrative Map for the Coronavirus Outbreak during January. Headlines in bold denote the starting event 🟢 and ending event 🟣. The blue dashed line corresponds to the main storyline ✏️ (the maximum likelihood path). The width of each edge depends on its weight (coherence). Events highlighted in green are representative landmarks ⬭ (members of the first maximum antichain). The markers 1, 2, and 3 highlight some storylines that we reference throughout our discussion, we display them in distinctive colors. The ⭐ markers show events that share the theme of questioning or criticizing government responses.

(a) Main Route *[Constituent Events]*



**01/03/2020 - China pneumonia outbreak: Mystery virus probed in Wuhan** | 01/05/2020 - China pneumonia: Sars ruled out as dozens fall in Wuhan | 01/16/2020 - Second person dies from myster Wuhan virus in China: official | 01/17/2020 - Coronavirus: more cases and second death reported in China | 01/27/2020 - Oil prices have fallen 10% as coronavirus raises fears for global growth | **01/31/2020 - Coronavirus death toll rises to 258 as US and others impose strict travel restrictions**

**Major Themes:** Spread of the virus and economic Impacts

(b) Representative Landmarks *[Supplementary Events]*



01/17/2020 - CDC to screen at three US airports for signs of new virus from China | 01/17/2020 - Coronavirus: more cases and second death reported in China | 01/24/2020 - Calls for global ban on wild animal markets amid coronavirus outbreak | 01/26/2020 - China says coronavirus can spread before symptoms show – calling into question US containment strategy | 01/30/2020 - Coronavirus Anger Boils Over in China and Doctors Plead for Supplies

**Major Themes:** Government containment measures to contain the outbreak, spread of the virus, economic impacts, scientific information about the virus, and questioning or criticizing official measures.
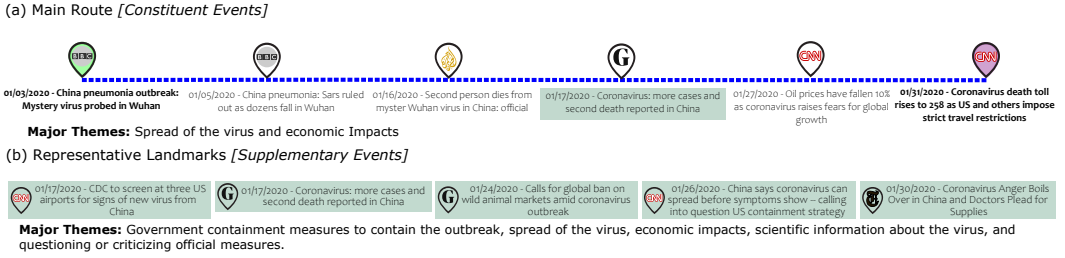
Fig. 6. Analysis summary of the main route and representative landmarks for the January narrative map. We extract the major themes out of these elements.

## 4 CASE STUDY: CORONAVIRUS NARRATIVE

In this section, we showcase our narrative extraction algorithm in action. We present a case study to demonstrate how we can leverage the narrative route map representation and our computational extraction method to understand the narrative as it evolves and changes. First, we present the data used to generate the narrative map. Then, we show our results and analysis.

### 4.1 Overview of the Case Study

We focus on the Coronavirus outbreak due to its high impact and how clearly defined the major events are (e.g., first case, lockdown in Wuhan, the closing of borders, stay at home orders). Note that it is not a requirement to have such clearly defined major events to create a narrative map, but it facilitates the analysis and thus it makes for an informative example.

We obtained data related to the Coronavirus outbreak from the COVID-19 archive[1]. This archive maintains a list of news articles related to the outbreak and comprised 1576 articles from 373 unique sources spanning three months—Jan'20 to Mar'20. For the purposes of this case study, we focused on the first month of the outbreak and retained articles from December 31st to January 31st, resulting in a data set of 607 news articles.

To reduce the search space of the optimization algorithm, we selected articles from the top 5 news domains in terms of publications (BBC, Al Jazeera, The Guardian, CNN, and New York Times), resulting in 102 news articles from the first month of the outbreak. Focusing on the major sources allows us to obtain an overview of the most prominent storylines and the major themes in the narrative while ignoring less important stories from more obscure sources. For the starting and ending events, we simply took the chronological first and last elements in the corpus. Out of the 102 articles, our narrative extraction algorithm returned an optimal solution with 22 events, in addition to the fixed starting and ending events. The rest were not selected because they lowered the overall coherence when added to any of the extracted storylines. Figure 5 shows our results.

### 4.2 Analysis: Narratives underlying the Coronavirus Outbreak

What are the narratives present in the first month of the Coronavirus outbreak? Figure 5 shows the full narrative map generated by our algorithm. Figure 6 shows a summary of the major themes of the narrative according to the main route ⟡ and the representative landmarks ⬡.

We begin by analyzing the **main route** in Figure 6(a), which starts by referring to the new "mysterious pneumonia virus" in Wuhan. At this point in the outbreak, information about the virus is scarce, but at the very least SARS was ruled out. Then, less than a week later, there are already two deaths from this mystery virus. The next day news outlets report the second death and the

---

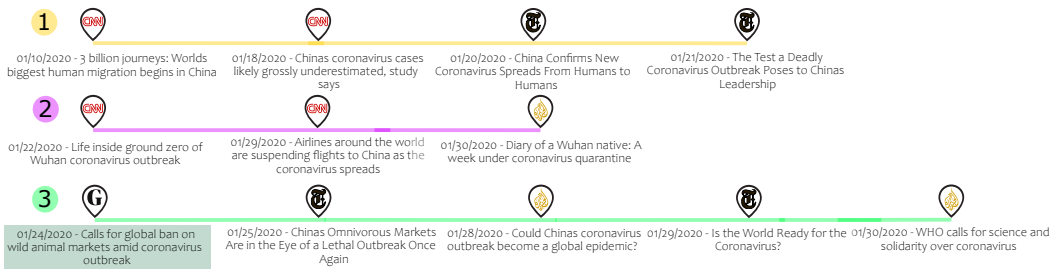[1]https://www.covid19-archive.com/

Fig. 7. Examples of substories from the Narrative Map for the Coronavirus Outbreak during January. For brevity the start event ⬡ and ending event ⬡ are not shown in detail. Events highlighted in green are representative landmarks ⬡. We use markers ①, ②, and ③ to refer to the storylines as in Figure 5.

virus has now been called "Coronavirus". After this point, the story changes its focus and talks about the drop in oil prices as the Coronavirus raises fears for global growth. Finally, this leads to the final event with a death toll of 258 and the US and other countries imposing restrictions. Thus, the major themes of the main route are the general spread of the virus and its economic impacts.

The **representative landmarks** in Figure 6(b) reveal a thematic overlap with the main route as well as surface new elements, where common topics overlapping with the main route include the "spread of the virus" and its "economic impact." New themes in the representative landmarks refer to "scientific information about the virus" *[China says Coronavirus can spread before symptoms show (...)]*, "government measures to contain the outbreak" *[CDC to screen at three US airports (...)]*, and "criticisms and questioning of government responses" *[Coronavirus anger boils over in China (...)]*.

Next, we turn our attention to some of the **stories** on the map. For instance, see ① in Figures 5 and 7. Here we see references to mass migration as people try to escape from the virus. Then we find out that the number of cases might be grossly underestimated and that humans can transmit the virus to other humans. Thus, this storyline implies that the virus is likely to spread from human mass migration.

Another storyline deals with the lockdown in Wuhan and the surrounding provinces (see ② in Figures 5 and 7). At this point, we see multiple articles about how life goes on during the quarantine. This storyline focuses on a more human aspect of the outbreak, seeing how it affects the daily lives of people. This is a theme that we did not identify directly by looking at the main route and the representative landmarks. This implies that some themes are abstracted away when we look at the narrative exclusively from the main route and representative landmarks perspective.

Amidst the battle with the virus, we see a route dealing with wild animal markets (see ③ in Figures 5 and 7), one of the potential causes of the outbreak according to the information available at that time. Then, this turns into a discussion about whether the virus could become a global pandemic and whether the world is ready to handle it. This is followed by the World Health Organization (WHO) asking for science and solidarity over the virus.

There are also events dealing with anger over how the Chinese government is handling the situation, talks about the challenges the outbreak could pose to the Chinese government, and questioning the United States containment strategy. These events are part of different storylines (see the ★ markers in Figure 5), yet they share a common theme in terms of questioning governments and how they are handling the outbreak.

| Dimension | Element | Questionnaire Item | Source |
|---|---|---|---|
| **Representation** | Coherence | The map presents a coherent overview of the narrative. | Shahaf (2010) |
| | Relevance | The map presents relevant information about the narrative. | Shahaf (2010) |
| | Redundancy | The map has too much redundant information. | Shahaf (2010) |
| | Effectiveness | I feel more familiar with the narrative after using this map. | Shahaf (2010) |
| | Completeness | The amount of information on the map is appropriate to represent the narrative. | Burkhard (2005) |
| | Size | The size of the map is appropriate to represent the narrative. | Burkhard (2005) |
| **Metaphor** | Nature | The metaphor is natural (i.e., it is similar to the real-world). | García (2015) |
| | Understanding | The metaphor is understandable (i.e., easy to understand and extract information). | García (2015) |
| | Interpretability | It is easy to draw conclusions and get new insights from the metaphor. | García (2015) |
| | Landmarks | Presenting narrative events as landmarks makes sense. | Burkhard (2005) |
| | Main Route | The main storyline serves well as an overview of the most important events in the narrative. | New |
| | Representative Landmarks | The representative landmarks serve as an overview of all the stories in the narrative. | New |
| **Visualization** | Visual Impact | The visualization is eye-pleasing. | Shamim (2015) |
| | Overall Performance | The visualization is easy to understand. | Shamim (2015) |
| | | The visualization is user-friendly. | Shamim (2015) |
| | Overall Design Style | The visualization is informative. | Shamim (2015) |
| | | The visualization is intuitive. | Shamim (2015) |
| | Information Quality | The visualization is useful. | Shamim (2015) |
| | | The comprehensiveness of the data is good. | Shamim (2015) |
| | Visual Representation Model | The visualization allows us to easily compare storylines. | Shamim (2015) |
| | | The representation style of the data is good. | Shamim (2015) |
| | Information Presentation Model | Previous knowledge is required to understand the visualization. | Shamim (2015) |

Table 2. Table containing the evaluation dimensions and their elements, as well as the questionnaire items and their source (if applicable). Some questions taken from a source have been adapted to our metaphor. The questionnaire has a total of 22 elements. Subjects were asked to evaluate each statement according to a 5-point Likert scale, ranging from "Strongly Disagree" to "Strongly Agree."

## 5   USER EVALUATION

Finally, to provide additional support for the effectiveness of our narrative map representation and answer RQ3, we perform a user evaluation. Here, we present the guidelines driving our evaluation followed by the results of our user study.

### 5.1   Evaluation Guidelines

The narrative map representation should be an effective method to communicate complex narratives to individuals. Hence, to measure the effectiveness of the representation, the underlying metaphor, and the overall algorithmically generated narrative-map visual, we evaluate three dimensions:

(1) **Representation**: How well does the map represent the narrative?
(2) **Metaphor**: How well does the route map metaphor work for narratives?
(3) **Visualization**: How well does the visualization work in general?

For each of these dimensions, we adapt evaluation questionnaires from a range of prior scholarly work. Table 2 shows an overview of the questionnaire. For example, to evaluate along the representation dimension, we adopt Shahaf et al's [53] evaluation scheme from their Connect-the-Dots

algorithm—an algorithm that heavily informed our extraction approach. We also take elements from Burkhard and Meier's work [8], where they present a framework for evaluating knowledge visualization tools that are based on visual metaphors. To evaluate the effectiveness of our metaphor, we adopted elements from García et al.'s [20] work, where they focused on visualization metaphors to process complex information and interpret data. Finally, we also take elements from the visualization evaluation questionnaire of Shamim et al. [59]—a survey questionnaire designed to assess the usability of visualization systems pertaining to news articles. To measure the internal consistency of our final 22-item evaluation questionnaire, we carried out a reliability analysis. Cronbach's alpha with $\alpha = 0.9560$ demonstrates the high internal consistency of our evaluation questionnaire.

## 5.2 Evaluation Procedure

Can users with no training in narrative or news analysis use and interpret our algorithmically generated narrative map? To answer, we asked Amazon Mechanical Turk workers (MTurkers) to do short tasks. Recent work by Cheng et al. [10] has shown the promise in employing MTurkers in the context of explaining complex AI algorithms, where they leverage the results from an MTurk evaluation with non-expert users to guide the design of their systems. Moreover, prior scholarly works have used MTurk in intelligence analysis tasks and have obtained high-quality results, demonstrating the capabilities of MTurk [32]. For example, MTurkers have performed at par with intelligence analysts in connecting entities in a visualization map [33] and in identifying relationships in large sets of data[43]. Therefore, even if MTurkers are not a replacement for expert feedback, they could still act as good proxies and provide valuable insights towards how our narrative maps proposal representation, metaphor, and visualization work in the context of sensemaking and intelligence analysis tasks. Appendix A contains snapshots detailing the MTurk job.

Prior research [26, 46] indicates that it is possible to obtain good quality data from MTurk. Specifically, as long as we follow certain guidelines, the data quality is similar to traditional sources (e.g., student recruitment pools). Furthermore, research using MTurk faces many ethical challenges [22, 66], in particular with regards to payment. These challenges have given rise to various initiatives to tackle them, such as Dynamo [52]. We have followed the general Dynamo guidelines for academic requesters in MTurk. We have ensured that our MTurk workers receive appropriate compensation, above the federal minimum wage according to the estimated response time. Furthermore, we provided workers with an explanation in case they were rejected due to failing attention check questions.

The task required that MTurkers spent time to interact with the visualization, understand the structured data it surfaced, infer one of the major themes of the main route and two themes of the whole narrative map (different from the main route one), followed by responding to our 22-item evaluation questionnaire (shown in Table 2). Subjects were required to answer how much they agreed with the statements shown in Table 2 on a 5-point Likert scale with the following options: Strongly Disagree, Disagree, Undecided, Agree, and Strongly Agree. In addition to the 22 questions, we added one attention check question. Three workers responded incorrectly to the attention check question. We discarded their answers and requested additional workers to keep our total to 20 valid responses for each narrative map task evaluation. The entire duration of the task along with reading instructions was approximately 30 minutes. Workers were compensated for $4 for each task, adhering to federal minimum wage standards.

To warrant high-quality responses, we required MTurkers to be Master Turkers, have at least a 95% approval rating, and to have completed at least 50 tasks on MTurk. To ensure that the workers correctly understood the task, we first trained them by providing an explanation of the basic elements of the narrative route map metaphor. We also showed them a simple example narrative
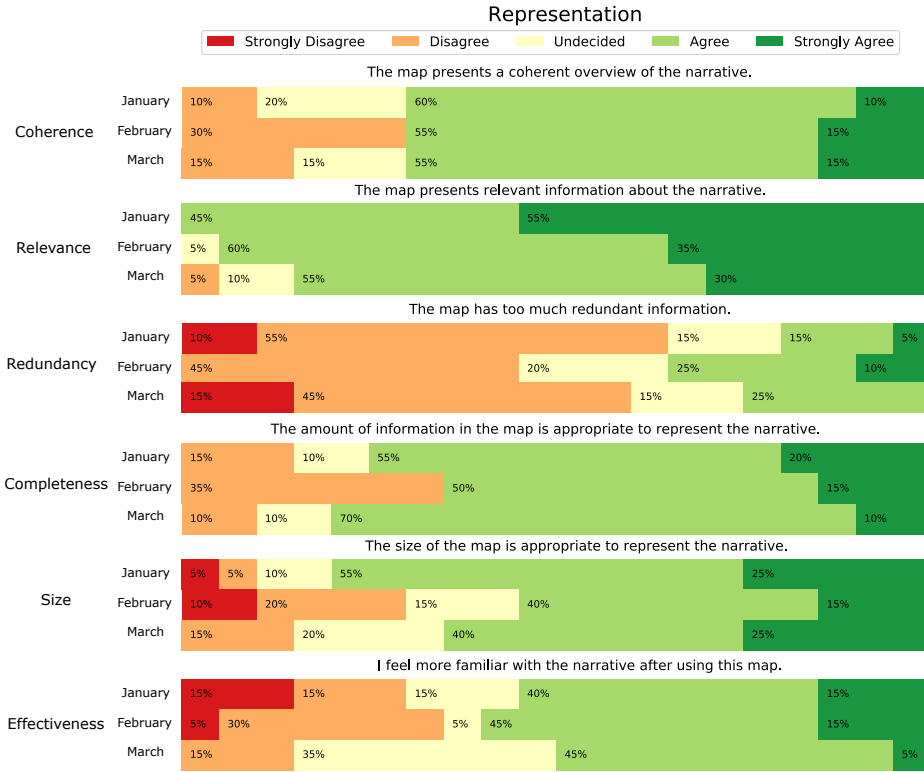
Fig. 8. User evaluation results for the representation dimension. The bars show the percentage of users per response for each question.

map with just three nodes, including a short discussion of the common themes that could be extracted from the example map. Prior work in crowdsourcing suggests that training workers by showing them examples results in high-quality responses [31, 41].

We performed our MTurk evaluation with three different narrative maps, each requiring 20 responses. All the maps were focused on the Coronavirus outbreak but spread across three different months to facilitate the evaluation: January, February, and March 2020.

## 5.3 Evaluation Results

Here, we present the results of our user study across each dimension (representation, metaphor, and visualization). Overall, all dimensions were evaluated well by the MTurkers as shown in Figures 8, 9, and 10. These results are mostly consistent throughout the three months. For brevity, we focus on a few interesting results from our analysis.

*5.3.1 Representation.* First, we analyze the results of the representation dimension. Overall, the maps are considered as a *coherent* overview of the narrative for each month. Next, we note that there is a slight drop in *relevance* during March, which might be because the March map has 40 events and is significantly larger than the other maps. Thus, there might be too many irrelevant events compared to the other maps. In contrast, we see that most users did not consider our representation *redundant*. This result is encouraging considering that one of the potential issues of using a similarity-based approach was that our representation could be redundant since it
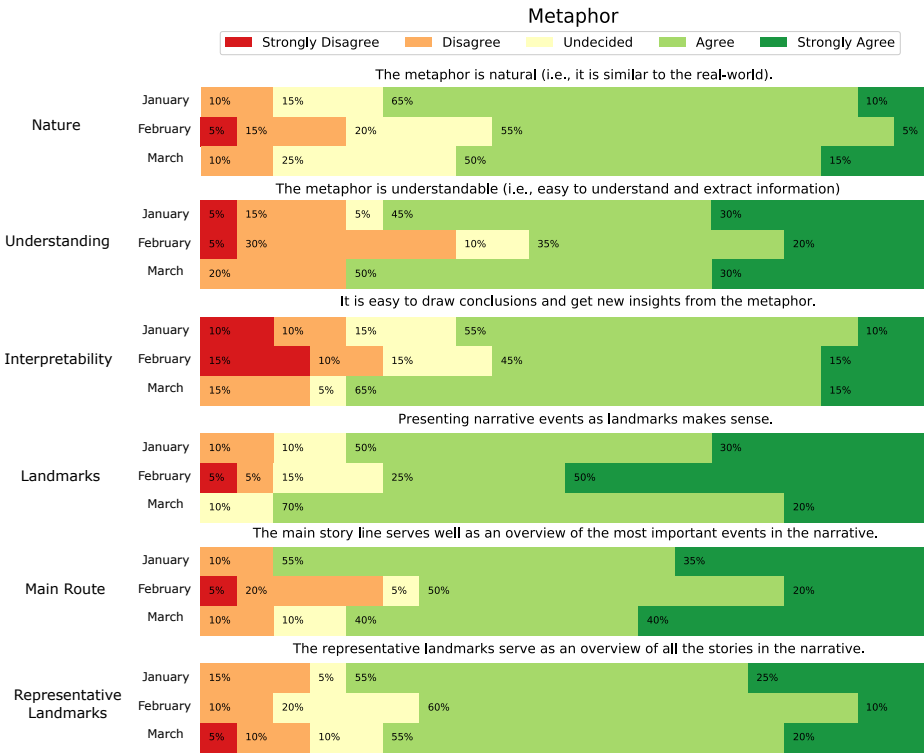
Fig. 9. User evaluation results for the metaphor dimension. The bars show the percentage of users per response for each question.

would naturally keep trying to connect highly similar items. In terms of *size*, we see that March had better results than February, despite being the larger map. This difference in results could be due to layout differences rather than the number of nodes itself. Moreover, we see that there is a trade-off between an appropriate map size and *completeness* of the data. In this context, and considering our results, our narrative extraction method balances both of these requirements adequately. Finally, we see that at least half of the users felt more familiar with the topic after using the visualization, thus showing the *effectiveness* of our representation. While we expected a more positive response to this question, it should be noted that due to how pervasive Coronavirus news was at the time we did our user study, it is possible that Turkers were already very well acquainted with the information at hand. With a more obscure topic, the visualization could have had a greater impact on the users by providing new knowledge.

*5.3.2  Metaphor.* Second, we analyze the results of the metaphor dimension. In general, the results for this dimension were positive. Users agree that our route map metaphor as well as the more specific metaphors (*landmarks*, the *main route*, and *representative landmarks*) are indeed an intuitive and *natural* way to represent a narrative. Furthermore, the users agree that it is easy to draw conclusions from the metaphor, thus it is highly *interpretable*. Moreover, most users found the metaphor to be *understandable* for January and March. However, fewer users agreed to this for the February map. We believe that this is due to the complex layout of the February map compared to the other two.
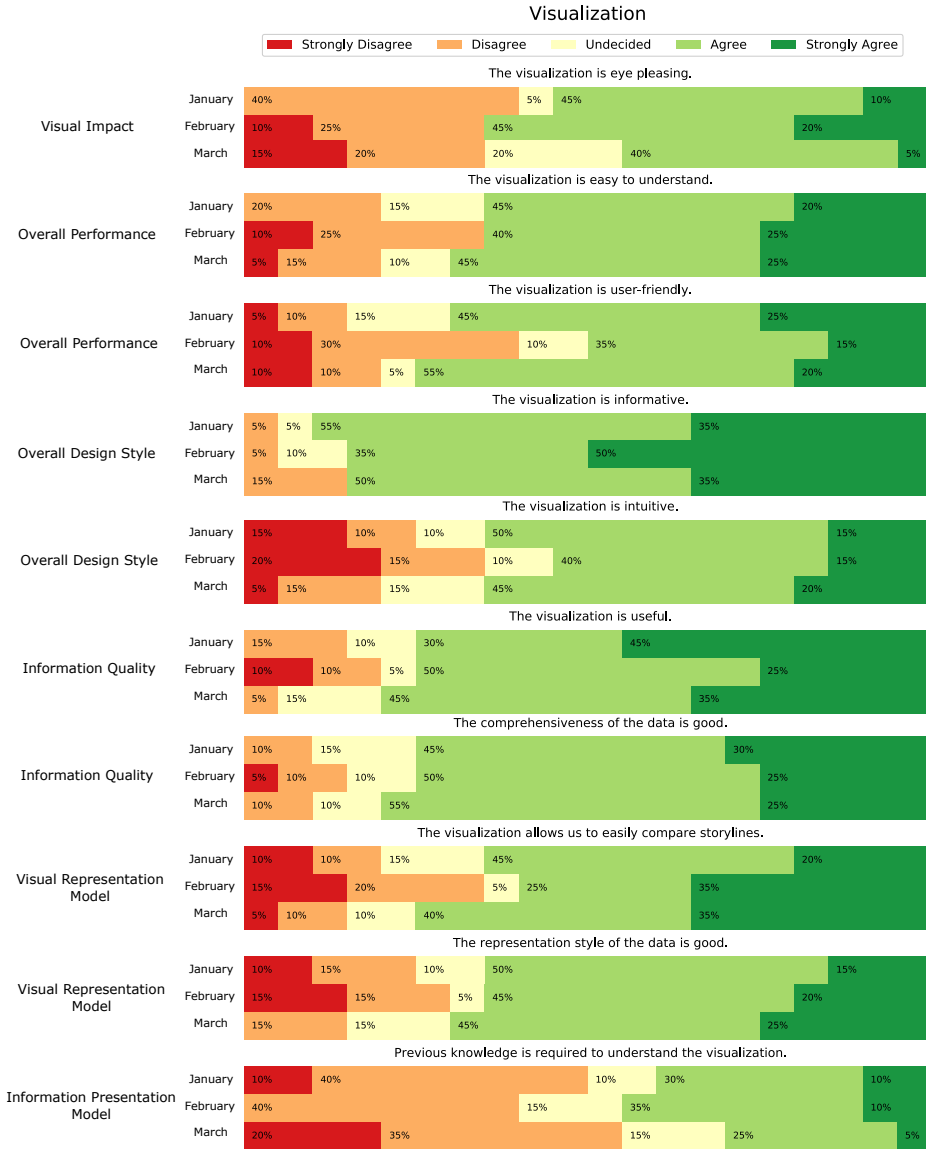
Fig. 10. User evaluation results for the visualization dimension. The bars show the percentage of users per response for each question.

*5.3.3 Visualization.* Finally, we turn to our analysis of the visualization dimension. Overall, the evaluation results for this dimension were positive as well. In terms of *visual impact*, most users found the visualization to be eye-pleasing, except for March. This difference could be due to the size of the map. In terms of its *overall performance*, the visualization had good results in both ease of understanding and user-friendliness. Regarding the *overall design style*, users mostly agreed that the current design is informative and intuitive. Likewise, in terms of *information quality*, the visualization had good results regarding its usefulness and the comprehensiveness of the data. The *visual representation model* was also evaluated well by the users in the context of comparing

storylines. Finally, with respect to the *information presentation model*, around half of the Turkers considered that the narrative representation did not require previous knowledge. The rest of the users believed that previous knowledge would be required to understand the visualization. This split in the user evaluation indicates that additional measures might be needed to facilitate using the visualization without prior knowledge. However, it might also reflect the fact that most people likely had previous knowledge about Coronavirus due to the constant barrage of news surrounding this topic. We leave additional evaluations, with a more obscure topic, for future work.

## 6 DISCUSSION

Our work provides a computational representation of narratives grounded in formal narrative theory. We also offer an algorithmic approach to extract these representations from data using a linear programming technique. We summarize our approach and outline the key implications.

### 6.1 RQ1: Narrative Representation

Unlike previous attempts at extracting storylines or narratives from data [53, 58, 64, 70, 74], our work is based on a comprehensive survey of narrative theory. Hence our narrative representation provides a general framework to represent narratives and can serve as a starting point for understanding narratives through a computational lens.

Our representational framework emulates the real-world metaphor of a route map, making it more usable and understandable while displaying abstract unstructured knowledge (also demonstrated by our user evaluation results). Moreover, the computational representation of our framework—graph-based—can be effortlessly extended to encompass other narrative phenomena. For example, we can augment our data set to contain social media information such as comments, tweets, or posts related to our news events. By feeding this additional data to our narrative building model, we can incorporate the reactions of people towards certain articles. In turn, this provides us with insights into the issue of narrative reception [51] (i.e., how the public perceives narratives) and how it relates to the projected narrative by the media.

Additionally, we could also add new discourse elements in the form of contextual information through text summaries and data plots. For example, in the case of the Coronavirus narrative, our narrative map could be augmented by including a text summary of the outbreak during that month and traditional data plots to show the growth of the virus as the narrative advances (see Figure 11).

### 6.2 RQ2: Algorithmically Extracting Narratives

The optimization-driven approach for generating narrative maps uses an objective function that seeks to maximize the coherence of the narrative based on similarity and clustering information. The resulting solution of this optimization problem selects real-world events that are displayed in the narrative map along with their connections. Our user evaluation study shows that our approach constructs a primarily coherent and relevant narrative from hundreds of news articles.

Moreover, we include a coverage constraint to force the narrative map to cover many topics. This constraint could be modified and expanded to consider different kinds of coverages, similar to the idea of user-personalized coverage from previous scholarly work [56]. For example, by adding a term to consider the political bias of the news sources, we could require a certain level of coverage for each side of the political spectrum. Thus, ensuring that all sides of the story are covered.

To help users interpret the storylines, major themes, and the connection between events, we provide them with the main route and representative landmarks of our map. These elements offer a way to characterize narratives by helping identify the storylines and major themes of the narrative. Our user study confirmed that both the main route and the representative landmarks served well in providing an overview and highlighting key elements of the narrative. While the *main route*

supplies with a *longitudinal view* of the narrative, *representative landmarks* created by the maximum antichain offer a *transverse view*. In essence, the representative landmarks and the main route provide us with an understanding of both the breadth and the depth of the narrative, respectively.

## 6.3 RQ3: Evaluation

The evaluation proves that the Narrative Map representation, alongside the extraction method, is an effective mechanism to help users understand complex news narratives. The route map metaphor was considered natural by the users, providing a coherent overview of the narrative. Moreover, the main route and the representative landmarks are able to convey the key elements of the narrative.

Although the user evaluation demonstrated the effectiveness of our method, there is room for improvement. We focus in particular on the possibilities for enhancing the visualization method. As noted in the user evaluation, it is necessary to refine its visual appeal. Moreover, reducing the need for previous knowledge in the visualization is an important consideration. This could be achieved by including auxiliary discourse elements to provide additional context (e.g., news summaries and complementary data visualization).

Further improvements could be made by adding more interactivity to the visualization. For example, a simple tweak would be to highlight all edges of a node when an edge is selected. More complex upgrades could allow users to move the nodes along with its edges or allow them to hide or merge certain nodes while preserving structure. For instance, an approach similar to the recursive summarization technique for discussion forums [72] could be adopted for narratives, enabling a user to explore different narrative storylines based on their interests, with varying levels of details. Investigating these extensions could be fruitful avenues for future research. Nevertheless, in its current state, the visualization correctly represents the narrative and is able to provide users with new insights, as shown by our user evaluation study with MTurkers.

## 6.4 Implementation Parameters and Constraints

Our linear programming problem has a total number of variables in the order of $O(|D|^2 \cdot k)$, where $|D|$ is the number of documents and $k$ is the number of clusters. In practical terms, the number of clusters is small enough to be considered constant. Thus, our main computational bottleneck lies in generating the pairwise combinations of nodes required by the optimization problem. Future work could deal with improvements to the linear programming approach or even applying a completely different optimization technique, similar to how the Connect-the-Dots method [53] evolved from its basic linear programming formulation into a divide-and-conquer approximation approach [57]

We also note that there are two parameters of interest in our linear program proposal: the expected length of the main route ($K$, based on the Connect-the-Dots formulation [57]), and the minimum average coverage of clusters (mincover). Tuning these parameters is required to obtain a good narrative map representation. Moreover, since our method depends on document embeddings and clustering, the choice of the vector embedding space representation, the clustering algorithm, and all the corresponding parameters for clustering can also influence the final representation. Thus, special care is required to choose an appropriate vector representation for the events, as this can influence both the similarity and the clustering results.

## 6.5 Limitations

One limitation of our study is that we constructed the event representation based on the headlines of the articles. Although headlines generally contain key event information, it would be useful to consider either the full text of the article or an appropriate summary. In particular, considering the headline and the lead of the article might be enough to capture the main event descriptors [45].

In terms of the user evaluation of our narrative maps, we only focused on the Coronavirus narrative with only 20 users for each map. In order to properly generalize our results, future work should focus on evaluating different narrative maps with more users and with different topics, ranging from well-known issues (e.g., Coronavirus) to more obscure ones. Nevertheless, our user evaluation provides an important baseline for future works involving this narrative representation.

### 6.6 Implications and Future Work

*6.6.1 A General Framework to Extract and Represent Narratives.* Our method is grounded in the scholarly definitions of narrative theory. We start with abstract theoretical definitions and move toward concrete representations of the essential elements in a narrative. Hence, our representation should be able to capture any kind of narrative as defined in the narrative theory literature. As long as data is in the form of (event, timestamp) tuples, we can use our algorithm to construct a map representation. Moreover, we can generalize this approach to any kind of data provided it has a temporal ordering and that we can define a measure (i.e., the evaluation metric of our metaphor) that allows us to compute coherence through similarity and clustering. At a technical level, the generalizability of our approach stems from the abstract mathematical formulation used to extract the map, which could, in turn, be used for other types of data, such as image data (e.g., creating a narrative map out of pictures from an embedding representation and timestamps).

*6.6.2 Design Implications.* Grounded in theoretical constructs of narratology, our approach can automatically extract structured information from multiple data points and represent it as a narrative map—a visual metaphor similar to physical maps. We believe that this capability offered by our approach can be used in the design of a wide range of systems. For example, imagine a news reading tool that accompanies articles with a narrative map visualization to help readers understand the big picture of the narrative, or consider a social media monitoring system that shows how the online narrative surrounding a certain issue evolves over time. Indeed, literary theorists, while outlining the importance of narratives noted [1]: "*we do not have any mental record of who we are until narrative is present as a kind of armature, giving shape to that record.*"

Furthermore, narrative maps could be used in the context of computational journalism, which focuses on providing tools for various computational needs in the journalistic context [12], such as topic detection, visualization, and sensemaking. Moreover, the theoretical basis of our representation, along with the general approach of extracting narratives from data, ensures that a designer can easily adapt our proposal to a different domain. For example, our model could be adopted by psychologists to help analyze personal narratives [30] or by narratologists to study literary works.

*6.6.3 Extending the Narrative Map Representation.* Using the foundations and methods presented in this paper, future work can create an enhanced representation, for example, by considering the bias and credibility of the sources of each event. See for instance how we added this information to the landmarks of the narrative maps in Figure 11. With these enhanced maps, intelligence analysts would have an additional tool at their disposal to help detect potential online misinformation campaigns by analyzing emergent narratives. The rapidly shifting media ecology has made strategic narratives more difficult to control [38], thus analysts would benefit from being able to detect counter-narratives as they emerge. Additionally, fact-checkers could use this enhanced representation to obtain an overview of how facts, both credible and questionable, spread throughout the narrative. This could help in their fact assessment process. Moreover, they could augment our representation by manually modifying it to add credibility annotations.

Furthermore, narrative maps can also be used as a summarization tool to identify high-level patterns in narratives. For example, for the Coronavirus narrative shown in our case study, we identified "economic impacts" and the "spread of the virus" as major themes. In conjunction with
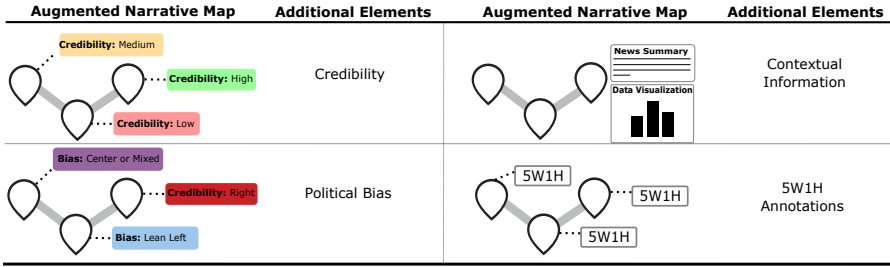
Fig. 11. Potential extensions to narrative maps including different elements. The top-left map considers adding credibility annotations to each of the events, representing how credible or questionable the sources and their claims are. The bottom-left map considers adding political bias annotations, thus helping users understand how bias affects the storylines of the map. The top-right map considers additional information in the form of a news summary and data visualization. The bottom-right map considers adding the answers to the 5W1H questions (who, what, where, when, why, and how), which are the main descriptors of an event.

more advanced graph-theoretical elements, it would also be possible to identify more complex patterns in the narrative. For instance, we could use network motifs—frequent patterns in the structure of a graph [37]—to encode interesting properties and relationships of its elements [6].

## 7  CONCLUSIONS

This paper presents an algorithmic approach to represent and extract narratives using *narrative maps*. Based on a comprehensive survey of formal narrative theory, we outline a computational representation of narratives comprising eight elements. We used an intuitive route map metaphor to connect the elements, their metaphoric renditions, and their computational counterparts. Building upon existing approaches of story extraction, we designed a novel optimization-based approach to extract the underlying graph representing a narrative map. In particular, we seek to maximize the coherence of the map, subject to structural and coverage constraints. Using this graph, we then extract the main storylines and the representative landmarks of the narrative map. These help us characterize the major themes in a narrative.

We presented a detailed case study following the Coronavirus outbreak during its first month. We explored the major events and identified the major themes surrounding the outbreak. This extended example helped us demonstrate how an analyst could leverage our representation to understand the evolution of a narrative. Next, we presented a user evaluation of our method considering three aspects: the narrative representation itself, the effectiveness of the metaphor, and the usability of the visualization. While there are still some aspects that could be improved, our evaluation results demonstrate that our narrative map proposal fulfills its intended purpose.

This work opens up multiple avenues for future research, such as adapting the current approach to another type of data (e.g., images) or to a different field (e.g., helping psychologists analyze personal narratives) or extending the narrative map representation by adding contextual information. In particular, incorporating measures of source bias and credibility into our representation could pave the way towards modeling how misinformation propagates in news narratives.

## ACKNOWLEDGMENTS

## REFERENCES

[1] H Porter Abbott. 2008. *The Cambridge introduction to narrative.* Cambridge University Press, One Liberty Plaza, New York, NY, USA.

[2] James Abello, Frank Van Ham, and Neeraj Krishnan. 2006. Ask-graphview: A large scale graph visualization system. *IEEE transactions on visualization and computer graphics* 12, 5 (2006), 669–676.

[3] Jeffery Ansah, Lin Liu, Wei Kang, Selasie Kwashie, Jixue Li, and Jiuyong Li. 2019. A Graph is Worth a Thousand Words: Telling Event Stories Using Timeline Summarization Graphs. In *The World Wide Web Conference (WWW '19)*. ACM, New York, NY, USA, 2565–2571.

[4] Chris Baber, Dan Andrews, Tom Duffy, and Richard McMaster. 2011. Sensemaking as narrative: Visualization for collaboration. *VAW2011, University London College* VAW2011 (2011), 7–8.

[5] Roland Barthes and Lionel Duisit. 1975. An Introduction to the Structural Analysis of Narrative. *New Literary History* 6, 2 (1975), 237–272.

[6] Austin R Benson, David F Gleich, and Jure Leskovec. 2016. Higher-order organization of complex networks. *Science* 353, 6295 (2016), 163–166.

[7] Kenneth Burke. 1969. *A grammar of motives.* Vol. 177. Univ of California Press, 155 Grand Ave. Suite 400. Oakland, CA.

[8] Remo Aslak Burkhard and Michael Meier. 2005. Tube Map Visualization: Evaluation of a Novel Knowledge Visualization Application for the Transfer of Knowledge in Long-Term Projects. *J. UCS* 11, 4 (2005), 473–494.

[9] Daniel Cer, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St. John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar, Yun-Hsuan Sung, Brian Strope, and Ray Kurzweil. 2018. Universal Sentence Encoder. *CoRR* abs/1803.11175 (2018), 1–7. arXiv:1803.11175

[10] Hao-Fei Cheng, Ruotong Wang, Zheng Zhang, Fiona O'Connell, Terrance Gray, F. Maxwell Harper, and Haiyi Zhu. 2019. Explaining Decision-Making Algorithms through UI: Strategies to Help Non-Expert Stakeholders. In *Proc. of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, 1–12.

[11] Giovanni Luca Ciampaglia. 2018. Fighting fake news: a role for computational social science in the fight against digital misinformation. *Journal of Computational Social Science* 1, 1 (2018), 147–153.

[12] Sarah Cohen, James T Hamilton, and Fred Turner. 2011. Computational journalism. *Commun. ACM* 54, 10 (2011), 66–71.

[13] Robert P Dilworth. 1987. *A Decomposition Theorem for Partially Ordered Sets.* Birkhäuser Boston, Boston, MA, 139–144.

[14] Peter Eades and Mao Lin Huang. 2004. Navigating Clustered Graphs Using Force-Directed Methods. In *Graph Algorithms and Applications 2*. World Scientific, 5 Toh Tuck Link, Singapore, 596224, Singapore, 191–215.

[15] John Ellson, Emden R. Gansner, Eleftherios Koutsofios, Stephen C. North, and Gordon Woodhull. 2004. Graphviz and Dynagraph — Static and Dynamic Graph Drawing Tools. In *Graph Drawing Software*, Michael Jünger and Petra Mutzel (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 127–148.

[16] Christos Faloutsos, Kevin S. McCurley, and Andrew Tomkins. 2004. Fast Discovery of Connection Subgraphs. In *Proc. of the Tenth ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD '04)*. ACM, New York, NY, USA, 118–127.

[17] Jonathan P. Forsyth, Rachel M. Bittner, Michael Musick, and Juan P. Bello. 2015. Improving and adapting finite state transducer methods for musical accompaniment. In *41st Int. Computer Music Conf., ICMC 2015*, Richard Dudas (Ed.). Int. Computer Music Association, Denton, TX, 290–297.

[18] Emden R Gansner, Yehuda Koren, and Stephen C North. 2005. Topological fisheye views for visualizing large graphs. *IEEE Transactions on Visualization and Computer Graphics* 11, 4 (2005), 457–468.

[19] Emden R Gansner, Eleftherios Koutsofios, Stephen C North, and K-P Vo. 1993. A technique for drawing directed graphs. *IEEE Transactions on Software Engineering* 19, 3 (1993), 214–230.

[20] Félix García, Mª Ángeles Moraga, Manuel Serrano, and Mario Piattini. 2015. Visualisation environment for global software development management. *IET Software* 9, 2 (2015), 51–64.

[21] Gérard Genette. 1983. *Narrative discourse: An essay in method.* Vol. 3. Cornell University Press, 512 East State St., Ithaca, NY.

[22] Ilka H Gleibs. 2017. Are all "research fields" equal? Rethinking practice for the use of data from crowdsourcing market places. *Behavior Research Methods* 49, 4 (2017), 1333–1342.

[23] Fethallah Hadjila, Amine Belabed, and Mohammed Merzoug. 2019. Flexible service discovery based on multiple matching algorithms. *Int. Journal of Web Engineering and Technology* 14, 4 (2019), 315–340.

[24] Aric Hagberg, Pieter Swart, and Daniel S Chult. 2008. *Exploring network structure, dynamics, and function using NetworkX.* Technical Report. Los Alamos National Lab. (LANL).

[25] Jeffry Halverson, Steven Corman, and H Lloyd Goodall. 2011. *Master narratives of Islamist extremism.* Springer, 175 5th Ave., New York, NY, USA.

[26] David Hauser, Gabriele Paolacci, and Jesse J Chandler. 2018. Common Concerns with MTurk as a Participant Pool: Evidence and Solutions.

[27] Jinwon Ho and Rong Tang. 2001. Towards an Optimal Resolution to Information Overload: An Infomediary Approach. In *Proc. of the 2001 Int. ACM SIGGROUP Conference on Supporting Group Work (GROUP '01)*. ACM, New York, NY, USA, 91–96.

[28] Thore Husfeldt. 1995. Fully Dynamic Transitive Closure in plane dags with one source and one sink. In *Algorithms — ESA '95*, Paul Spirakis (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 199–212.

[29] Dongwoo Kim and Alice Oh. 2011. Topic Chains for Understanding a News Corpus. In *Computational Linguistics and Intelligent Text Processing*, Alexander Gelbukh (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 163–176.

[30] Catherine Kohler Riessman. 2000. Analysis of personal narratives. In *Qualitative research in social work*. Columbia University Press, 61 West 62 St., New York, NY, USA, Chapter 7, 168–191.

[31] Walter S. Lasecki, Mitchell Gordon, Danai Koutra, Malte F. Jung, Steven P. Dow, and Jeffrey P. Bigham. 2014. Glance: Rapidly Coding Behavioral Video with the Crowd. In *Proc. of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 551–562.

[32] Tianyi Li, Chandler J Manns, Chris North, and Kurt Luther. 2019. Dropping the baton? Understanding errors and bottlenecks in a crowdsourced sensemaking pipeline. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–26.

[33] Tianyi Li, Asmita Shah, Kurt Luther, and Chris North. 2018. Crowdsourcing Intelligence Analysis with Context Slices. In *CHI 2018 Sensemaking Workshop*.

[34] Bang Liu, Di Niu, Kunfeng Lai, Linglong Kong, and Yu Xu. 2017. Growing Story Forest Online from Massive Breaking News. In *Proc. of the 2017 ACM on Conference on Information and Knowledge Management (CIKM '17)*. ACM, New York, NY, USA, 777–785.

[35] Leland McInnes, John Healy, and Steve Astels. 2017. hdbscan: Hierarchical density based clustering. *Journal of Open Source Software* 2, 11 (2017), 205.

[36] Leland McInnes, John Healy, and James Melville. 2018. Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426.

[37] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon. 2002. Network motifs: simple building blocks of complex networks. *Science* 298, 5594 (2002), 824–827.

[38] Alister Miskimmon, Ben O'loughlin, and Laura Roselle. 2014. *Strategic narratives: Communication power and the new world order*. Routledge, 711 3rd Ave. #8, New York, NY, USA.

[39] Alister Miskimmon, Ben O'Loughlin, and Laura Roselle. 2017. *Forging the world: Strategic narratives and international relations*. University of Michigan Press, 839 Greene St., Ann Arbor, MI.

[40] Tanushree Mitra and Eric Gilbert. 2015. CREDBANK: A Large-Scale Social Media Corpus With Associated Credibility Annotations. *Int. AAAI Conf. on Web and Social Media* 2015 (2015), 258–267.

[41] Tanushree Mitra, C.J. Hutto, and Eric Gilbert. 2015. Comparing Person- and Process-Centric Strategies for Obtaining Quality Data on Amazon Mechanical Turk. In *Proc. of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1345–1354.

[42] Ramesh Nallapati, Ao Feng, Fuchun Peng, and James Allan. 2004. Event Threading within News Topics. In *Proc. of the Thirteenth ACM Int. Conf. on Information and Knowledge Management (CIKM '04)*. ACM, New York, NY, USA, 446–453.

[43] Alexis R Neigel, Justine P Caylor, Sue E Kase, Michelle T Vanni, and Jefferson Hoye. 2018. The role of trust and automation in an intelligence analyst decisional guidance paradigm. *Journal of Cognitive Engineering and Decision Making* 12, 4 (2018), 239–247.

[44] A. Conrad Nied, Leo Stewart, Emma Spiro, and Kate Starbird. 2017. Alternative Narratives of Crisis Events: Communities and Social Botnets Engaged on Social Media. In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17 Companion)*. ACM, New York, NY, USA, 263–266.

[45] Brian Keith Norambuena, Michael Horning, and Tanushree Mitra. 2020. Evaluating the Inverted Pyramid Structure through Automatic 5W1H Extraction and Summarization. In *Proc. of the 2020 Computation + Journalism Symposium*. Computation + Journalism 2020, Boston, MA, USA, 1–7.

[46] Gabriele Paolacci, Jesse Chandler, and Panagiotis G Ipeirotis. 2010. Running experiments on amazon mechanical turk. *Judgment and Decision making* 5, 5 (2010), 411–419.

[47] Wim Pijls and Rob Potharst. 2013. Another Note on Dilworth's Decomposition Theorem. *Journal of Discrete Mathematics* 2013 (2013), 692–645.

[48] Kent Puckett. 2016. *Narrative theory*. Cambridge University Press, One Liberty Plaza, New York, NY, USA.

[49] Shlomith Rimmon-Kenan. 2003. *Narrative fiction: Contemporary poetics*. Routledge, 711 3rd Ave. #8, New York, NY, USA.

[50] Corina Robbins. 2019. *The Weaponization of Social Media*. Technical Report. Mercy Corps.

[51] Laura Roselle, Alister Miskimmon, and Ben O'Loughlin. 2014. Strategic narrative: A new means to understand soft power. *Media, War & Conflict* 7, 1 (2014), 70–84.

[52] Niloufar Salehi, Lilly C. Irani, Michael S. Bernstein, Ali Alkhatib, Eva Ogbe, Kristy Milland, and Clickhappier. 2015. We Are Dynamo: Overcoming Stalling and Friction in Collective Action for Crowd Workers. In *Proc. of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1621–1630.

[53] Dafna Shahaf and Carlos Guestrin. 2010. Connecting the Dots between News Articles. In *Proc. of the 16th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD '10)*. ACM, New York, NY, USA, 623–632.

[54] Dafna Shahaf and Carlos Guestrin. 2012. Connecting two (or less) dots: Discovering structure in news articles. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 5, 4 (2012), 1–31.

[55] Dafna Shahaf, Carlos Guestrin, and Eric Horvitz. 2012. Trains of Thought: Generating Information Maps. In *Proc. of the 21st Int. Conf. on World Wide Web (WWW '12)*. ACM, New York, NY, USA, 899–908.

[56] Dafna Shahaf, Carlos Guestrin, and Eric Horvitz. 2013. Metro Maps of Information. *SIGWEB Newsl.* Spring, Article 4 (2013), 9 pages.

[57] Dafna Shahaf, Carlos Guestrin, Eric Horvitz, and Jure Leskovec. 2015. Information cartography. *Commun. ACM* 58, 11 (2015), 62–73.

[58] Dafna Shahaf, Jaewon Yang, Caroline Suen, Jeff Jacobs, Heidi Wang, and Jure Leskovec. 2013. Information Cartography: Creating Zoomable, Large-Scale Maps of Information. In *Proc. of the 19th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD '13)*. ACM, New York, NY, USA, 1097–1105.

[59] Azra Shamim, Vimala Balakrishnan, and Muhammad Tahir. 2015. Evaluation of opinion visualization techniques. *Information visualization* 14, 4 (2015), 339–358.

[60] Barbara Herrnstein Smith. 1980. Narrative versions, narrative theories. *Critical inquiry* 7, 1 (1980), 213–236.

[61] Sandeep Soni, Tanushree Mitra, Eric Gilbert, and Jacob Eisenstein. 2014. Modeling Factuality Judgments in Social Media Text. In *Proc. of the 52nd Annual Meeting of the ACL (Volume 2: Short Papers)*. ACL, Baltimore, Maryland, 415–420.

[62] Joanna Szostek. 2017. Defence and Promotion of Desired State Identity in Russia's Strategic Narrative. *Geopolitics* 22, 3 (2017), 571–593.

[63] Barbara Tversky. 2004. Narratives of Space, Time, and Life. *Mind & Language* 19, 4 (2004), 380–392.

[64] Lu Wang, Claire Cardie, and Galen Marchetti. 2015. Socially-Informed Timeline Generation for Complex Events. In *Proc. of the 2015 Conference of the North American Chapter of the ACL: Human Language Technologies*. ACL, Denver, Colorado, 1055–1065.

[65] Jen Weedon, William Nuland, and Alex Stamos. 2017. *Information operations and Facebook*. Technical Report. Facebook.

[66] Vanessa Williamson. 2016. On the ethics of crowdsourced research. *PS: Political Science & Politics* 49, 1 (2016), 77–81.

[67] Tom Wilson, Kaitlyn Zhou, and Kate Starbird. 2018. Assembling strategic narratives: Information operations as collaborative work within an online community. *Proc. of the ACM on HCI* 2, CSCW (2018), 1–26.

[68] Sarah E. Worth. 2008. Storytelling and Narrative Knowing: An Examination of the Epistemic Benefits of Well-Told Stories. *Journal of Aesthetic Education* 42, 3 (2008), 42–56.

[69] Shize Xu, Shanshan Wang, and Yan Zhang. 2013. Summarizing Complex Events: a Cross-Modal Solution of Storylines Extraction and Reconstruction. In *Proc. of the 2013 Conference on Empirical Methods in Natural Language Processing*. ACL, Seattle, Washington, USA, 1281–1291.

[70] Rui Yan, Xiaojun Wan, Jahna Otterbacher, Liang Kong, Xiaoming Li, and Yan Zhang. 2011. Evolutionary Timeline Summarization: A Balanced Optimization Framework via Iterative Substitution. In *Proc. of the 34th Int. ACM SIGIR Conference on Research and Development in Info. Retrieval (SIGIR '11)*. ACM, New York, NY, USA, 745–754.

[71] Christopher C Yang, Xiaodong Shi, and Chih-Ping Wei. 2009. Discovering event evolution graphs from news corpora. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 39, 4 (2009), 850–863.

[72] Amy X. Zhang, Lea Verou, and David Karger. 2017. Wikum: Bridging Discussion Forums and Wikis Using Recursive Summarization. In *Proc. of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 2082–2096.

[73] Houkui Zhou, Huimin Yu, Roland Hu, and Junguo Hu. 2017. A survey on trends of cross-media topic evolution map. *Knowledge-Based Systems* 124 (2017), 164–175.

[74] P. Zhou, B. Wu, and Z. Cao. 2017. EMMBTT: A Novel Event Evolution Model Based on TFxIEF and TDC in Tracking News Streams. In *2017 IEEE Second Int. Conf. on Data Science in Cyberspace (DSC)*. IEEE, Shenzhen, China, 102–107.

## A  MECHANICAL TURK JOB



Fig. 12.  Description of the MTurk job shown to the workers.
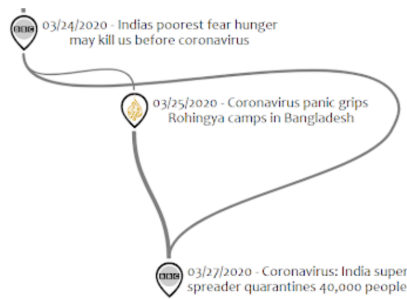


Fig. 13.  Narrative Map Description and Short Example

**Example Description and Analysis**

In this short example of a map, you see 3 news events represented by 3 landmarks. For example, the first event comes from the BBC and talks about how India's poorest fear hunger. The second event comes from Al-Jazeera and talks about panic in Rohingya camps. The third event talks about a super spreader in India which led to a quarantine of 40,000 people. The landmark icons show the news source that published those articles. Also, these events are joined together by edges, with a higher width representing a stronger relationship between two events. In this short example, we can quickly identify an important theme: the panic provoked by the Coronavirus spread (e.g., fear of food shortages and panic gripping the camps).

**Task Description**

In this task, we will consider the news narrative during the second month of the Coronavirus outbreak (February 2020). Your objective is to analyze the major themes that emerged in the news during this month. You can do this by looking at the main route and the representative landmarks highlighted in green. However, feel free to navigate through the map as you please.

This task should take between 10 to 15 minutes to complete.

Bigger and interactive version of the narrative map is available here:



Fig. 14. Example Analysis and Assigned Task

Please identify a major theme of the main route in the narrative map. *

Your answer

Please identify two major themes in the narrative map (different from the previous one).
Write down your responses in the separate text boxes below.

Theme 1 *

Your answer

Theme 2 *

Your answer

Fig. 15. Theme identification task.