Advances in Human Event Modeling: From Graph Neural Networks to Language Models

Songgaojun Deng
s.deng@uva.nl
University of Amsterdam
Amsterdam, The Netherlands

Maarten de Rijke
m.derijke@uva.nl
University of Amsterdam
Amsterdam, The Netherlands

Yue Ning
yue.ning@stevens.edu
Stevens Institute of Technology
Hoboken, New Jersey, USA

ABSTRACT

Human events such as hospital visits, protests, and epidemic outbreaks directly affect individuals, communities, and societies. These events are often influenced by factors such as economics, politics, and public policies of our society. The abundance of online data sources such as social networks, official news articles, and personal blogs chronicle societal events, facilitating the development of AI models for social science, public health care, and decision making. Human event modeling generally comprises both the forecasting stage, which estimates future events based on historical data, and interpretation, which seeks to identify influential factors of such events to understand their causative attributes. Recent achievements, fueled by deep learning and the availability of public data, have significantly advanced the field of human event modeling.

This survey offers a systematic overview of deep learning technologies for forecasting and interpreting human events, with a primary focus on political events. We first introduce the existing challenges and background in this domain. We then present the problem formulation of event forecasting and interpretation. We investigate recent achievements in graph neural networks, owing to the prevalence of relational data and the efficacy of graph learning models. We also discuss the latest studies that utilize large language models for event reasoning. Lastly, we provide summaries of data resources, open challenges, and future research directions in the study of human event modeling.

CCS CONCEPTS

• **Information systems** → **Data mining**; • **Computing methodologies** → **Machine learning**; • **Applied computing** → *Sociology*.

KEYWORDS

Event Forecasting, Graph Neural Networks, Language Models

ACM Reference Format:

Songgaojun Deng, Maarten de Rijke, and Yue Ning. 2024. Advances in Human Event Modeling: From Graph Neural Networks to Language Models. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24), August 25–29, 2024, Barcelona, Spain.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3637528.3671466

KDD '24, August 25–29, 2024, Barcelona, Spain

ACM ISBN 979-8-4007-0490-1/24/08
https://doi.org/10.1145/3637528.3671466

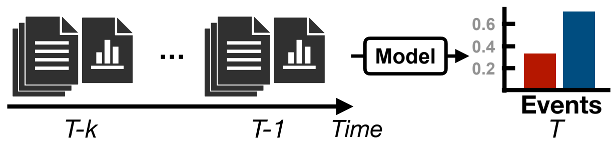


Figure 1: An example of event forecasting. Uses event indicators such as news articles and event statistics in the past to predict the occurrence of events of interest in the future.

1 INTRODUCTION

Human events can be broadly categorized into offline events (e.g., strikes) and online events (e.g., cyber attacks) [110]. Offline events occur at specific locations and times, impacting both local communities and global society in different ways. Understanding such events and their recurring patterns is an urgent issue for many stakeholders such as investors, suppliers, and decision-makers. Modeling such events through forecasting focuses on anticipating events in the future based on historical event indicators as shown in Figure 1, which is different from retrospective studies such as event detection [104] and summarization [11]. Accurate and reliable prediction of future events is conducive to effective allocations of public resources, reducing economic loss and social damage. Generally, it can bring benefits to society and individuals, supporting effective disaster response and socio-economic growth.

The field of human event modeling, particularly in the realm of political event forecasting [20, 22, 23], has witnessed significant progress over the years. Early on, researchers primarily relied on statistical [68, 100] and machine learning methods [42, 58, 70] to analyze different types of data and predict future events. With the advent of deep learning [49] and the availability of large amounts of (multi-modal) data coupled with increased computational power, there has been a notable shift towards leveraging deep learning techniques. Many researchers have studied recurrent neural networks [33, 84], and attention models [27], leading to substantial advances over previous approaches.

Despite these achievements, several challenges exist in the domain of political event forecasting:

- **Leveraging heterogeneous data.** As the availability of open-source data continues to increase, researchers have begun to resort to heterogeneous data to develop predictive models [57, 85, 97]. Human events occur in a dynamic social environment, and their corresponding key information can be published in various forms, such as traditional news reports, social media, and government official reports. It brings unique challenges of efficient processing and learning from heterogeneous data to make accurate event predictions.

- **Studying complex event dependencies.** Human events exhibit geographical properties and also have a high degree of temporal dependency [111]. Modeling event contextual information requires an in-depth investigation of the spatiotemporal dependencies of events. Traditional methodologies have shown limitations in modeling complex event data [63, 84], which encourages the development of advanced models to address this challenge.
- **Interpreting event predictions.** Most machine learning models focus on improving predictive accuracy. However, interpreting predicted events is an equally important task, as it can assist practitioners in understanding prediction results and making reasonable and practical decisions. Data-driven event prediction models with adequate interpretability support downstream users in event analysis and decision-making.

In recent years, there has been a surge of advanced deep learning approaches in the human event domain in response to the challenges listed above. Graph neural networks [103], in particular, have attracted considerable attention because of their ability to effectively capture complex relationships and patterns inherent in event data. The structured nature of graph data also facilitates interpretation due to its intuitive format. Additionally, large language models (LLMs) [117] have emerged as powerful tools capable of generating text, answering questions, and performing various linguistic tasks. LLMs have begun to demonstrate effectiveness across diverse fields [13, 43, 87], including human event forecasting [82, 109].

This survey concentrates on recent advances in graph neural networks and language models for human event modeling. We delve into the methodologies, benefits, and challenges associated with these two prominent paradigms, providing a comprehensive overview of recent progress and future directions in the field.

1.1 Previous Work and Contributions

There are several other surveys on human event predictions. Phillips et al. [65] provided a literature review that examines the problems and techniques for predictive analysis using social media data. Recently, Zhao [110] provided a systematic survey of existing data-driven event prediction methods, covering challenges, techniques, applications, evaluations, and open problems. As event prediction methods are typically motivated by specific application areas, some surveys instead focus on event technologies in particular domains. Ganar and Ardhapurkar [28] summarized studies of civil unrest prediction in social media analysis based on keyword filtering.

In contrast to previous work, this survey focuses on recent advances in graph neural networks and language models specifically in forecasting political events. Political events are influential human events dependent on evolving social environments. The main contributions of this work include (1) A systematic review of current graph neural network models in human event modeling. These models are categorized based on their emphasis on fundamental graph operations or additional data enhancements. (2) A summary of recent advancements in large language models assisting human event prediction. (3) A comprehensive summary of data resources that will help researchers and human experts in several ways, such as selecting appropriate data sources for specific research needs. (4) A discussion on ongoing challenges and prospects in the field.

1.2 Outline

The rest of this paper is organized as follows. Section 2 provides background information about graph neural networks and language models. Section 3 formulates the problem of human event prediction. Section 4 provides a succinct summary of early approaches, followed by a comprehensive review of graph neural networks and language models developed for human event forecasting. This section also discusses interpretation methods. Section 5 summarizes data resources related to political events. Section 6 lists open challenges and opportunities for future research directions. The survey concludes with a summary in Section 7.

2 BACKGROUND

2.1 Graph Neural Networks

Graph neural networks (GNNs) aim to project elements (e.g., nodes, edges) in a graph to a low-dimensional continuous space while preserving graph structure and inherent properties. Such models have gone through significant developments and have been applied to a wide range of fields such as computer vision [81], natural language processing [105], bioinformatics [71], etc.

GNNs can be categorized into spectral-based and spatial-based models depending on the types of convolutions. Spectral-based GNNs have graph signal filters in the spectral domain [73, 83]. In this line of research, Defferrard et al. [21] proposed ChebNet, which uses K-polynomial filters in the convolutional layers for localization. Later on, Graph convolutional networks (GCNs) [46] were proposed and introduced a first-order approximation of ChebNet. Spatial-based methods define graph convolutions based on a node's spatial relations [29, 60]. In this field, the aggregation-based inductive representation learning model (GraphSAGE) [35] and attention-based graph neural network (GAT) [92] were proposed to effectively learn graph representations and solve various graph-related tasks.

Some studies have attempted to extend GNNs to model heterogeneous graphs [108], which contain multiple types of nodes or edges and have become ubiquitous in real-world scenarios (e.g., knowledge graphs). The relational graph convolutional network (RGCN) [75] was proposed to model knowledge graphs by learning different weight matrices for each edge type. The composition-based multi-relational graph convolutional network (CompGCN) [90] was designed to embed both nodes and edges in a relational graph.

2.2 Language Models

A language model (LM) is designed to predict the probability of a sequence of words or characters in a language. It learns the patterns and structures of natural language by analyzing large corpora of text data. Language models are widely used in natural language processing tasks such as machine translation, text generation, speech recognition, and sentiment analysis.

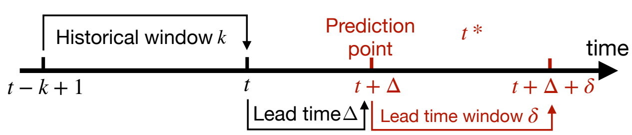
One of the most significant advancements in language models is the development of the transformer [91], which has led to the emergence of large language models (LLMs) [117]. LLMs are transformer-based language models that contain hundreds of billions (or more) parameters, trained on massive text datasets [80], such as GPT-3 [7], PaLM [16] and LLaMA [88]. These LLMs exhibit remarkable capabilities in understanding natural language and performing complex tasks, e.g., they excel in generating coherent and contextually relevant text that closely resembles human ability. LLMs have been explored in various domains, such as in medicine [87], education [43], and programming [13].

3 PROBLEM FORMULATION

Event forecasting can be formulated as a supervised learning task in machine learning, which aims at learning a function that maps an input (e.g., a historical window of records) to an output (e.g., an event occurrence in the future). To achieve this, the process uses labeled training data, which includes a set of training instances. These instances contain historical data collected before the timestamp of the output variable. The goal is to infer a function from this data that can accurately predict future events. We first introduce the preliminaries, including terminology and mathematical notation.

Historical data. Suppose there are L locations (e.g., cities, states) of interest, and each location l can be represented by a set of features used for prediction. We divide features into two categories, *static* and *dynamic*.[1] Static features such as population and political ideology remain constant or change slowly over a long time, while dynamic features such as frequency of events or number of tweets expressing "angry" emotions are updated for each time interval t (e.g., day, week). Let $\mathcal{S}_l$ denote the set of static features of location l, and $\mathcal{X}_{t,l}$ be the collection of dynamic features for location l at time t. The collection of dynamic features from location l within a historical window (i.e., observing time window) with size k up to time t can be represented as $\mathcal{X}_{t-k+1:t,l} = (\mathcal{X}_{t-k+1,l}, \ldots, \mathcal{X}_{t,l})$.

Ground-truth event occurrence. A target variable $\mathcal{Y}_{t^*,l}$ indicates the occurrence of a future political event (e.g., civil unrest) for each location l at time t^*. Note that t^* can be either a time point $(t+\Delta)$ or a time window in the future $(t+\Delta : t+\Delta+\delta)$. $\Delta \geq 1$ is the **lead time** that denotes the number of time steps in advance for a prediction. We use $\delta \geq 0$ to denote the **lead time window** that represents whether an event will occur between time $t+\Delta$ and $t+\Delta+\delta$.



Definition 3.1. **Binary event prediction**. Given static and dynamic input features, learn a classifier $f(\mathcal{S}_l, \mathcal{X}_{t-k+1:t,l}) \rightarrow \mathcal{Y}_{t^*,l}$ that maps the input to a binary event variable $\mathcal{Y}_{t^*,l} \in \{0,1\}$ at the future time t^* for the target location l.

Definition 3.2. **Concurrent event prediction**. Given static and dynamic input features, learn a classifier $f(\mathcal{S}_l, \mathcal{X}_{t-k+1:t,l}) \rightarrow \mathcal{Y}_{t^*,l}$ that maps the input to a binary vector for events $\mathcal{Y}_{t^*,l} \in \{0,1\}^M$ at the future time t^* for the target location l. Here, $M > 2$ denotes the number of event types that can occur concurrently (e.g., Appeal for judicial cooperation, accuse of crime). This problem is typically formulated as a multi-label classification task.

For *a concrete example* of the problem formulation, consider the GSR dataset [70] (More dataset details are in Section 5) and the binary event prediction task. We use the collection of n_t news articles related to protest published on a given day t at city l by $\mathcal{X}_{t,l} = \{\mathbf{x}_{t,1}, \ldots, \mathbf{x}_{t,n_t}\}$ where the j-th news article is represented by $\mathbf{x}_{t,j}$. No static features are used thus $\mathcal{S}_l = \emptyset$. Given the historical window size k, the input is $\mathcal{X}_{t-k+1:t,l} = (\mathcal{X}_{t-k+1,l}, \ldots, \mathcal{X}_{t,l})$ including news articles at city l in the previous k days up to day t. To predict the protest on day $t+\Delta$, the output is $\mathcal{Y}_{t+\Delta,l} \in \{0,1\}$.

Difference from temporal knowledge graph completion. With the rise of heterogeneous graph neural networks [75, 90], many recent methods for human event forecasting use temporal knowledge graphs (TKGs) [9] to enhance prediction and interpretability. This brings the human event prediction problem close to temporal knowledge graph completion (TKGC) [9]. TKGC, a subfield of knowledge graph completion (KGC) [40] is dedicated to predicting missing links or entities in evolving knowledge graphs over time. For instance, it seeks to predict the object $(s, r, ?, T)$, relation $(s, ?, o, T)$, or subject $(?, r, o, T)$ using all past event quadruples $\{(s_i, r_i, o_i, t_i)\}_i, \forall t_i < T$. TKGC is often considered a ranking problem rather than a classification problem. While both TKGC and human event forecasting aim to uncover patterns and dependencies within temporal data, they have distinct objectives and methodologies. Human event forecasting specifically targets the prediction of specific events based on temporal data sources such as news articles [22, 95], social media posts [102, 111], and historical event records [23, 30].

4 METHODOLOGY

4.1 Early Approaches

Over the years, researchers have leveraged various predictive techniques for human event predictions ranging from statistical methods to more sophisticated methods such as deep neural networks.

4.1.1 Statistical methods. In a pioneering study, Radinsky and Horvitz [68] mined chains of events from massive news archives and proposed a probabilistic method that predicts the likelihood of future worldwide events of interest. Manrique et al. [55] introduced a simple threshold-based methods for forecasting civil unrest. Jin et al. [41] characterized mass protest propagation using a bispace diffusion model. Chen and Neill [12] explored a nonparametric graph scan algorithm to the problem of civil unrest detection and forecasting using heterogeneous social media graphs. Temporal statistical models such as autoregressive [77, 100, 107] and hidden Markov models (HMM) were also proposed [66, 67, 111] to model political events.

4.1.2 Machine learning methods. Subsequently, people applied traditional machine learning models to civil unrest prediction, such as random forests [42] and logistic regression [8, 47, 70, 102, 113]. Early model based event recognition using surrogates (EMBERS) [58, 70] is an automated system developed for generating forecasts about civil unrest from massive and multiple data sources. More advanced methodologies such as multi-task learning [30, 62, 112–115] and multi-instance learning [61] were incorporated in forecasting spatial-temporal protest events. More recently, Zhao et al. [116] presented a group-Lasso based hierarchical feature learning model to characterize feature dependence, feature sparsity, and interactions among missing values.

[1]Dynamic features are usually involved in civil unrest prediction studies, while static features are sometimes not considered.

4.1.3 Early deep learning methods. A number of deep learning-based approaches have been proposed to predict political events and have demonstrated improved predictive capability and interpretability. Due to the temporal nature of event occurrence, researchers utilize variant recurrent neural networks (such as LSTM [38] and GRU [17]) to forecast political events, which have shown superior expressiveness and power compared to autoregressive models. Smith et al. [84] and Halkia et al. [33] applied LSTM models to predict material conflict events such as armed attacks and destruction of property. Parrish et al. [64] studied a GRU-based multi-feature driven approach to predict disruptive events. Meng and Srihari [57] combines convolutional layers and LSTM layers for predicting civil unrest. To tackle the challenges of limited predictive ability and explainability in RNN-based approaches, researchers have integrated attention mechanisms [4] into RNN models. Wang et al. [96] proposed a context-aware attention-based LSTM framework to study different contributions of data points in the time series for predicting civil unrest events. Attention mechanisms [4] were introduced to dynamically highlight relevant features of the input data, mimicking cognitive attention in humans. Such methods enhance the important parts of input data and fade out the rest. Ertugrul et al. [27] introduced a hierarchical attention-based spatiotemporal learning approach for predicting future protest occurrences and explaining feature importance.

4.2 Graph Neural Networks

There has been a surge of research focusing on studying graph neural networks for modeling political events. To provide a comprehensive overview of this body of work, we categorize recent studies into three groups based on their approach to graph learning: (1) **Vanilla graph learning**, which focuses solely on basic graph operations without additional data enhancements. (2) **Graph learning with contextual information**, which incorporates contextual data such as text summaries to enrich graph-based models. (3) **Graph learning with causal reasoning**, which integrates causal effect estimation or causal relationships to enhance interpretability and accuracy. This categorization aims to offer valuable insights into the diverse and popular methodologies used in leveraging graph neural networks for political event prediction.

4.2.1 Vanilla graph learning. Methods in this category primarily concentrate on graph operations, which involve leveraging node and/or edge embeddings as well as graph-level embedding techniques. Vanilla graph learning methods for human event forecasting typically follow four key steps:

- **Graph construction**. Begin by constructing a sequence of graphs from historical event data $(G_{t-k+1}, \ldots, G_t)$ where k is the sequence length. $G_t = (V_t, E_t)$ is the graph at time t with the node set V and edge set E. The graphs can be represented using adjacency matrices or tuple/quadruple sets. Next, define features for graph elements such as nodes and/or edges. For example, we can use word embeddings for word node features at time t that can be denoted as $\mathbf{X}_t^{N\times d}$, where $N = |V_t|$ is the number of words and d is the feature dimension.
- **Graph learning**. Update node and/or edge representations based on information from neighboring nodes. For instance, the procedures for updating nodes can be written as follows:
$$\mathbf{H}_t^{(l)}[i] \leftarrow \underset{\forall j\in\mathcal{N}_t(i),\forall e\in E_t(j,i))}{\mathrm{AGG}} \left(\mathrm{EX}\left(\mathbf{H}_t^{(l-1)}[i];\mathbf{H}_t^{(l-1)}[j],e\right)\right) \quad (1)$$
where $\mathbf{H}_t^{(l)}[i]$ is the node representation of node i at the l-th GNN layer. $\mathcal{N}_t(i)$ is the set of neighboring nodes of i at time t, and $E_t(i,j)$ denotes all the edges from node j to i at time t. EX$(\cdot)$ represents the neighbor information extractor. AGG$(\cdot)$ gathers the neighborhood information of source nodes via some aggregation operators, such as mean, sum, max, or more sophisticated pooling and normalization functions.
- **Temporal dependency learning**. This step involves capturing temporal dependencies from past events to effectively forecast future events. Techniques commonly employed include time series analysis, recurrent neural networks, attention mechanisms, or customized temporal methods.
- **Output prediction**. Use the learned node or graph representations for downstream tasks, such as predicting binary events or concurrent events, by applying an appropriate output layer.

Following the aforementioned framework, Deng et al. [22] proposed the first work to integrate graph learning into human events modeling. This work develops a dynamic graph convolutional network (DynamicGCN) [22] for predicting protest events and identifying key context graphs to understand their progression. The graph construction is achieved by an encoding method that encodes historical news articles into a sequence of undirected and weighted semantic word graphs, where each node is a keyword, and the weighted edge between two words is calculated by point-wise mutual information (PMI) [18]. Pre-trained word embeddings are used as initial node features. The dynamic graph model proposed in this study consists of novel temporal encoded features to re-encode input features for graph convolutional layers at each time step. These features encompass semantic information from word nodes and learned graph embeddings that aggregate neighboring information from previous graphs. The graph-level embeddings, generated at the final layer, are fed into an output layer for binary event classification.

Chen and Wang [15] proposed GasNet, a graphical and sequential network. Instead of learning from word semantic graphs, this work constructs dynamic event graphs comprising three primary types of nodes: attribute, event, and date nodes. These nodes are organized into three levels, with each previous node type pointing to the subsequent node type, and the date node serves as the terminal node. The graph learning is inspired by RGCN [75] which is the first graph neural network designed to operate on relational graph structures. In this approach, the representation of the date node in the event graph is used as the feature of this date. A sequence of date features is then fed into convolutional layers and LSTM [38] for estimating civil unrest events in future days.

Yinsen et al. [106] proposed a temporal attention-based graph sequence feature learning model (TAGS) for interpretable event prediction. This work leverages dynamic relational event graphs (e.g., knowledge graphs) to learn compressed graph embeddings at different times utilizing the CompGCN [90]. The authors employed a temporal-aware attention mechanism inspired by the transformer [91] to capture temporal dependencies in graph sequences.

Given the effectiveness of attention mechanisms in handling temporal data, researchers also introduced a hierarchical attention-based feature learning framework (HAFL) [98] for protest event forecasting. HAFL integrates attention mechanisms at the graph structure, node, and temporal level to enhance event prediction performance and explanation.

4.2.2 *Graph learning with contextual information.*

Given the vast and diverse open-source data available today, researchers have been actively exploring the integration of additional features to enhance graph learning in the context of event forecasting. These additional features provide contextual information about events and focus on non-causal factors, contrasting with the methods discussed in the next section. We classify the methods in this section into two categories based on when contextual information enhancement occurs: *context in graph construction* and *context in graph learning*.

Context in graph construction. Huai et al. [39] proposed a spatial and temporal knowledge graph neural network (STKGN) to explore both trans-regional influence and temporal sequencing patterns. Specifically, this work introduces a novel spatial-temporal event graph with cross-regional connection, where each region is denoted as a node and trans-regional influences are reflected by bidirectional edges. STKGN extracts semantics from event descriptions to enhance event representations, achieved by a text convolution procedure to learn low-dimensional vectors, inspired by textCNN [14]. The approach further introduces a continuous-time dynamic graph neural network to simulate and forecast the evolving process of entities.

Instead of using additional textual information in graph construction, Ma et al. [53] assumed the availability of external prior knowledge, i.e., the categorical context such as Covid-19, Olympics 2016, and G20 2022 Summit. They designed a novel framework, separation and collaboration graph disentanglement (SeCoGD) for context-aware event forecasting. SeCoGD is a two-stage framework including separation and collaboration. The separation stage uses the context as prior guidance to disentangle the event graph into multiple sub-graphs, followed by a context-specific modeling module for capturing each context's relational and temporal patterns. The collaboration stage leverages hypergraphs to model the cross-context collaborative associations, and then perform context-aware prediction and optimization.

Context in graph learning. Deng et al. [23] proposed Glean, a graph learning framework based on event knowledge graphs to incorporate both relational and word contexts. The proposed method is for predicting concurrent events of multiple types and event participants. The authors utilized event tuples and corresponding text summaries for graph construction and learning. For example, (*Citizen*, *Criticizes*, *Government*, 02/26/2015) with description "*A Politician attacked the state government on various fronts such as fertilizer crunch and land acquisition act.*". The authors construct temporal event knowledge graphs built upon a sequence of event sets in ascending time order, and a sequence of word graphs using text summaries. The model uses GCN [46] and CompGCN [90] to learn node embeddings from event knowledge graphs and semantic word graphs, respectively at each historical timestamp. To leverage the rich information encompassed in event texts, a context-aware embedding fusion module is proposed to enhance representations of nodes and edges in event knowledge graphs by blending embeddings of contextual word nodes. The fusion procedures are achieved through attention [52]. The enhanced representations of nodes and edges, as well as node embeddings of word graphs are then aggregated (e.g., pooling) and fed into a recurrent neural network for final event prediction. This approach can also infer the potential participants of the event of interest, providing additional information on the event prediction.

Deng et al. [24] introduced a contextualized multilevel feature learning framework (CMF) for event prediction and explanation by leveraging diverse contextual information in events. This approach models features ranging from coarse to refined granularity, including event frequencies, documents, and event graphs for a historical window denoted as $\{\mathbf{x}, D, G\}_{t-k+1:t}$. The approach hierarchically models heterogeneous data, capturing dependencies between different data types by propagating signals from higher-level (coarser) features to lower-level (refined) ones. Each type of feature is modeled at one level and then integrated into the next level. The authors also proposed an event explainer to provide post-hoc temporal and multi-level explanations for the prediction model.

Most work is based on the Markov assumption that the probability of an event is only influenced by the state of its last time step (or recent history). Han and Ning [37] proposed a text-enriched graph learning model, MTG, that takes into account multiple temporal granularities beyond just recent events. The authors proposed integrating news texts as auxiliary features during graph learning. The proposed framework consists of a cache module to learn medium-term tendencies from past events and news texts, a memory module to learn long-term statistics from past events and cached memories, as well as a dynamic CompGCN module to capture short-term triggers through interactions among entities and the corresponding auxiliary news texts.

4.2.3 *Graph learning with causal reasoning.*

Causal reasoning is a promising direction for improving prediction accuracy and interpretability in event forecasting [32]. Unlike conventional machine learning methods that primarily focus on capturing correlations and patterns within data, models integrating causal reasoning mechanisms aim to discern causal relationships and/or causal effects among various factors. Explicitly modeling causal dependencies can offer deeper insights into the underlying mechanisms driving events, enabling more robust and interpretable predictions. In this section, we discuss recent studies in graph learning with causal reasoning. Existing methods typically perform causal reasoning before the prediction stage, i.e., using a **two-stage approach**.

Deng et al. [26] proposed a novel framework CAPE, which incorporates causal inference into the prediction of future event occurrences in a spatiotemporal environment. CAPE presents a novel causal inference model to estimate conducts individual treatment effect (ITE) [79] from observational event data with spatiotemporal attributes. The learned event-related causal information is then incorporated into event prediction as prior knowledge. The prior causal knowledge (e.g., the estimated ITEs) is injected into the event forecasting stage via a feature reweighting module and an approximate constraint loss. The proposed method is validated on real-world event datasets by integrating learned causal prior knowledge into different base models for event forecasting.

Another study investigated the causal relationship between event occurrences and news topics. Researchers extracted topics from event-related documents and represented them as probability distributions of words [25]. They then introduced a causal discovery approach based on propensity score matching (PSM) [10] to discover evolving causal topics that causally impact future events from observational data. Such topics, together with words and documents are represented as nodes with changing edges in the dynamic heterogeneous graphs. The authors then proposed a dynamic heterogeneous graph model with causality-enhanced node representation, HGC [25] for forecasting civil unrest. To address temporal dependencies in dynamic graphs, they introduced a novel temporal information learning module that updates node representations based on their evolving context and heterogeneous semantics.

4.2.4 Advantages and limitations. GNN-based models can capture intricate interactions and dependencies in complex event data, thereby enhancing prediction accuracy. However, the computational demands for training GNNs on massive graphs pose scalability challenges, when considering the vast volume of event data. GNN models may also encounter difficulties in generalizing to unseen graphs [51]. This limitation can impede their ability to capture relevant patterns or dependencies in unfamiliar event contexts, such as underrepresented regions with limited training data.

4.3 Large Language Models

Large language models (LLMs) have recently garnered significant attention across various fields. These advanced models, capable of generating coherent and contextually relevant text, offer unique opportunities for extracting insights from vast amounts of textual data sourced from news articles, social media posts, and other textual sources. Researchers have started to explore LLMs for event predictive tasks. Next, we explore the burgeoning field of utilizing LLMs for human event forecasting, recognizing **LLMs as valuable tools** within the event forecasting pipeline.

In recent work, Shi et al. [82] investigated the capabilities of prompting large language models (e.g., GPT-3.5) in reasoning about real-world events. The authors proposed LAMP, a framework that integrates an LLM in event prediction. In the framework, an event sequence model first proposes predictions (e.g., what events will happen?). Then, a large language model suggests cause events, which will pattern-match against actual previous events and retrieve the most relevant. In the end, a neural model learns to assign high scores to the proposed predictions that are strongly supported by the retrieved evidence. LAMP has been demonstrated to outperform state-of-the-art event sequence models on real-world datasets significantly.

Ma et al. [54] explored LLMs in a different way for human event modeling. They introduced a novel formulation for structured, complex, and time-complete temporal event (SCTc-TE), along with a simple and fully automated pipeline for constructing such SCTcTEs from a large amount of news articles. A key step in this pipeline is entity extraction, where the authors employed LLMs for entity extraction in a zero-shot paradigm. Specifically, they took open-sourced Vicuna-13b for entity extraction and used GPT-4 for entity linking to merge the same entities. The LLM-based event construction method shows promise in replacing rule-based extraction systems and human-annotated processes. In addition, the authors propose a novel model, named LoGo [54], that leverages both local and global contexts for event forecasting based on SCTcTE.

Zhang and Ning [109] proposed two event forecasting tasks: object prediction and multi-event forecasting and presented a unified framework LEAF that uses LLMs to simplify the design of temporal event predictions.

4.3.1 Advantages and limitations. LLMs provide powerful tools for human event forecasting by extracting insights and performing reasoning from textual data. However, the answers from LLMs may lack factuality [93] and causality [45], as they generate responses based on statistical patterns rather than explicit knowledge. While LLMs have advantages in capturing linguistic nuances and contextual information, ensuring the accuracy and reliability of their outputs remains a challenge. Therefore, it is essential to critically evaluate LLM outputs in event forecasting applications.

4.4 Interpretation in Human Event Forecasting

Interpretation is crucial in human event forecasting, as it helps illuminate the underlying mechanisms driving predictions and provides actionable insights for decision-makers. We categorize the existing literature on event interpretation methods into three groups: *post-hoc non-trainable precursor discovery*, *dedicated explanation models* and *language model-driven reasoning*.

4.4.1 Post-hoc non-trainable precursor discovery. In this category, researchers have used separate methods to uncover clues of predicted events after the forecasting process [22, 98, 106]. These clues are extracted from historical data often using learned model weights to identify important elements, such as keywords and documents. Deng et al. [22] proposed a heuristic subgraph extraction to help explain event prediction results. It first extracts the important nodes from the trained model and then construct the subgraph of the input dynamic graph. Other researchers [98, 106] proposed to analyze subgraphs associated with important dates, thus potential clues affecting event occurrence can be mined. They used the product of attention values and the weights of MLP layers to identify the top-k important historical time steps. They then searched for subgraphs corresponding to high-frequency protest participants.

4.4.2 Dedicated explanation models. Such approaches employ separate trainable procedures to identify significant event precursors. Researchers introduced a separate multi-actor prediction model to estimate potential actors involved in the predicted event, given its relevant historical data [23], and a post-hoc multi-level event explanation module to extract relevant news articles and historical events related to predicted events via learning important feature masks [24].

4.4.3 Language model-driven reasoning. Some researchers have harnessed the powerful capabilities of LLMs for explaining event predictions. Shi et al. [82] proposed a method that first estimates the causes of predicted events using LLMs, followed by matching these causes with historical events to provide explanations. Although studies in this category are currently limited, it is anticipated that it will attract increased research efforts in the future.

5 DATA RESOURCES

We summarize data sources commonly used in human event studies, especially for political events. These data serve as the ground truth for event occurrences. Historical events are key indicators of future events. We also discuss external indicators used as historical input data for predicting events.

5.1 Event Data

Since the last century, various event data projects with different data collection, coding, and analysis processes have emerged. We organize political event data in Table 1. According to the coding process of event data, we divide these data into human-coded and machine-coded data. Human-coded data depend on human research teams with specific knowledge of the local context, while machine-coded data rely entirely on automated event encoding systems.

5.1.1 Human-encoded events. Many human-coded event datasets have been developed and maintained, which allow researchers to build forecasters at specific sub-state geographic units [57, 77, 100]. In the earliest days, due to technological limitations (i.e., the lack of electronic articles and computational power), the World Event Interaction Survey (WEIS) [56] and the Conflict and Peace Data Bank (COPDAB) [3] projects hire human analysts to physically collect newspaper clippings, press reports, and summary accounts from Western news sources to obtain news stories. These projects focus on daily international and domestic events or interactions.

Later, there were more event data projects focused on specific areas. The Social, Political, Economic Event Database (SPEED) Project [59] is a technology-intensive effort to extract event data from a global archive of news reports covering the Post WWII era. Salehyan et al. developed the Social Conflict in Africa Database (SCAD) [72], which contains instances of protests, riots, strikes, government repression, communal violence, and other forms of unrest that happened mainly in Africa. The Gold Standard Report (GSR) is a collection of human-classified civil unrest news reports from the most influential newspaper outlets in Latin America [70]. The Armed Conflict Location and Event Data Project (ACLED) [69] collects the dates, actors, locations, fatalities, and types of all reported political events (e.g., violence, protests) around the world. Urdal and Hoelscher [89] introduced an event dataset on urban unrest at the city level that covers 55 major cities in Asia and Sub-Saharan African countries.

Some datasets focus on violent events motivated by political grievances. Daly [19] collected violent events at the municipality-month level in Colombia. The Global Terrorism Database (GTD) [48] provides information on domestic and international terrorist attacks around the world. The Konstanz One-Sided Event Dataset (KOSVED) [76] provides detailed information on the magnitude and locations of one-sided violent events in 20 civil wars. The Uppsala Conflict Data Program Georeferenced Event Dataset (UCDP GED) [86] is an event dataset that classifies three types of organized violence (state-based conflict, non-state conflict, and one-sided violence) both spatially and temporally.

5.1.2 Machine-encoded events. Manual approaches began to be replaced with automated coding with the first iteration of the Kansas event data set (KEDS) [78] project in the late 1980s. KEDS uses the automated coding of English-language news reports to generate political event data. These data are used in statistical early warning models to predict political changes. The Integrated Crisis Early Warning System (ICEWS) [6] includes a database of political events with global coverage. Similar to ICEWS, the Global Dataset of Events, Location, and Tone (GDELT) [50] has been developed and compiled a comprehensive list of electronic news sources. Both GDELT and ICEWS are active automatic systems that identify and classify events from public data following the Conflict and Mediation Event Observations (CAMEO) [31] which is a framework for coding event data. These two datasets have been extensively studied in various fields. Some researchers compared ICEWS with GDELT [2, 94, 99] and pointed out limitations on local conflict processes that rely too heavily on machine-coded data [33, 36]. The Historical Phoenix Event Data [1] includes events extracted from The New York Times, BBC Monitoring's Summary of World Broadcasts, and the CIA's Foreign Broadcast Information Service. It also uses the CAMEO methodology to encode events. Given the large scale and extensive spatial coverage of machine-coded data, many researchers have used machine-coded event data to build forecasts for political eventst [22, 24–26, 30, 53, 62, 66, 67, 107, 112–115]. There is also an emerging trend in using LLMs for event encoding [54], driven by their superior linguistic comprehension compared to rule-based machine encoding methods, and their efficiency in saving human effort compared to manual encoding.

5.2 External Event Indicators

Researchers also incorporate public media data as input features for human event forecasting. Such data is known as Open Source Indicators (OSI). OSI includes traditional media data such as digital newspapers, blogs, and social media data such as posts from Twitter and Facebook. These data provide a wealth of background information that helps one understand the social context and public opinion of political events. Economic indicators and other meta-data sources have also been explored in this line of research. Studies show that exogenous political and economic variables can serve as the necessary underlying drivers of political events besides social media [57, 101]. Social and economic features derived from the World Development Indicators (WDI) [5] and Worldwide Governance Indicators (WGI) [44] have been investigated [64]. Researchers have also utilized Google Trends (GT) to uncover social dynamics associated with behavior that precedes episodes of civil unrest [55]. Google Trends analyzes the popularity of top search queries in Google Search across various regions and languages.

6 OPEN CHALLENGES AND FUTURE DIRECTIONS

Data dynamics, sufficiency, and reliability. Data-driven approaches for human event prediction depend heavily on data quality, making them subject to several data challenges. The dynamic nature of data is one of the main challenges. For text data, language, vocabulary, and mainstream slang are constantly evolving. In geographic data, location names and area boundaries may change due to major political events. The sufficiency of data is another challenge. Researchers have investigated various external data sources in addition to historical event occurrences to improve the accuracy

Table 1: A summary of political event datasets used for human event modeling. All listed datasets are labeled with geolocation information. The start/end time in temporal coverage indicates the earliest/latest time of data collection. "/" means unavailable.

Dataset	Event Types	Temporal coverage	Spatial coverage	Coding process	Open sourced
COPDAB [3]	Political events	1948-1978	Near global	Human	Yes
WEIS [56]	Political events	1966-1978	Global	Human	Yes
SPEED [59]	Social, political and economic events	1945-2008	Global	Human	Yes
SCAD [72]	Social conflict events	1990-2017	Africa and Latin America	Human	Yes
ACLED [69]	Political violence and demonstrations	1997-	Near global	Human	Yes
Daly [19]	Violent events/Rebellion	1964-1984	Colombia	Human	No
Urdal and Hoelscher [89]	Civil unrest events	1960-2009	Asia and Sub-Saharan Africa	Human	No
KOSVED [76]	One-sided violence	1991-2008	Africa and Europe	Human	Yes
GTD [48]	Terrorism	1970-2019	Global	Human	Yes
UCDP GED [86]	Organized violence	1989-2020	Near global	Human	Yes
GSR [70]	Civil unrest events	/	Latin America	Human	No
KEDS [78]	Political events	1979-1997	Middle East, Balkans, and West Africa	Machine	Yes
Phoenix [1]	Political events	1945-2019	Global	Machine	Yes
GDELT [50]	Various events	1979-	Global	Machine	Yes
ICEWS [6]	Political events	1995-	Global	Machine	Yes
AutoGSR [74]	Civil unrest events	/	Latin Americ	Machine	No
POLECAT [34]	Political events	2010-	Global	Machine	Yes

of predictions. Collecting external data from multiple sources and distinguishing correlated data from noisy data is expensive in terms of time, material, and computational costs. Moreover, the spatial scarcity of human events will also hinder event prediction studies in underrepresented areas. Data reliability is also a fundamental issue in forecasting problems. Missing or incorrect data often occurs during manual or automated data collection. For example, automated event collection systems may lose events due to unexpected network failures. Social media posts provide valuable resources for tracking user behavior and social activities. However, such data include typos, chit-chat, and misinformation that can mislead predictive models. Hence, over-reliance on data can make prediction models vulnerable to real-world applications.

Unfaithful explanations. While explainable models have emerged to improve model transparency and assist in decision-making, ensuring the faithfulness of event explanations remains a challenge. When explanations provided by interpretable models fail to reflect the true underlying factors driving predictions accurately, they can lead to misinterpretations and erroneous conclusions. This lack of faithfulness may arise due to various reasons, including oversimplification of complex relationships, biases in the training data, or limitations in the interpretability techniques themselves. Language models (LLMs) offer promise in this regard, with their ability to generate coherent and contextually relevant text. However, leveraging LLMs for event prediction requires careful consideration of their limitations (e.g., hallucinations and reliability).

Causality study in human events. There have been advancements in human event prediction that leverage estimated causal information to enhance event prediction [25, 26]. However, there remains a significant gap in deeply mining the underlying causal mechanisms driving event occurrences. More advanced approaches that delve into understanding the intricate causal relationships among various factors for human events are expected. By gaining deeper insights into the causal structure underlying human events, it becomes possible to develop more robust and accurate predictive models capable of capturing the true causal drivers of events, thereby advancing the field of human event forecasting

7 CONCLUSION

In this article, we present a comprehensive survey of current methods for human event modeling, with a particular emphasis on graph neural networks and language models. We outline the existing challenges in human event prediction, summarize recent research papers, and discuss both traditional and advanced predictive techniques studied for human events. Additionally, we offer an extensive overview of available data resources. At the end, we highlight open challenges and propose promising avenues for future investigation. With the emergence of large language models, which are changing the problem space, we anticipate a surge of research in this domain. We believe our survey offers a valuable overview for researchers, driving new ideas for future endeavors.

ACKNOWLEDGMENTS

This work is supported in part by Ahold Delhaize, by the Dutch Research Council (NWO) under project nrs. 024.004.022, NWA.1389.20.183, and KICH3.LTP.20.006, by the European Union's Horizon Europe program under grant agreement No 101070212, and by the US National Science Foundation under grant 2047843. All content represents the opinion of the authors, which is not necessarily shared or endorsed by their respective employers and/or sponsors.

REFERENCES

[1] Scott Althaus, Joseph Bajjalieh, John Carter, Buddy Peyton, and Dan Shalmon. 2020. Cline Center Historical Phoenix Event Data. Cline Center for Advanced Social Research. v1.3.0. May 4.

[2] Bryan Arva, John Beieler, Bejamin Fisher, Gustavo Lara, Philip A Schrodt, Wonjun Song, Marsha Sowell, and Sam Stehle. 2013. Improving Forecasts of International Events of Interest. In *EPSA 2013 Annual General Conference Paper*, Vol. 78.

[3] Edward E. Azar. 1980. The Conflict and Peace Data Bank (COPDAB) project. *Journal of Conflict Resolution* 24, 1 (1980), 143–152.

[4] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2014. Neural Machine Translation by Jointly Learning to Align and Translate. *arXiv preprint arXiv:1409.0473* (2014).

[5] World Bank. 2016. World Development Report 2016: Digital Dividends.

[6] Elizabeth Boschee, Jennifer Lautenschlager, Sean O'Brien, Steve Shellman, James Starz, and Michael Ward. 2015. ICEWS Coded Event Data.

[7] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language Models Are Few-shot Learners. *Advances in neural information processing systems* 33 (2020), 1877–1901.

[8] Jose Cadena, Gizem Korkmaz, Chris J Kuhlman, Achla Marathe, Naren Ramakrishnan, and Anil Vullikanti. 2015. Forecasting Social Unrest Using Activity Cascades. *PloS one* 10, 6 (2015), e0128879.

[9] Borui Cai, Yong Xiang, Longxiang Gao, He Zhang, Yunfeng Li, and Jianxin Li. 2022. Temporal Knowledge Graph Completion: A Survey. *arXiv preprint arXiv:2201.08236* (2022).

[10] Marco Caliendo and Sabine Kopeinig. 2008. Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of economic surveys* 22, 1 (2008), 31–72.

[11] Deepayan Chakrabarti and Kunal Punera. 2011. Event Summarization Using Tweets. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 5.

[12] Feng Chen and Daniel B Neill. 2014. Non-parametric Scan Statistics for Event Detection and Forecasting in Heterogeneous Social Media Graphs. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. 1166–1175.

[13] Mark Chen, Jerry Tworek, Heewoo Jun, Qiming Yuan, Henrique Ponde de Oliveira Pinto, Jared Kaplan, Harri Edwards, Yuri Burda, Nicholas Joseph, Greg Brockman, et al. 2021. Evaluating Large Language Models Trained on Code. *arXiv preprint arXiv:2107.03374* (2021).

[14] Yahui Chen. 2015. *Convolutional Neural Network for Sentence Classification*. Master's thesis. University of Waterloo.

[15] Zheng Chen and Yifan Wang. 2021. Civil Unrest Event Forecasting Using Graphical and Sequential Neural Networks. In *Artificial Neural Networks and Machine Learning–ICANN 2021: 30th International Conference on Artificial Neural Networks, Bratislava, Slovakia, September 14–17, 2021, Proceedings, Part III 30*. Springer, 192–203.

[16] Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, et al. 2023. Palm: Scaling Language Modeling with Pathways. *Journal of Machine Learning Research* 24, 240 (2023), 1–113.

[17] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. 2014. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *arXiv preprint arXiv:1412.3555* (2014).

[18] Kenneth Ward Church and Patrick Hanks. 1990. Word Asociation Norms, Mutual Information, and Lexicography. *Comput. Linguist.* 16, 1 (March 1990), 22–29.

[19] Sarah Zukerman Daly. 2012. Organizational Legacies of Violence: Conditions Favoring Insurgency Onset in Colombia, 1964–1984. *Journal of Peace Research* 49, 3 (2012), 473–491.

[20] Bruce Bueno De Mesquita, David Newman, and Alvin Rabushka. 1985. Forecasting Political Events. *The Future of Hong* (1985).

[21] Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. 2016. Convolutional neural networks on graphs with fast localized spectral filtering. *arXiv preprint arXiv:1606.09375* (2016).

[22] Songgaojun Deng, Huzefa Rangwala, and Yue Ning. 2019. Learning Dynamic Context Graphs for Predicting Social Events. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1007–1016.

[23] Songgaojun Deng, Huzefa Rangwala, and Yue Ning. 2020. Dynamic Knowledge Graph based Multi-event Forecasting. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1585–1595.

[24] Songgaojun Deng, Huzefa Rangwala, and Yue Ning. 2021. Understanding Event Predictions via Contextualized Multilevel Feature Learning. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. 342–351.

[25] Songgaojun Deng, Huzefa Rangwala, and Yue Ning. 2022. Causality Enhanced Societal Event Forecasting with Heterogeneous Graph Learning. In *2022 IEEE International Conference on Data Mining (ICDM)*. IEEE, 91–100.

[26] Songgaojun Deng, Huzefa Rangwala, and Yue Ning. 2022. Robust Event Forecasting with Spatiotemporal Confounder Learning. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 294–304.

[27] Ali Mert Ertugrul, Yu-Ru Lin, Wen-Ting Chung, Muheng Yan, and Ang Li. 2019. Activism via Attention: Interpretable Spatiotemporal Learning to Forecast Protest Activities. *EPJ Data Science* 8, 1 (2019), 1–26.

[28] Ruchika Parmarth Ganar and Shrikant Ardhapurkar. 2016. Prediction of Civil Unrest by Analysing Social Network Using Keyword Filtering: A Survey. In *2016 Online International Conference on Green Engineering and Technologies (IC-GET)*. IEEE, 1–4.

[29] Hongyang Gao, Zhengyang Wang, and Shuiwang Ji. 2018. Large-scale Learnable Graph Convolutional Networks. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 1416–1424.

[30] Yuyang Gao and Liang Zhao. 2018. Incomplete Label Multi-task Ordinal Regression for Spatial Event Scale Forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32.

[31] Deborah J Gerner, Philip A Schrodt, Omür Yilmaz, and Rajaa Abu-Jabr. 2002. Conflict and Mediation Event Observations (CAMEO): A New Event Data Framework for the Analysis of Foreign Policy Interactions. *International Studies Association, New Orleans* (2002).

[32] Ruocheng Guo, Lu Cheng, Jundong Li, P Richard Hahn, and Huan Liu. 2020. A Survey of Learning Causality with Data: Problems and Methods. *ACM Computing Surveys (CSUR)* 53, 4 (2020), 1–37.

[33] Matina Halkia, Stefano Ferri, Michail Papazoglou, Marie-Sophie Van Damme, and Dimitrios Thomakos. 2020. Conflict Event Modelling: Research Experiment and Event Data Limitations. In *Proceedings of the Workshop on Automated Extraction of Socio-political Events from News 2020*. 42–48.

[34] Andrew Halterman, Benjamin E Bagozzi, Andreas Beger, Phil Schrodt, and Grace Scraborough. 2023. PLOVER and POLECAT: A new political event ontology and dataset. In *International Studies Association Conference Paper*.

[35] William L Hamilton, Rex Ying, and Jure Leskovec. 2017. Inductive Representation Learning on Large Graphs. *arXiv preprint arXiv:1706.02216* (2017).

[36] Jesse Hammond and Nils B Weidmann. 2014. Using Machine-coded Event Data for the Micro-level Study of Political Violence. *Research & Politics* 1, 2 (2014), 2053168014539924.

[37] Xiaoxue Han and Yue Ning. 2022. Text-Enhanced Multi-granularity Temporal Graph Learning for Event Prediction. In *2022 IEEE International Conference on Data Mining (ICDM)*. IEEE, 171–180.

[38] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Computation* 9, 8 (1997), 1735–1780.

[39] Zepeng Huai, Guohua Yang, Jianhua Tao, et al. 2023. Spatial-temporal Knowledge Graph Network for Event Prediction. *Neurocomputing* 553 (2023), 126557.

[40] Shaoxiong Ji, Shirui Pan, Erik Cambria, Pekka Marttinen, and Philip S. Yu. 2021. A Survey on Knowledge Graphs: Representation, Acquisition, and Applications. *IEEE transactions on neural networks and learning systems* 33, 2 (2021), 494–514.

[41] Fang Jin, Rupinder Paul Khandpur, Nathan Self, Edward Dougherty, Sheng Guo, Feng Chen, B Aditya Prakash, and Naren Ramakrishnan. 2014. Modeling Mass Protest Adoption in Social Network Communities using Geometric Brownian Motion. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. 1660–1669.

[42] Nathan Kallus. 2014. Predicting Crowd Behavior with Big Public Data. In *Proceedings of the 23rd International Conference on World Wide Web*. 625–630.

[43] Enkelejda Kasneci, Kathrin Seßler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günnemann, Eyke Hüllermeier, et al. 2023. ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education. *Learning and individual differences* 103 (2023), 102274.

[44] Daniel Kaufmann, Aart Kraay, and Massimo Mastruzzi. 2011. The Worldwide Governance Indicators: Methodology and Analytical Issues1. *Hague journal on the rule of law* 3, 2 (2011), 220–246.

[45] Emre Kıcıman, Robert Ness, Amit Sharma, and Chenhao Tan. 2023. Causal Reasoning and Large Language Models: Opening a New Frontier for Causality. *arXiv preprint arXiv:2305.00050* (2023).

[46] Thomas N Kipf and Max Welling. 2017. Semi-supervised Classification with Graph Convolutional Networks. In *International Conference on Learning Representations*.

[47] Gizem Korkmaz, Jose Cadena, Chris J Kuhlman, Achla Marathe, Anil Vullikanti, and Naren Ramakrishnan. 2015. Combining Heterogeneous Data Sources for Civil Unrest Forecasting. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*. 258–265.

[48] Gary LaFree and Laura Dugan. 2007. Introducing the Global Terrorism Database. *Terrorism and Political Violence* 19, 2 (2007), 181–204.

[49] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. 2015. Deep Learning. 521, 7553 (2015), 436–444.

[50] Kalev Leetaru and Philip A Schrodt. 2013. GDELT: Global Data on Events, Location, and Tone, 1979–2012. In *ISA Annual Convention*, Vol. 2. Citeseer, 1–49.

[51] Haoyang Li, Xin Wang, Ziwei Zhang, and Wenwu Zhu. 2022. Ood-gnn: Out-of-distribution Generalized Graph Neural Network. *IEEE Transactions on Knowledge*

and Data Engineering (2022).
[52] Minh-Thang Luong, Hieu Pham, and Christopher D Manning. 2015. Effective Approaches to Attention-based Neural Machine Translation. *arXiv preprint arXiv:1508.04025* (2015).
[53] Yunshan Ma, Chenchen Ye, Zijian Wu, Xiang Wang, Yixin Cao, and Tat-Seng Chua. 2023. Context-aware Event Forecasting via Graph Disentanglement. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 1643–1652.
[54] Yunshan Ma, Chenchen Ye, Zijian Wu, Xiang Wang, Yixin Cao, Liang Pang, and Tat-Seng Chua. 2023. Structured, complex and time-complete temporal event forecasting. *arXiv preprint arXiv:2312.01052* (2023).
[55] Pedro Manrique, Hong Qi, Ana Morgenstern, Nicolas Velasquez, Tsai-Ching Lu, and Neil Johnson. 2013. Context Matters: Improving the Uses of Big Data for Forecasting Civil Unrest: Emerging Phenomena and Big Data. In *2013 IEEE International Conference on Intelligence and Security Informatics*. IEEE, 169–172.
[56] Charles McClelland. 1978. World Event/Interaction Survey, 1966-1978. *WEIS Codebook ICPSR* 5211 (1978).
[57] Lu Meng and Rohini K Srihari. 2019. Leveraging heterogeneous data sources for civil unrest prediction. Social, Cultural, and Behavioral Modeling.
[58] Sathappan Muthiah, Patrick Butler, Rupinder Paul Khandpur, Parang Saraf, Nathan Self, Alla Rozovskaya, Liang Zhao, Jose Cadena, Chang-Tien Lu, Anil Vullikanti, et al. 2016. Embers at 4 Years: Experiences Operating an Open Source Indicators Forecasting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 205–214.
[59] Peter F. Nardulli, Kalev H. Leetaru, and Matthew Hayes. 2011. Event Data, Civil Unrest and the Social, Political and Economic Event Database (SPEED) Project: Post World War II Trends in Political Protests and Violence. In *Annual meeting of the International Studies Association, Quebec, Canada*.
[60] Mathias Niepert, Mohamed Ahmed, and Konstantin Kutzkov. 2016. Learning Convolutional Neural Networks for Graphs. In *International conference on machine learning*. PMLR, 2014–2023.
[61] Yue Ning, Sathappan Muthiah, Huzefa Rangwala, and Naren Ramakrishnan. 2016. Modeling Precursors for Event Forecasting via Nested Multi-instance Learning. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 1095–1104.
[62] Yue Ning, Rongrong Tao, Chandan K. Reddy, Huzefa Rangwala, James C. Starz, and Naren Ramakrishnan. 2018. STAPLE: Spatio-temporal Precursor Learning for Event Forecasting. In *Proceedings of the 2018 SIAM International Conference on Data Mining*. SIAM, 99–107.
[63] Andreas M Olligschlaeger. 1997. Artificial Neural Networks and Crime Mapping. *Crime mapping and crime prevention* 1 (1997), 313.
[64] Nathan H. Parrish, Anna L. Buczak, Jared T. Zook, James P. Howard, Brian J. Ellison, and Benjamin D. Baugher. 2018. Crystal Cube: Multidisciplinary Approach to Disruptive Events Prediction. In *International Conference on Applied Human Factors and Ergonomics*. Springer, 571–581.
[65] Lawrence Phillips, Chase Dowling, Kyle Shaffer, Nathan Hodas, and Svitlana Volkova. 2017. Using Social Media to Predict the Future: A Systematic Literature Review. *arXiv preprint arXiv:1706.06134* (2017).
[66] Fengcai Qiao, Pei Li, Xin Zhang, Zhaoyun Ding, Jiajun Cheng, and Hui Wang. 2017. Predicting Social Unrest Events with Hidden Markov Models Using GDELT. *Discrete Dynamics in Nature and Society* 2017 (2017).
[67] Fengcai Qiao, Xin Zhang, and Jinsheng Deng. 2020. Learning Evolutionary Stages with Hidden Semi-Markov Model for Predicting Social Unrest Events. *Discrete Dynamics in Nature and Society* 2020 (2020).
[68] Kira Radinsky and Eric Horvitz. 2013. Mining the web to predict future events. In *Proceedings of the sixth ACM international conference on Web search and data mining*. 255–264.
[69] Clionadh Raleigh and Caitriona Dowd. 2015. Armed Conflict Location and Event Data Project (ACLED) codebook. *Find this resource* (2015).
[70] Naren Ramakrishnan, Patrick Butler, Sathappan Muthiah, Nathan Self, Rupinder Khandpur, Parang Saraf, Wei Wang, Jose Cadena, Anil Vullikanti, Gizem Korkmaz, et al. 2014. 'Beating the News' with EMBERS: Forecasting Civil Unrest Using Open Source Indicators. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 1799–1808.
[71] Sungmin Rhee, Seokjun Seo, and Sun Kim. 2017. Hybrid Approach of Relation Network and Localized Graph Convolutional Filtering for Breast Cancer Subtype Classification. *arXiv preprint arXiv:1711.05859* (2017).
[72] Idean Salehyan, Cullen S Hendrix, Jesse Hamner, Christina Case, Christopher Linebarger, Emily Stull, and Jennifer Williams. 2012. Social Conflict in Africa: A New Database. *International Interactions* 38, 4 (2012), 503–511.
[73] Aliaksei Sandryhaila and José MF Moura. 2013. Discrete Signal Processing on Graphs. *IEEE transactions on signal processing* 61, 7 (2013), 1644–1656.
[74] Parang Saraf and Naren Ramakrishnan. 2016. EMBERS autogsr: Automated Coding of Civil Unrest Events. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 599–608.
[75] Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne Van Den Berg, Ivan Titov, and Max Welling. 2018. Modeling Relational Data with Graph Convolutional Networks. In *The semantic web: 15th international conference, ESWC 2018, Heraklion, Crete, Greece, June 3–7, 2018, proceedings 15*. Springer, 593–607.
[76] Gerald Schneider and Margit Bussmann. 2013. Accounting for the Dynamics of One-sided Violence: Introducing KOSVED. *Journal of Peace Research* 50, 5 (2013), 635–644.
[77] Gerald Schneider, Margit Bussmann, and Constantin Ruhe. 2012. The Dynamics of Mass Killings: Testing Time-series Models of One-sided Violence in the Bosnian Civil War. *International Interactions* 38, 4 (2012), 443–461.
[78] Philip A Schrodt, Shannon G Davis, and Judith L Weddle. 1994. Political Science: KEDS—A Program for the Machine Coding of Event Data. *Social Science Computer Review* 12, 4 (1994), 561–587.
[79] Uri Shalit, Fredrik D Johansson, and David Sontag. 2017. Estimating Individual Treatment Effect: Generalization Bounds and Algorithms. In *International Conference on Machine Learning*. PMLR, 3076–3085.
[80] Murray Shanahan. 2024. Talking about Large Language Models. *Commun. ACM* 67, 2 (2024), 68–79.
[81] Yantao Shen, Hongsheng Li, Shuai Yi, Dapeng Chen, and Xiaogang Wang. 2018. Person Re-identification with Deep Similarity-guided Graph Neural Network. In *Proceedings of the European Conference on Computer Vision (ECCV)*. 486–504.
[82] Xiaoming Shi, Siqiao Xue, Kangrui Wang, Fan Zhou, James Zhang, Jun Zhou, Chenhao Tan, and Hongyuan Mei. 2024. Language Models Can Improve Event Prediction by Few-shot Abductive Reasoning. *Advances in Neural Information Processing Systems* 36 (2024).
[83] David I. Shuman, Sunil K. Narang, Pascal Frossard, Antonio Ortega, and Pierre Vandergheynst. 2013. The Emerging Field of Signal Processing on Graphs: Extending High-dimensional Data Analysis to Networks and Other Irregular Domains. *IEEE Signal Processing Magazine* 30, 3 (2013), 83–98.
[84] Emmanuel M. Smith, Jim Smith, Phil Legg, and Simon Francis. 2017. Predicting the Occurrence of World News Events Using Recurrent Neural Networks and Auto-regressive Moving Average Models. In *UK Workshop on Computational Intelligence*. Springer, 191–202.
[85] Alexander Stec and Diego Klabjan. 2018. Forecasting Crime with Deep Learning. *arXiv preprint arXiv:1806.01486* (2018).
[86] Ralph Sundberg and Erik Melander. 2013. Introducing the UCDP Georeferenced Event Dataset. *Journal of Peace Research* 50, 4 (2013), 523–532.
[87] Arun James Thirunavukarasu, Darren Shu Jeng Ting, Kabilan Elangovan, Laura Gutierrez, Ting Fang Tan, and Daniel Shu Wei Ting. 2023. Large Language Models in Medicine. *Nature Medicine* 29, 8 (2023), 1930–1940.
[88] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. 2023. Llama: Open and Efficient Foundation Language Models. *arXiv preprint arXiv:2302.13971* (2023).
[89] Henrik Urdal and Kristian Hoelscher. 2012. Explaining Urban Social Disorder and Violence: An Empirical Study of Event Data from Asian and Sub-saharan African Cities. *International Interactions* 38, 4 (2012), 512–528.
[90] Shikhar Vashishth, Soumya Sanyal, Vikram Nitin, and Partha Talukdar. 2019. Composition-based Multi-relational Graph Convolutional Networks. *arXiv preprint arXiv:1911.03082* (2019).
[91] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is All You Need. In *Advances in neural information processing systems*. 5998–6008.
[92] Petar Velickovic, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2018. Graph Attention Networks. *ICLR*.
[93] Cunxiang Wang, Xiaoze Liu, Yuanhao Yue, Xiangru Tang, Tianhang Zhang, Cheng Jiayang, Yunzhi Yao, Wenyang Gao, Xuming Hu, Zehan Qi, et al. 2023. Survey on Factuality in Large Language Models: Knowledge, Retrieval and Domain-specificity. *arXiv preprint arXiv:2310.07521* (2023).
[94] Wei Wang, Ryan Kennedy, David Lazer, and Naren Ramakrishnan. 2016. Growing Pains for Global Monitoring of Societal Events. *Science* 353, 6307 (2016), 1502–1503.
[95] Wei Wang, Yue Ning, Huzefa Rangwala, and Naren Ramakrishnan. 2016. A Multiple Instance Learning Framework for Identifying Key Sentences and Detecting Events. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. 509–518.
[96] Xiuling Wang, Hao Chen, Zhoujun Li, and Zhonghua Zhao. 2018. Unrest News Amount Prediction with Context-aware Attention LSTM. In *Pacific Rim International Conference on Artificial Intelligence*. Springer, 369–377.
[97] Yaqian Wang, Liang Ge, Siyu Li, and Feng Chang. 2020. Deep Temporal Multi-graph Convolutional Network for Crime Prediction. In *International Conference on Conceptual Modeling*. Springer, 525–538.
[98] Yinsen Wang, Xin Zhang, Yan Pan, and Zexin Fu. 2023. Hierarchical Attention based Feature Learning for Interpretable Social Event Prediction. In *Proceedings of the 2023 Asia Conference on Computer Vision, Image Processing and Pattern Recognition*. 1–7.
[99] Michael D. Ward, Andreas Beger, Josh Cutler, Matthew Dickenson, Cassy Dorff, and Ben Radford. 2013. Comparing GDELT and ICEWS Event Data. *Analysis* 21, 1 (2013), 267–297.

[100] Nils B. Weidmann and Michael D. Ward. 2010. Predicting Conflict in Space and Time. *Journal of Conflict Resolution* 54, 6 (2010), 883–901.

[101] Gadi Wolfsfeld, Elad Segev, and Tamir Sheafer. 2013. Social Media and the Arab Spring: Politics Comes First. *The International Journal of Press/Politics* 18, 2 (2013), 115–137.

[102] Congyu Wu and Matthew S. Gerber. 2017. Forecasting Civil Unrest Using Social Media and Protest Participation Theory. *IEEE Transactions on Computational Social Systems* 5, 1 (2017), 82–94.

[103] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and Philip S. Yu. 2020. A Comprehensive Survey on Graph Neural Networks. *IEEE transactions on neural networks and learning systems* 32, 1 (2020), 4–24.

[104] Yiming Yang, Tom Pierce, and Jaime Carbonell. 1998. A Study of Retrospective and On-line Event Detection. In *Proceedings of the 21st annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. 28–36.

[105] Liang Yao, Chengsheng Mao, and Yuan Luo. 2019. Graph Convolutional Networks for Text Classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 7370–7377.

[106] Wang Yinsen, Zhang Xin, Pan Yan, and Fu Zexin. 2023. A Temporal Attention-based Model for Social Event Prediction. In *2023 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 1–8.

[107] James E Yonamine. 2013. Predicting Future Levels of Violence in Afghanistan Districts using GDELT. *Unpublished manuscript* (2013).

[108] Chuxu Zhang, Dongjin Song, Chao Huang, Ananthram Swami, and Nitesh V Chawla. 2019. Heterogeneous Graph Neural Network. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*. 793–803.

[109] Libo Zhang and Yue Ning. 2024. Large Language Models as Event Forecasters. *arXiv preprint arXiv:2406.10492* (2024). arXiv:2406.10492

[110] Liang Zhao. 2021. Event Prediction in the Big Data Era: A Systematic Survey. *ACM Computing Surveys (CSUR)* 54, 5 (2021), 1–37.

[111] Liang Zhao, Feng Chen, Chang-Tien Lu, and Naren Ramakrishnan. 2015. Spatiotemporal Event Forecasting in Social Media. In *Proceedings of the 2015 SIAM International Conference on Data Mining*. SIAM, 963–971.

[112] Liang Zhao, Feng Chen, Chang-Tien Lu, and Naren Ramakrishnan. 2016. Multi-resolution Spatial Event Forecasting in Social Media. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*. IEEE, 689–698.

[113] Liang Zhao, Qian Sun, Jieping Ye, Feng Chen, Chang-Tien Lu, and Naren Ramakrishnan. 2015. Multi-task Learning for Spatio-temporal Event Forecasting. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 1503–1512.

[114] Liang Zhao, Qian Sun, Jieping Ye, Feng Chen, Chang-Tien Lu, and Naren Ramakrishnan. 2017. Feature Constrained Multi-task Learning Models for Spatiotemporal Event Forecasting. *IEEE Transactions on Knowledge and Data Engineering* 29, 5 (2017), 1059–1072.

[115] Liang Zhao, Junxiang Wang, and Xiaojie Guo. 2018. Distant-supervision of Heterogeneous Multitask Learning for Social Event Forecasting with Multilingual Indicators. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32.

[116] Liang Zhao, Jieping Ye, Feng Chen, Chang-Tien Lu, and Naren Ramakrishnan. 2016. Hierarchical Incomplete Multi-source Feature Learning for Spatiotemporal Event Forecasting. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2085–2094.

[117] Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, et al. 2023. A Survey of Large Language Models. *arXiv preprint arXiv:2303.18223* (2023).