



Communication Disaster Prediction Knowledge Graph Based on Multi-Source Spatio-Temporal Information

Xingtong Ge^{1,2}, Yi Yang³, Jiahui Chen^{1,2}, Weichao Li¹, Zhisheng Huang^{4,5,6}, Wenyue Zhang^{1,2} and Ling Peng^{1,*}

- ¹ Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100049, China; gexingtong21@mails.ucas.ac.cn (X.G.); chenjiahui21@mails.ucas.ac.cn (J.C.); liweichao@rsai.tech (W.L.); Zhangwenyue@rsai.tech (W.Z.)
- ² College of Resources and Environment (CRE), University of Chinese Academy of Sciences, Beijing 100049, China
- ³ Institute of Automation, Chinese Academy of Sciences, Beijing 100049, China; yangyi@ia.ac.cn
- ⁴ Knowledge Representation and Reasoning (KR&R) Group, Vrije Universiteit Amsterdam, 1081 HD Amsterdam, The Netherlands; z.huang@vu.nl
- ⁵ School of Computer Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, China
- ⁵ Ztone International BV, 1448 VC Purmerend, The Netherlands
- * Correspondence: pengling@aircas.ac.cn

Abstract: Natural disasters have frequently occurred and caused great harm. Although the remote sensing technology can effectively provide disaster data, it still needs to consider the relevant information from multiple aspects for disaster analysis. It is hard to build an analysis model that can integrate the remote sensing and the large-scale relevant information, particularly at the sematic level. This paper proposes a disaster prediction knowledge graph for disaster prediction by integrating remote sensing information, relevant geographic information, with the expert knowledge in the field of disaster analysis. This paper constructs the conceptual layer and instance layer of the knowledge graph by building a common semantic ontology of disasters and a unified spatio-temporal framework benchmark. Moreover, this paper represents the disaster prediction model in the forms of knowledge of disaster prediction. This paper demonstrates experiments and cases studies regarding the forest fire and geological landslide risk. These investigations show that the proposed method is beneficial to multi-source spatio-temporal information integration and disaster prediction.

Keywords: disaster prediction knowledge graph; spatio-temporal; disaster dynamic prediction; multi-source data fusion; forest fire risk prediction; geological landslide risk prediction

1. Introduction

Disaster is a general term for events that can have a destructive impact on humans and the environments on which they depend. Although natural disasters are inevitable, it must take efforts to greatly reduce losses. However, due to the wide distribution of natural disasters, time, location, and scale of natural disasters have great uncertainty. It greatly increases the difficulty to resist natural disasters. Therefore, the critical issue is the prediction of disasters with multi-source data of disaster scenarios for formulating emergency plans.

Remote sensing technology can obtain a large amount of strongly dynamic information in large observation ranges with high speed in real-time [1]. It can identify the fine features of ground objects, and extract information such as buildings, roads, and cultivated land by combining remote sensing technology and artificial intelligence technology. It plays an important role in the fields of disaster monitoring, early warning, and emergency decision-making. The critical problem is the fusion of different data types, different data structures, and different types of expert knowledge when the remote sensing information extraction for the multi-scale, multi-temporal, multi-type, multi-precision remote sensing



Citation: Ge, X.; Yang, Y.; Chen, J.; Li, W.; Huang, Z.; Zhang, W.; Peng, L. Disaster Prediction Knowledge Graph Based on Multi-Source Spatio-Temporal Information. *Remote Sens.* 2022, *14*, 1214. https://doi.org/10.3390/rs14051214

Academic Editors: Hideomi Gokon, Yudai Honma and Shunichi Koshimura

Received: 31 January 2022 Accepted: 26 February 2022 Published: 1 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). extraction information, other geographic information, and other spatio-temporal data in disaster scenes. There are studies for addressing the problem of spatio-temporal disaster information fusion, including knowledge representation of spatio-temporal disaster data, attribute analysis based on data spatio-temporal relations, and spatio-temporal dynamic geographic information of and expert knowledge.

In the real-world disaster scene, there exists large-scale multi-source heterogeneous data. There are complex associations between the data, as well as between the data and the disaster knowledge. It can effectively represent and model the data as well as the in-between relations using a knowledge graph. This is an artificial intelligence technology proposed by Google in 2012. Knowledge graph is a data modeling method that represents knowledge as concept, entity, and semantic relations with each other in the form of a graph [2]. Knowledge graphs can connect multi-source spatio-temporal disaster data and expert knowledge by a graph in order to support the analysis of the multi-source heterogeneous data under dynamic and static situations in natural disaster scenarios.

The spatio-temporal knowledge graph is a knowledge graph that can realize the semantic association of spatio-temporal data, which is the basis of disaster prediction. Sun et al. proposed a spatio-temporal knowledge graph ontology model which can manage and query spatio-temporal data in specific fields by adding spatio-temporal attribute information to entities [3]. Geographic information is a type of spatio-temporal data with spatial and temporal characteristics. Therefore, the construction technology of a geographic knowledge graph is very meaningful for this study. Zhang et al. proposed a method for constructing a geographic knowledge graph that takes the spatio-temporal characteristics into account. They studied the formal representation of geographic knowledge and laid a foundation for the acquisition, fusion, reasoning, and application of the spatio-temporal data [4]. Based on ontology, Liu et al. proposed a fast retrieval method that takes the semantic knowledge of geospatial data into account for reducing users' dependence on data storage methods and database grammar rules [5]. Jiang et al. proposed the construction process of a geographic knowledge graph, which helps to realize the knowledge of geographic information by reviewing the geographic knowledge graph [6].

At present, there exists research work regarding the construction of ontology and knowledge graph for specific disasters. Qiu et al. constructed model ontology and data ontology for flood management based on the four phases of disaster management [7]. Xu et al. constructed a conceptual model for earthquake disasters based on geo-ontology and relevant rules of earthquake emergency decision-making [8]. Scheuer et al. constructed flood risk assessment ontology by integrating SWEET ontology and MONITOR ontology [9]. Tao et al. summarized the knowledge graph construction process and key methods for the integrated comprehensive disaster reduction, and took the Jiuzhaigou earthquake as an example to demonstrate the knowledge graph construction process and construction results for Jiuzhaigou earthquake disaster reduction [10]. However, the above studies only focused on the knowledge model without considering the unified knowledge graph for different disaster types.

Furthermore, the existing knowledge graph models usually do not focus on strategies and rules based on common semantics so that it is hard to provide specific disasterhandling strategies in disaster emergency scenarios for supporting decision-making. The disaster prediction knowledge graph developed by Qiu et al. [7] visualizes the disaster situation. It can more autonomously determine the environmental models and disasterrelated data used in different disaster phases using the semantic constraints defined in the ontology [11]. Li et al. constructed a knowledge graph for flow disasters with the personalized visualization of different scenes for multilevel users [12]. Xie et al. combined the complex semantic relationship between earthquake prevention and control entities by extracting various earthquake disaster information, and established a semantic network by constructing a knowledge graph of earthquake disaster prevention and mitigation as the core [13]. Du et al. proposed a top-down and bottom-up method for constructing a natural disaster emergency knowledge graph [14]. They realized the transformation from multi-source data to interconnected knowledge by taking the flood disaster emergency knowledge graph as an example for experimental verification. At present, the most related research on the disaster knowledge graph respectively works on the perspectives of disaster ontology, disaster knowledge, and disaster events, particularly paying attention to construction methods. In applications, there are only domain knowledge-driven analyses. This rarely integrates dynamic disaster data and model methods deeply for the analysis based on the fused multi-feature fusion. It leads to the problem that the emergency plan cannot be flexibly adapted to the actual situations.

Affected by the attribute characteristics of surface elements, the development of disasters will show certain patterns when the disasters spatially migrate and change over time. This study selects forest fires and geological landslides as typical research objects. This paper proposes the Disaster Prediction Knowledge Graph (DPKG) that facilitates multi-source spatio-temporal data management and knowledge association. The proposed method supports semantic association of spatio-temporal data, efficient query, and quantitative analysis in order to implement information analysis with multiple types of disasters. The proposed method can be used for intelligent deduction of disaster situation and intelligent identification of disaster risk. This paper demonstrates experiments for verifying the benefits of the proposed method with respect to forest fire prediction and geological landslide risk prediction.

The main contributions of this paper are as follows: (1) This paper proposes a disaster prediction method which realizes dynamic data-driven disaster prediction based on the DPKG. (2) This paper proposes a fusion method for the dynamically changing spatiotemporal data and knowledge model. The proposed method is beneficial to improve the ability of natural disaster monitoring, early warning, and emergency response. It can provide a quantitative reference for disaster prediction and play a practical role to a certain extent.

2. Materials and Methods

2.1. DPKG Architecture

This section describes the theoretical way for mapping disaster prediction scenarios to knowledge graphs, mainly including the hierarchical semantic model of disaster prediction, and the mapping relationship between disaster prediction scenarios and DPKG architecture.

2.1.1. Hierarchical Semantic Model of Disaster Prediction Scenarios

The occurrence of disasters is related to a variety of factors. It facilities disaster prediction and decision-making by the analysis of disaster-inducing factors and their interrelationships.

The affect factors of occurrence of natural disasters are composed of common factors and characteristic factors. The common factors are closely related to different types of disasters. The characteristic factors are closely related to the occurrence of a certain natural disaster and weakly related to other types of natural disasters. This study establishes a hierarchical semantic model of disaster prediction considering the affect factors, which provided a theoretical basis for the construction of a conceptual layer for forest fire risk prediction and geological landslide risk prediction, as shown in Figure 1. Aiming at the prediction of forest fires and geological landslides, this paper extracts the common factors and characteristic factors of disaster occurrence by considering remote sensing information (vegetation, farmland, roads, buildings, and water bodies), terrain, meteorology, and human factors closely related to the occurrence of disasters.



Figure 1. A hierarchical semantic model for disaster prediction scenarios regarding forest fires and geological landslides.

2.1.2. Mapping between Disaster Prediction Scenarios and DPKG Architecture

For modeling disaster prediction using a knowledge graph, it is necessary to represent the disaster prediction scenario as the hierarchical semantic model. It maps the data with spatio-temporal characteristics in the disaster prediction scenario into the DPKG according to the structure of the semantic model. The disaster analysis model can be transformed into knowledge. The DPKG also needs to integrate the transformed disaster analysis model and expert knowledge for disaster prediction.

In addition, the architecture design of the DPKG for disaster prediction needs to consider the fusion of multi-source heterogeneous spatio-temporal data including remote sensing extraction information, terrain, meteorology, vegetation, and human factors, as well as the fusion of spatio-temporal data with disaster analysis models and expert knowledge for linkage analysis. This study designs an architecture of DPKG for disaster prediction, as shown in Figure 2, in order to support the modeling of disaster prediction and the collaborative analysis with fused data.



Figure 2. DPKG architecture.

The DPKG for disaster prediction consists of two parts, namely the rule set (TBox) and the fact set (ABox). The TBox consists of the conceptual layer and the inference rules. The conceptual layer is the common semantic basis for describing the hierarchical relationship of various factors in disaster prediction scenarios. The inference rules are the common logic basis for multi-source spatio-temporal data that supports semantic reasoning. The ABox constitutes the instance layer of the knowledge graph, which contains the time, space, status, and other attribute information of various types of ground objects related to disaster emergency extracted from multi-source heterogeneous data. The ABox and the TBox constitute the reasoning basis of the DPKG. Through the rapid retrieval of spatio-temporal data in disaster areas, multi-source spatio-temporal data analysis can be realized. The architecture diagram will be described in detail later.

2.2. Construction of Disaster Prediction Knowledge Graph

This section introduces the knowledge representation language, the design of the conceptual layer and instance layer, and the knowledge representation of the disaster prediction model.

2.2.1. Knowledge Representation Language

Knowledge graph representation refers to the use of computer symbols to scientifically mark the objectively existing knowledge for facilitating the semantic reasoning [15].

OWL (Web Ontology Language) is a semantic description standard including three sub-languages: OWL-Lite, OWL-DL, and OWL-Full. At present, OWL is the most standardized, rigorous, and expressive language for a knowledge graph. The commonly used knowledge representation form is RDF (Resource description framework) triples. Each piece of knowledge can be represented as the following triple format: (subject, predicate, object), such as (vegetation, yes, combustible). This study chooses OWL as the knowledge representation language.

2.2.2. Design of Conceptual Layer

The conceptual layer of the DPKG is a logical structure for multi-source spatiotemporal data. It contains semantic concepts and their interrelationships. Based on the semantic associations between different concepts, such as the subordinate relationship, attribute subject-object relationship, a tree-like hierarchical concept triple interconnection network is constructed for ensuring the consistence of the intrinsic semantic concepts of multi-source spatio-temporal data. The conceptual layer of disaster prediction DPKG is composed of time ontology, space ontology, and disaster prediction common semantic ontology.

1. Common Semantic Ontology of Disaster Prediction

The dynamic prediction of forest fires and geological landslide disasters needs the semantic reasoning and collaborative computing of multi-source spatio-temporal data.

For this goal, this paper proposes a tree-like taxonomy of multi-source concepts of geographic entity associated with multi-source spatio-temporal data. According to the taxonomy, concepts of geographic entity are divided into five fields. This paper defines the domain attribute predicates related to the concepts of each geographic entity.

The concepts of geographic entity constitute a common semantic ontology for the prediction of forest fire and geological landslide disaster.

The storage model of common semantic ontology is hierarchical and expandable. It builds the foundation for the predicate logic reasoning based on the common hierarchical semantic relationship of disaster prediction, as shown in Figure 3.



Figure 3. Common semantic ontology of disaster prediction.

2. Time Ontology

Time ontology provides a specification of unified time semantic representation in order to ensure that the time information of entities is comparable and computable. This paper leverages the Semantic Web Rule Language (SWRL) time ontology proposed by Stanford University for representing the common time concepts of the DPKG [16]. The logical structure of the SWRL time ontology is shown in Figure 4.



Figure 4. SWRL temporal ontology.

3. Space Ontology

The expression of spatial ontology applies A Geographic Query Language for RDF Data (GeoSPARQL) [17] that is a geographic semantic query specification proposed by OGC (Open Geospatial Consortium) [18]. The geographic semantic query mainly includes the following three common contents.

OWL Ontology Vocabulary

It represents all feature entities as subclasses of Spatial Objects. All geometric objects can be subdivided into points, linestrings, and polygons. Geometric objects are represented using features and geometric objects (Geometry) and their defined relationships, as shown in Figure 5.



Figure 5. GeoSPARQL part of the ontology table.

Geometry extension expression

The GeoSPARQL specification follows the OGC standard for the expression format of point, line, and surface. For example, the coordinate information of a surface data will be recorded in the attribute of *ogc:asWKT*, where *WKT* stands for Well Known Text, which is the ASCII code representation method of spatial objects.

Topology extension query

GeoSPARQL specifies eight query relationships for all spatial objects, as shown in Table 1. In addition, it provides many commonly used distance query functions (*geof:distance*), buffer query functions (*geof:buffer*), and convex Package constructor (*geof:convexHull*), to enhance spatial query.

Table 1. GeoSPARQL spatial topological relationship.

geof:sfEquals	geof:sfDisjoint	geof:sfIntersects	geof:sfTouches
geof:sfCrosses	geof:sfWithin	geof:sfContains	geof:sfOverlaps

2.2.3. Construction of Disaster Prediction Inference Rules

1. Construction of First-Order Logical Inference rules

First-order logic (FOL) is a formal system used in mathematics, philosophy, linguistics, and computer science. Specifically, hierarchical relationships and object attribute relationships can be described as first-order predicate logic, as shown in Figure 6. The common ontology, time ontology, and space ontology of disaster prediction are formally modeled with OWL language. The hierarchical relationship and attribute relationship implies a series of first-order logical reasoning.

Figure 6. Hierarchical relationship and object attribute relationship.

2. Construction of Production Inference rule

Due to the limitation of the knowledge representation of first-order logic, this paper applies SWRL (Semantic Web Rule Language). It extends the set of OWL axioms to include Horn-like rules. In this study, SWRL is used to further enhance the expression of rules for DPKG. A rule has the form, where both antecedent and subsequent are conjunctions of atoms written. Using this syntax, a rule asserting that the combination of the tree species' properties and the tree species' forest fire risk-level properties implies the tree's forest fire risk-level properties, as shown in Figure 7.

```
<species(?x,?y) ∧ forestFireRiskLevel(?y,?z) ⇒ forestFireRiskLevel(?x,?z)>
```

Figure 7. SWRL inference rule example.

3. Construction of Spatio-Temporal Semantic Inference Rule

Although SWRL can support basic quantitative calculations, it cannot support quantitative analysis of space and time including spatio-temporal semantics. This study leverages the construction method of production inference rules to build the spatio-temporal semantic inference rules. The spatio-temporal semantic inference rule *RuleObject* is a set of rules for the automatic execution of reasoning programs by the DPKG. Each rule is composed of an event object *TriggerObject* and an action object *ActionObject*, which are expressed as RuleObject = (Tr, Ac), where *Tr* represents the event contained in the *RuleObject*, *Ac* represents the action object *ActionObject* contained in the *RuleObject*, and *R* is the reasoning result. The event *TriggerObject* in this article is defined as a triple that is represented as TriggerObject = (O, T, S), where *O* represents the geographic entity set contained in the event object, *T* and *S* respectively represent the intersection of the geographic entity set in the time dimension and the space dimension. A spatio-temporal co-occurrence scene with a set of geographic entities can be described as an event object that is a definition of the applicable conditions of an inference rule.

The event (or action) object concept is further subdivided into independent events (or actions) and event (or action) combinations, which together form the concept of knowledge inference rules. The logical structure is shown in Figure 8.



Figure 8. Schematic diagram of the conceptual hierarchy of knowledge inference rules.

- 2.2.4. Design of Instance Layer
- 1. Knowledge Extraction from Unstructured Data

The dynamic disaster prediction of forest fires and geological landslides requires high spatio-temporal resolution of land cover data. This paper uses Gaofen-2 satellite remote sensing images with the resolution of 0.8 m as the data source. The spatial distribution of disaster-bearing bodies such as buildings and roads is extracted by deep learning methods. The spatial distribution of surface vegetation is obtained by NDVI numerical calculation method [1,19]. It converts the spatial information into triples according to the specification of spatio-temporal and professional attribute representation defined by the concept layer. The triples are stored in the graph database *GraphDB*.

2. Knowledge Extraction from Semi-Structured Data

The terrain data for dynamic disaster prediction of forest fires and geological landslide is the raster geographic data with GeoTIFF format. It converts all types of raster data into vector data of surface elements. It converts geological and lithological data including stratigraphic age, fault, and lithology distribution data into vector geographic information with SHP format. For all types of vector geographic information, it converts the spatial information and feature attributes into triple that are stored in *GraphDB*.

3. Knowledge Extraction from Structured Data

The meteorological data for the dynamic disaster prediction of forest fires and geological landslide is multi-field structured data, which has a direct mapping relationship with the spatio-temporal and professional attributes. Since the vast majority of meteorological data is normal non-hazardous data, this paper only converts the spatial information element attributes of potential disaster-causing meteorological indicators into a triple, so as to avoid the low reasoning speed caused by a large amount of irrelevant data to the reasoning. The triple is stored in *GraphDB*.

4. Knowledge Extraction from Disaster Prediction Reasoning Criterion

This paper introduces the knowledge extraction method by taking the geological landslide prediction model as an example. It is hard to build a general prediction model for a geological landslide because a single model with idiosyncratic conditions cannot deal with the rich types of geological landslide prediction models and the involved data. This paper defines the inference rules for the geological landslide disaster prediction according to the geological landslide risk probability calculation using machine learning and the effective precipitation models. Based on the general law of geological landslide risk analysis [20], this paper selects the H-index shown in Formula (1) as the metric for the quantitative evaluation of the geological landslide risk:

$$\mathbf{H} = P_{(T)} \times P_{(S)} \times P_{(I)},\tag{1}$$

where H is the risk probability of geological landslide disasters within a specific time and space range; $P_{(T)}$ is the time probability, i.e., the probability of geological landslide disasters occurring within a specific time range; $P_{(s)}$ is spatial probability, i.e., the probability of geological landslide disasters occurring within a specific spatial range; $P_{(I)}$ is the intensity probability, i.e., the intensity of possible geological landslide disasters.

This paper introduces a time probability model based on dynamic precipitation data. The model is mainly composed of two parts. One is the effective precipitation; another is the fitting relationship between the effective precipitation and the frequency of geological landslides. The effective precipitation *R* in the region considers the respective precipitation contributions of the previous days to the cumulative precipitation that causes the geological landslide. The later the time, the greater the contribution rate to the geological landslide.

$$R = R_0 + \alpha_1 R_1 + \ldots + \alpha_n R_n, \tag{2}$$

where R_0 is the daily precipitation; α is the contribution rate; n is the number of precipitation days, and R_n is the precipitation on the n-th day before. A linear fitting relationship between the effective precipitation and the frequency of geological landslides is defined for the time probability of geological landslides associated with dynamic precipitation data.

The geological landslide strength is considered as a parameter representing the destructive force of the geological landslide. It includes the volume of the geological landslide, and the sliding velocity. At present, there is no unified set of indicators to describe them. In this paper, the geological landslide volume parameter is used to obtain the geological landslide strength, and the formula is as follows:

$$n_L = \log V_L, \tag{3}$$

where V_L is the volume of a single geological landslide with the unit of m³; which is calculated based on DEM, and m_L represents the strength of a single geological landslide. By calculating the strength of each geological landslide in the study area by the above formula, it can obtain the statistics of geological landslide strength according to the frequency.

1

There are many common factors that lead to forest fires and geological landslides. Certain triggering factors can be monitored when forest fires and geological landslides occur. The development of the two types of disasters will develop and spread within a certain time and space range. Affected by the attribute characteristics of surface elements, both disasters will show certain change patterns when they dynamically migrate and change in space over time. Therefore, this paper extends the geological landslide risk calculation formula to forest fires by considering the above-mentioned commonalities of the two disasters. It builds the unified prediction model for forest fire according to Formula (1). Based on Analytic Hierarchy Process, it builds the forest fire-driven comprehensive index for forest fire risk prediction model.

2.3. *Query and Reasoning of Disaster Prediction DPKG*

2.3.1. Spatio-Temporal Semantic Query for Disaster Fusion Data

When a specific attribute of the external dynamic data satisfies the numerical conditions defined by the rule, it is regarded as an abnormal event candidate. Its time range and space range are treated as the basic condition of semantic query. It queries with SPARQL Protocol and RDF Query Language (SPARQL) [21] statements for a geographic entity that has an intersection with the abnormal event candidate and satisfies the reasoning conditions. When querying the spatio-temporal relationship in the graph database, it can quickly locate the affected area of the abnormal event candidate because the graph database is good at efficient depth-first query, as shown in Figure 9. It plays an important role in the emergency field with high response speed requirements.

Disaster prediction reasoning flow chart



Figure 9. Disaster prediction query flow chart.

With the help of the spatio-temporal semantic query of the disaster fusion data, it can quickly find geographic entities located within the temporal and spatial scope of the AbnormalTrigger based on the temporal-spatial intersection relationship in order to filter the buildings and roads affected by the AbnormalTrigger, as shown in Figure 10.



Figure 10. Schematic diagram of the spatio-temporal intersection of anomalous events and geographic entities.

In order to ensure the efficiency of spatial information query in disaster emergency scenarios, this paper builds multi-scale geocoding indexes for multi-source geographic entities in GeoSPARQL. It significantly improves the depth-first query speed of random geospatial information in comparison with the frequent joint query of multiple tables in the relational database.

- 2.3.2. Disaster Prediction Rule Reasoning
- 1. First-order logical reasoning

As defined in the Common Semantic Ontology of Disaster Prediction, surface combustibles include Woodland, Farmland, Building, and Grassland, and Coniferous Forest belongs to Woodland. Based on first-order logical reasoning, it concludes that Coniferous Forest belongs to Surface Combustibles. With the prefix of *hazard*, the above process can be formally defined as shown in Figure 11. Based on these expressions, *(hazard : Coniferous Forest rdf : subClassOf hazard : Surface Combustibles)* can be inferred.

- 1 <hazard:Woodland rdf:subClassOf hazard:Surface Combustibles>
- 2 <hazard:Farmland rdf:subClassOf hazard:Surface Combustibles>
- 3 <hazard:Building rdf:subClassOf hazard:Surface Combustibles>
- 4 <hazard:Grassland rdf:subClassOf hazard:Surface Combustibles>
- 5 <hazard:Coniferous Forest rdf:subClassOf hazard:Woodland>

Figure 11. First-order logical reasoning process.

2. Production reasoning

This paper takes SWRL rule reasoning as an example. The feasible reasoning on the DPKG includes reasoning of geographic entity from one attribute to other attributes; reasoning from one geographic entity and its attributes to geographic entities and their attributes.

For example, given the forest fire risk entity and its attribute, i.e., the comprehensive forest fire-driven index *fireDangerValue*, the goal is to build an inference rule for the forest fire risk-level prediction. For this task, the domain analysis rule "when fireDangerValue > 3.8, fireDangerLevel is level 3" is formalized as shown as line 1 in Figure 12.

- 1 <hazard:ForestFireGrid(?t) ∧hazard:fireDangerValue(?t,?frl) ∧swrlb:greaterThan(?frl,"3.8" ∧ ∧xsd:double) →hazard:fireDangerLevel(?t,"3")>
- 2 <hazard:aspect(?t) ∧ swrlb:greaterThanOrEqual(?a, "180" ∧ ∧xsd:int) ∧ swrlb:subtract(?u, ?a, "180" ∧ ∧xsd:int) ∧hazard:aspect(?t, ?a)→hazard:upAspect(?t, ?u)>

Figure 12. Production reasoning process.

One more example, knowing the value of aspect entity and aspect attribute, the goal is to construct an inference rule to infer the upslope aspect value. For this task, the domain analysis rule "When aspect is $<180^\circ$, upAspect = aspect + 180° " is formalized as shown as line 2 in Figure 12.

The prediction can be carried out using the SWRL reasoning engine with the inference rules.

3. Spatio-temporal semantic reasoning

This paper builds an automatic real-time monitoring mechanism for emergency information for disaster prediction. It extracts the dynamic information related to the occurrence and development of disasters, i.e., attributes such as space, time, and status of target objects, from structured, semi-structured, and unstructured data sources. It encapsulates the extracted information into real-time event message objects with GeoJSON format. The update of dynamic data will trigger the automatic judgment by the DPKG as shown in Figure 13. The results drive the chained reasoning workflow defined by the disaster prediction inference rules.

Calculation process



Figure 13. Workflow of disaster dynamic prediction.

This paper designs a chain-type automatic reasoning method based on the reasoning criterion. The method divides the calculation logic of the geological landslide prediction model into a series of antecedents and consequences. The antecedents are execution conditions that contain a series of geographic entities. When there exists an intersection of the geographic entities and the conditions set in the antecedent are met, the subsequent associated with the antecedent, i.e., the action object, will be triggered.

It establishes the automatic flow logic between independent criteria such as time probability, space probability, intensity probability, vulnerability of hazard-affected body, and value amount by defining for RuleObject. In this way, it can analyze the relations between time information and space information, which is not feasible using the production reasoning with SWRL.

3. Results

This section demonstrates experiments and cases studies of the forest fire and geological landslide risk based on DPKG. These investigations show that the proposed method is beneficial to multi-source spatio-temporal information integration and disaster predictions.

3.1. Case Study: Forest Fire Risk Prediction

In this case study, Yanyuan County, in the Liangshan Yi Autonomous Prefecture of China's Sichuan Province, was selected to carry out a dynamic prediction experiment of forest fire disasters. Yanyuan County is located between 100°42′09″–102°03′44″ east longitude and 27°06′31–28°16′31″ north latitude, with a total area of 8398.6 square kilometers. There are rich vegetation and many miscellaneous irrigation and thatch grass in the area. The forest fire disasters are easily caused by the inducing factors such as low rainfall. Therefore, Yanyuan County strongly needs to forecast the situation in advance.

In this case, multi-source heterogeneous data related to forest fire risk prediction are collected, such as vegetation data, terrain data, meteorological data, and land cover data. In order to fully represent the spatio-temporal features of the above data in the DPKG knowledge graph, this paper uses Protégé to construct a time ontology and a space ontology. The concepts related to forest fire prediction form the conceptual layer of DPKG. For the above-mentioned multi-source heterogeneous data, we built a diversified knowledge extraction method for transferring the data into triples according to the semantics of the conceptual layer.

The triples form an instance layer of the DPKG of disaster prediction. An example of the extraction process of the slope in the terrain data is shown in Figure 14.



Figure 14. Knowledge extraction process using slope as an example.

Forest fire risk prediction needs the theoretical support because the above spatiotemporal data is only a prerequisite for forest fire risk prediction. In this case, the AHP (Analytic Hierarchy Process) is used to evaluate forest fire risk indicators. This paper builds an indicator system for four types of forest fire risk factors: meteorology, terrain, vegetation, and man-made. The weights are shown in Figure 15.



Figure 15. Forest fire risk factor indicator system.

The comprehensive indicator of forest fire driving factors is defined as Formula (4):

$$CIFFD = \sum v_i \times w_i \ i \in [1, \ 10], \tag{4}$$

where *CIFFD* represents a composite index of forest fire drivers; v_i indicates the level corresponding to the original value of the factor, and w_i indicates the weight corresponding to the factor. The larger the CIFFD, the greater the risk of forest fires.

In this case, 10 factors in four categories of terrain, vegetation, meteorology, and human factors in the experimental area are used as the query input for the DPKG instance layer. The definition in Formula (4) is transformed into spatio-temporal semantic reasoning rules. The reference output is a comprehensive index of forest fire driving factors.

Define spatio-temporal semantic inference rule 1: If the update of meteorological data is detected, it divides the spatial range of the meteorological data update into several rectangular areas. The rectangular areas automatically triggers the query of 10 factors of

terrain, vegetation, meteorology, and human factors in the rectangular areas for obtaining the original value of the factor.

Define spatio-temporal semantic inference rule 2: It obtains the corresponding level value based on the original value of the factor. Based on Formula (4), the comprehensive index of forest fire driving factors is obtained.

Define spatio-temporal semantic inference rule 3: Based on the natural discontinuity method, the calculated comprehensive index of forest fire driving factors is divided into 7 intervals for representing different degrees from low to high. The forest fire risk warning information of the experimental area is obtained, as shown in Figure 16.



Figure 16. Forest fire risk factor index system.

The antecedent rules in the above spatio-temporal semantic inference rules can trigger the subsequent rules when satisfying the conditions. Therefore, the dynamic changes of meteorological data can automatically trigger the forest fire risk calculation. It realizes the automatic completion of the chain reasoning of the whole process of DPKG without manual effort.

In this case, the actual meteorological record data of Yanyuan County Meteorological Observatory on 3 January 2022 is used as the query input of the instance layer of the DPKG. The output is calculated based on the reasoning of the DPKG.

For the forest fire risk index of the whole county of Yanyuan, the larger the index value, the higher the forest fire risk. The experimental results are shown in Figure 17.



Figure 17. Map of landslide hazards in Yanyuan County based on multi-Graded Cascade Random Forest.

NASA's Fire Information for Resource Management System shows the actual fire point in Yanyuan County at 6:35 on 3 January 2022, as shown in Figure 17. The results show that the proposed method successfully hits the fire point of the forest fire species within the time and space range of the predicted forest fire risk index greater than 4.21.

The case uses the traditional spatial analysis tool ArcGIS as a baseline method to calculate the forest fire risk in Yanyuan County for the comparative experiments. It is necessary to load and query Yanyuan County data from 14 types of raster datasets. Steps such as uniform coordinate system, uniform pixel size, grid cropping, and stitching are performed on each raster data, and then the prediction calculation can be performed. There are many steps in the baseline operation, and the whole process takes more than 190 min. When using the grid calculator, the performance is low because it needs complex numerical computations in multiple steps.

In addition, the boundaries of various data are not completely coincident, which will lead to missing or abnormal results of the boundary area analysis, and wrong predictions. Although it is possible to forcibly align the pixel positions by means of translation, it will cause errors by an operation that lacks a realistic basis, and changes the spatial distribution of the original data. In addition, it has a low prediction efficiency when using the traditional spatial analysis tools because they cannot process other target regions in parallel.

3.2. Case Study: Geological Landslide Risk Prediction

In another case study, Xiji County in Guyuan City of China's Ningxia Hui Autonomous Region is selected to carry out an experiment regarding the dynamic prediction of geological landslide hazards. Xiji County locates between 105°20′–106°04′ east longitude and 35°35′–36°14′ north latitude, which has a total area of 3130 square kilometers. In this area, the loess landslide disaster caused by the Haiyuan earthquake is particularly serious because of the special geological environmental conditions. It has formed 765 geological landslides [22]. It is very dangerous that new geological landslide disasters are easily generated by the inducing factors, e.g., rainfall. Therefore, it is urgent to take an effective method to predict disaster in advance.

In this case, the basic geological data, basic geographic data, hydrological data, human activities, and land is collected for the prediction and calculation of geological landslide risk. In order to fully represent the spatio-temporal features of the above data in the knowledge graph, this paper uses Protégé [23] to construct a time ontology and a space ontology, and concepts related to geological landslide prediction form the conceptual layer of spatial-temporal knowledge graph for disaster prediction. For the above-mentioned multi-source heterogeneous data, a diversified knowledge extraction method is adopted to convert the data into triples according to the semantics of the conceptual layer. It forms the instance layer of the DPKG for disaster prediction using the triples.

The above spatio-temporal data is only a prerequisite for geological landslide risk prediction. It needs the theoretical support of the geological landslide risk prediction. The landslide risk assessment is defined as Formula (2). Geological landslide risk prediction involves time, spatial, and intensity probability. In this case, there are 13 factors belonging to four categories of topography, geology, lithology, meteorology and hydrology, land cover, and human activities in the experimental area. The information of the factors is queried from the knowledge graph. It takes the DEM factors in the topographic data as the query input of the instance layer of the knowledge graph. The linear fitting model between historical landslides and DEM data is transformed into spatio-temporal semantic reasoning rules, as shown in Figure 18. In this way, the spatial probability of landslide disasters in the experimental area can be calculated.



Figure 18. Geological landslide risk prediction calculation formula.

The time probability is a dynamic element. In this case, the time probability is defined as the empirical function with the independent variable of the local precipitation data in the past 10 days, as shown in Formula (3). It transforms Formula (3) into spatio-temporal semantic inference rules, and the time probability of landslide disaster occurrence in the experimental area.

Define spatio-temporal semantic inference rule 1: If the time probability satisfies specific conditions, it is defined as a dangerous time probability.

Define spatio-temporal semantic reasoning rule 2: For a new dangerous time probability, the calculation of the spatial probability and the intensity probability will be automatically triggered. The spatial distribution of the geological landslide risk value in the test area will be calculated according to Formula (1).

Define spatio-temporal semantic inference rule 3: Based on the natural discontinuity method, the geological landslide risk probability value calculated by the previous rule is divided into 10 intervals in ascending order, representing different degrees, every two intervals are combined into a grade for a total of five grades, as shown in Figure 19.



Figure 19. Space-time semantic reasoning logic with criteria (VL: very low, L: low, M: moderate, H: high, VH: very high).

The antecedent rules in the above spatio-temporal semantic inference rules can trigger the subsequent rules when the conditions are satisfied. Therefore, the dynamic change of precipitation data can automatically trigger the geological landslide risk calculation for realizing the automatic completion of the whole process of DPKG chain reasoning without manual effort.

In this case, the actual precipitation record data of Xiji County Meteorological Observatory on 18 September 2018 is used as the query input of the instance layer of the DPKG. Based on the reasoning of spatial-temporal knowledge graph, the geological landslide time probability was daily calculated for the whole county of Xiji from 8 to 28 September 2018. From 18 to 19 September 2018, the geological landslide time probability in the county met the above geological landslide risk time probability conditions for the two consecutive days. The experimental results are shown in Figure 20. Due to the setting of reasoning rule 2, the calculation of geological landslide risk will be triggered automatically. The calculation results of geological landslide risk probability on 18 September 2018 are shown in Figure 21.





Bar: The graph of the time probability of landslide in Xiji County from September 8 to 28, 2018 Broken line: The graph of dangerous time probability of landslide in Xiji Countyfrom September 8 to 28, 2018

Figure 20. The graph of the time probability of landslide. The bar is the graph of the time probability of landslide in Xiji County from 8–28 September 2018. The broken line is the graph of dangerous time probability of landslide in Xiji County from 8–28 September 2018.



Figure 21. Map of landslide hazards in Xiji County based on multi-Graded Cascade Random Forest.

In this case, since the distribution data of buildings and roads in Xiji County is included in the DPKG, it further filtered the buildings and roads in the high-risk areas of geological landslides on 18 September 2018, as shown in Figure 22. From Internet news, we learned that a landslide occurred on National Highway 309 on 18 September 2018, which caused road blockage. The actual geological landslide section is located in the dangerous road area that has been correctly predicted by our method, as shown in Figure 22. Experiments show that it successfully hit the geological landslide event within the time and space range of the warning.



Figure 22. Road distribution map of Xiji County included in the landslide disaster warning area on 18 September 2018.

Prediction was carried out in Xiji Country respectively for the area of 3130.00 km², the area of 1509.54 km², and the area of 800.88 km². The experimental results are shown in Figure 23. With the increase of the experimental range, the time of experimental calculation also increases. The performance is basically in a satisfied performance range. Therefore, our method is applicable for large-scale disaster prediction.



Figure 23. Comparison of forecast time for different areas in Xiji Country.

4. Discussion

Compared with the traditional prediction methods, the disaster prediction method based on the DPKG has the advantages of more accurate spatial overlay analysis, high degree of automation, and fast calculation speed, as shown in Table 2. However, it cannot be ignored that the proposed method in this paper still has certain limitations: (1) Before forecasting the target area, the data required for disaster prediction needs to be preprocessed and stored in advance, and data storage requires a certain amount of time. The larger the area, the more data dimensions required for prediction, and the longer the time. (2) The data used in this case is not accurate enough. For example, the spatial resolution of terrain

data is 30 m, and the spatial resolution of surface coverage data for disaster prediction is 10 m. The accuracy still needs to be improved. If higher-resolution data is obtained, and more detailed tree species, tree age, and other data can be involved in disaster prediction, the accuracy of disaster prediction can be further improved. (3) The comparison with traditional disaster forecasting methods is not sufficient in the experiment that only considers the most well-known tool ArcGIS.

	Experimental Procedure	Prediction Precision	Calculation Time (Calculation Area Is 3130.00 km ²)
ArcGIS	 Load and query County data from types of raster datasets. Uniform coordinate system, uniform pixel size. Grid cropping and stitching are performed on each raster data. Use the Raster Calculator tool to do a weighted sum of all factors. Eliminate outliers and get disaster prediction results. 	 The boundaries of various data are not completely coincident, which will lead to missing or abnormal results of the boundary area analysis, and wrong predictions. Although it is possible to forcibly align the pixel positions by means of translation, it will cause errors by an operation that lacks a realistic basis and changes the spatial distribution of the original data. 	190 min
DPKG	 Multi-source spatio-temporal data preprocessing; input into the knowledge graph. Represent the prediction model as an inference rule and input it into the knowledge graph. Calculate and obtain forecast results. 	 More accurate spatial overlay analysis, when multi-source raster data is not spatially aligned; it will not cause calculation errors. When multi-source data with different resolutions have overlapping areas, select data with higher resolution for calculation. 	7.81 min

Table 2. Comparison of DPKG and ArcGIS.

5. Conclusions

Compared with plain text data and structured databases, DPKG has the advantages of richer semantic representation, more accurate data content, and better query performance. The knowledge graph provides a new means for the organization, management, fusion, and analysis of multi-source heterogeneous data in complex disaster environments. It facilitates prediction, prevention, and mitigation of the natural disasters. This paper proposes the DPKG that integrates multi-source remote sensing information in disaster scenes with other multi-source heterogeneous information. It considers not only dynamic data describing spatio-temporal facts but also domain analysis models. Based on the analysis of disaster influencing factors, this paper proposes a common disaster prediction method based on the DPKG for the dynamic data-driven disaster prediction. This paper introduces new ideas by taking forest fires and geological landslides as examples. The proposed method can serve the research and application of disaster emergency response with the strong help of artificial intelligence technology.

- From the perspective of cross-domain knowledge integration, knowledge graph integrates remote sensing knowledge and expert knowledge through semantic technology. It effectively connects multi-source heterogeneous data (GIS, meteorology, terrain, ground sensors, etc.) with expert knowledge in the field of disasters;
- (2) The DPKG contains dynamically updated spatio-temporal facts that reflect changes in the real world. Through knowledge graph query, the query performance in disaster emergency scenarios can be improved, and dynamic data updates can automatically drive prediction;
- (3) It provides efficient data storage and management methods for practitioners in the fields of remote sensing and geo-information, which helps to improve the efficiency

of spatio-temporal data query. It reduces the manual effort by using reasoning of the knowledge graph.

The DPKG proposed in this paper aims at the integration of remote sensing information, geographic information, and the correlation between dynamic data and static knowledge in complex disaster environments. It lays the foundation for future in-depth research, particularly for knowledge extraction and knowledge discovery for disaster prediction. In future work, indexing and query performance need to be further verified in the case of more types and larger amounts of data. The depth of knowledge reasoning needs to be further improved. In future research, we will enhance the DPKG by leveraging graph neural networks and deep learning models, so as to integrate a large number of spatio-temporal facts, and realize disaster emergency decision-making, historical data verification, and new model construction; it needs to build a knowledge service system for natural disaster monitoring and prediction as well as the emergency decision-making. In addition, this article basically focuses on verifying the availability and usability of our model through case studies. We will compare with more traditional prediction methods in the future.

Author Contributions: Conceptualization, X.G. and Y.Y.; methodology, X.G. and W.L.; validation, X.G., W.L. and W.Z.; resources, W.L.; data curation, X.G. and W.L.; writing—original draft preparation, X.G. and J.C.; writing—review and editing, X.G., Y.Y., L.P. and Z.H.; funding acquisition, L.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Beijing Municipal Science and Technology Project (Z191100001419002), and Ningxia Key R&D Program (2020BFG02013).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Chen, D.-Y.; Peng, L.; Li, W.-C.; Wang, Y.-D. Building Extraction and Number Statistics in WUI Areas Based on UNet Structureand Ensemble Learning. *Remote Sens.* 2021, 13, 1172. [CrossRef]
- 2. Lu, F.; Yu, L.; Qiu, P.Y. On geographic knowledge graph. J. Geo-Inf. Sci. 2017, 19, 723–734.
- 3. Sun, Y.; Yu, L.; Guo, Z.; Chen, G. Construction and Application of spatio-temporal Knowledge Graph. J. Inf. Eng. Univ. 2020, 21, 464–469.
- 4. Zhang, X.Y.; Zhang, C.J.; Wu, M.G.; Lv, C. Spatio-temporal features based geographical knowledge graph construction. *Sci. Sin. Inf.* **2020**, *50*, 1019–1032. (In Chinese)
- Liu, J.; Liu, H.; Chen, X.; Zhao, Q.; Liu, J.; Kang, L.; Guo, X. Rapid Retrieval of Geospatial Data Considering Semantic Knowledge. *Geomat. Inf. Sci. Wuhan Univ.* 2020, 1–13. Available online: https://kns.cnki.net/kcms/detail/42.1676.tn.20201209.0957.001.html (accessed on 25 February 2022).
- Jiang, B.; Wan, G.; Xu, J.; Feng, L.I.; Wen, H. Geographic Knowledge Graph Building Extracted from Multi-sourced Heterogeneous Data. Acta Geod. Et Cartogr. Sin. 2018, 47, 1051–1061.
- 7. Qiu, L.; Du, Z.; Zhu, Q.; Fan, Y. An integrated flood management system based on linking environmental models and disasterrelated data. *Environ. Model. Softw.* 2017, 91, 111–126. [CrossRef]
- Xu, J.; Nyerges, T.L.; Nie, G. Modeling and representation for earthquake emergency response knowledge: Perspective for working with geo-ontology. *Int. J. Geogr. Inf. Sci.* 2014, 28, 185–205. [CrossRef]
- 9. Scheuer, S.; Haase, D.; Meyer, V. Towards a flood risk assessment ontology—Knowledge integration into a multi-criteria risk assessment approach. *Comput. Environ. Urban Syst.* 2013, 37, 82–94. [CrossRef]
- 10. Tao, K.; Zhao, Y.; Zhu, P.; Zhu, Y.; Liu, S.; Zhao, X. Knowledge Graph Construction for Integrated Disaster Reduction. *Geomat. Inf. Sci. Wuhan Univ.* **2020**, *45*, 1296–1302.
- 11. Yang, W.; Xu, K.; Lian, J.; Ma, C.; Bin, L. Integrated flood vulnerability assessment approach based on TOPSIS and Shannon entropy methods. *Ecol. Indic.* 2018, *89*, 269–280. [CrossRef]
- 12. Li, W.; Zhu, J.; Zhang, Y.; Fu, L.; Gong, Y.; Hu, Y.; Cao, Y. An on-demand construction method of disaster scenes for multilevel users. *Nat. Hazards* **2020**, *101*, 409–428. [CrossRef]
- 13. Xie, Y.; Wang, L.; Dong, C.; Kang, F. Research on the construction method of earthquake disaster prevention knowledge graph. *Sci. Surv. Mapp.* **2021**, *46*, 219–226.

- 14. Du, Z.; Li, Y.; Zhang, Y.; Tan, Y.; Zhao, W. Knowledge Graph Construction Method on Natural Disaster Emergency. *Geomat. Inf. Sci. Wuhan Univ.* **2020**, *45*, 1344–1355.
- 15. Sowa, J. Principles of Semantic Networks: Explorations in the Representation of Knowledge; Morgan Kaufmann Publishers, Inc.: San Mateo, CA, USA, 1991.
- O'Connor, M.J.; Das, A.K. A method for representing and querying temporal information in owl. In Proceedings of the International Joint Conference on Biomedical Engineering Systems and Technologies, Valencia, Spain, 20–23 January 2010; pp. 97–110.
- 17. GeoSPARQL. Available online: http://www.geosparql.org/ (accessed on 25 February 2022).
- 18. Duan, W. Key Technologies of Geographicity Query. Master's Thesis, Wuhan University, Wuhan, China, 2013.
- 19. Chen, R.; Li, X.; Hu, Y. Road Extraction from Remote Sensing Images in Wildland-Urban Interface Areas. *Geosci. Remote Sens. Lett.* **2020**, *19*, 3000705. [CrossRef]
- 20. Corominas, J.; Copons, R.; Moya, J. Quantitative assessment of the residual risk in a rockfall protected area. *Landslides* **2005**, *2*, 343–357. [CrossRef]
- 21. SPARQL-Semantic Web Standards. Available online: https://www.w3.org/2001/sw/wiki/SPARQL (accessed on 25 February 2022).
- Shang, H.; Ni, W.K.; Liu, H.S. Development Characteristics and Distribution of Earthquake-induced Loess Landslides in Xiji County of Ningxia Hui Autonomous Region. *Bull. Soil Water Conserv.* 2012, 32, 28–31.
- Musen, M.A. Encyclopedia of Systems Biology; Dubitzky, W., Wolkenhauer, O., Cho, K.-H., Yokota, H., The Protégé Team Protégé Ontology Editor, Eds.; Springer: New York, NY, USA, 2013; pp. 1763–1765. Available online: https://link.springer.com/ referenceworkentry/10.1007%2F978-1-4419-9863-7_1104 (accessed on 25 February 2022).