

Ontology-Based Landslide Knowledge Representation and Hazard Assessment Research

Chenchen Luo, Shixiu Huang*, Han Zhou, Kang Zhou

School Of Civil And Hydraulic Engineering, Hefei University Of Technology, Hefei, China, 230009

* Corresponding Author Email: huanghfut2024@163.com

Abstract. The occurrence of landslides is characterized by suddenness and unpredictability, with many complex relationships among various influencing factors. This study, by introducing the ontology of knowledge engineering, summarizes the characteristics of landslide knowledge and constructs a multidisciplinary, multi-level, and multi-perspective ontology of the landslide domain. Through in-depth analysis of existing landslide research systems and methods, the study defines the basic concepts, attributes, and their complex relationships and rules related to landslides. This research expresses landslide knowledge using the Web Ontology Language (OWL), enabling effective storage and retrieval of landslide knowledge, thus promoting the systematization and standardization of knowledge in the landslide domain. Furthermore, leveraging the high knowledge density feature of the constructed landslide domain ontology, this study extends traditional landslide evaluation and prediction methods by introducing more evaluation indicators based on the constructed landslide ontology. The comprehensive evaluation of landslide risk is conducted using the contribution-weighting method, providing new insights and tools for landslide risk management and decision-making.

Keywords: Landslide knowledge representation, Landslide evaluation, Ontology.

1. Introduction

Landslides, as a common natural disaster, generally refer to the geological hazard of unstable soil or rock mass on a slope, which, influenced by natural conditions or human factors, moves downward along one or multiple sliding shear surfaces under the force of gravity. In regions with complex terrain and adverse geological conditions, landslide disasters occur frequently, posing a serious threat to human society and economic development^[1]. Therefore, monitoring and evaluating landslides have significant economic and social benefits^[2]. Introducing ontology technology from knowledge engineering to supplement and improve the information at the knowledge level of landslides is a new approach to studying landslides.

Ontology technology is a standardized way of representing knowledge^[3]. As early as the beginning of this century, ontologies have been applied in the fields of geoscience and disasters^[4], focusing on the integration of multi-source data such as geology, topography, and meteorology. This technology has gained wide attention in fields such as knowledge graphs, biomedicine, and information retrieval^[5]. Some scholars have established knowledge graphs of landslides from the perspectives of engineering geology^[6], spatiotemporal mechanisms in geoscience, etc. However, they lack a comprehensive interdisciplinary perspective, and there is also a lack of computational and application aspects of knowledge. Traditional methods for landslide evaluation and prediction rely on representative judgment criteria such as slope and elevation, with weight assignments through expert scoring or statistical analysis^[7]. Although these methods are easy to operate and have strong feasibility, they are overly reliant on experience, possess subjective biases, and do not fully consider the mechanistic knowledge of landslide occurrence. By constructing an applied ontology-based knowledge base in the landslide domain, evaluation and prediction can be combined with knowledge in the landslide domain, reducing the subjectivity of landslide evaluation and improving the accuracy and reliability of prediction.

This study utilized the efficient knowledge representation and flexible, scalable characteristics of ontology to conduct research on landslide domain ontology construction and hazard assessment.

Through in-depth analysis of the causes and characteristics of landslide disasters, a multidisciplinary ontology in the landslide domain involving geology, mechanics, hydrology, ecology, and other fields was constructed. A tripartite knowledge expression framework of landslide basic information, mechanisms, and deformation mechanisms was defined within the ontology, along with corresponding hierarchical structures and expression methods. Based on the constructed landslide ontology, rule intersection nodes were selected as indicators, and the contribution weighting method was applied to assess landslide hazards, which to some extent improved the accuracy of landslide evaluation. Through this research, it is aimed to enrich the conceptual knowledge of landslides at the theoretical level and improve the landslide knowledge system. It is expected to provide a theoretical basis and technical support for the prevention of landslide disasters, and to provide decision-making reference tools for geological disaster management departments and decision-makers.

2. Landslide Knowledge System

2.1. Landslide Knowledge Framework

Landslides, as a special and complex geological disaster system, encompass three elements: causative factors, predisposing environments, and susceptible bodies^[8]. However, these three elements fail to consider the dynamic changes during landslide development, leading to an insufficient understanding of the landslide destruction chain. This study, through summarization and induction, found that landslide knowledge exhibits characteristics of multidisciplinary, multilevel structure, interconnectedness, and rich information content. It includes basic attributes such as landslide volume, location, and damage losses, while also containing information on the geological, mechanical, and ecological characteristics and dynamic evolution of landslides across multiple disciplines. Therefore, based on the three elements of landslides, this study divides landslide knowledge into a tripartite knowledge framework of basic information, mechanisms of action, and deformation mechanisms, to meet the needs of landslide domain ontology construction and dynamic mechanism research.

The basic information of landslides covers various attributes such as the type, velocity, volume, and damage losses of landslides, which constitute a comprehensive and supplementary understanding of the causative factors and susceptible bodies of landslides. The mechanisms of landslide action expand upon the predisposing environments, involving physical parameter models of landslide incubation environments, surrounding anthropogenic factors, and the sliding deformation, reflecting the conditions for landslide incubation and dynamic correlation information. As for the deformation mechanisms of landslides, different classifications are made based on the predisposing environments or mechanisms of action of landslides, considering the destructive stages in which landslide bodies are located and the possible or already occurring movement states. This classification research reflects the adjustment trajectory and phenomena of internal stress states during the evolution process of landslides.

2.2. Types of Slippery Slope Knowledge

Landslide knowledge is a unified summary of landslide-related concepts, rules, relationships, and instances. Based on the different characteristics and functions of landslide knowledge, the research methods within the field of knowledge graphs are referenced to classify landslide knowledge^[9]. This classification aims to improve the organization, management efficiency, and quality of use of knowledge. Landslide knowledge is divided into four categories: conceptual knowledge, relational knowledge, rule knowledge, and instance knowledge. These categories are integrated into a landslide knowledge expression model, as shown in Fig 1.

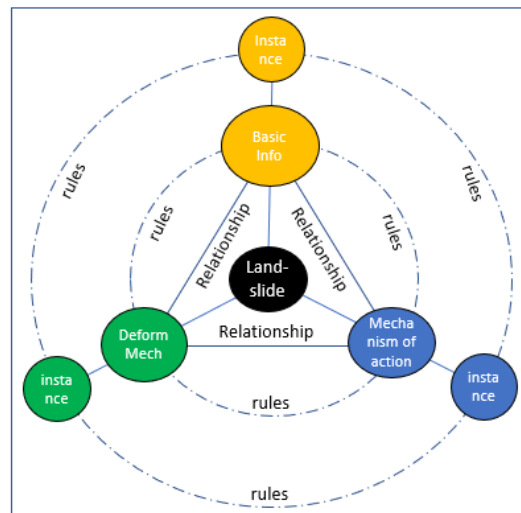


Fig. 1 Landslide Knowledge Expression Model

Conceptual knowledge serves as the foundation for constructing the landslide knowledge system, encompassing various concepts defined within the landslide domain. For instance, landslides consist of elements such as landslide bodies, landslide walls, and sliding zones. These concepts are subordinate to basic landslide information or deformation mechanisms and are further defined at different levels according to certain classification rules.

Relational knowledge constitutes a crucial part of establishing a comprehensive landslide ontology, comprising a collection of relationships describing various connections between concepts or instances, including semantic relationships and part-whole relationships. For example, secondary disasters of landslides encompass phenomena such as barrier lakes, secondary landslides, and collapses. Through diverse relationships, different concepts and instances are interconnected, forming an internal relational network of landslide knowledge.

Rule knowledge refers to the constraints and limitations defined between different hierarchical concepts of landslides, constituting the core of landslide knowledge. Different concepts need to be divided into sub-concepts according to certain rules. For example, landslide bodies are classified into extra-large, large, medium, and small-scale landslides based on their size. In addition to providing clear boundaries and unified standards, landslide rule knowledge also includes mechanistic knowledge of landslide occurrence, interactions between conceptual factors, and semantic rules such as destruction, liquefaction, and influence, as shown in Table 1. Numerous rule constraints connect different concepts and instances, forming a complex interwoven system of landslide knowledge, revealing the processes and mechanisms of landslide destruction.

Table 1. Main Rules of Landslides

Rule Name	Rule Description
Intrinsic Factors	Refers to the fundamental conditions for the formation and development of landslides. This includes the physical and chemical properties of the landslide body, the slope angle of the strata, the degree of rock weathering, etc.
Triggering Factors	Refers to the factors that directly influence the stability of landslides from the external environment or activities. This includes rainfall, earthquakes, human activities, etc.
Positive Promotion	Refers to the positive impact of a concept's gainful change on another concept, such as enhancement, promotion, elevation, or enlargement.
Negative Inhibition	The reverse rule of positive promotion, such as weakening, inhibition, reduction, or diminishment.
Influence	Refers to the impact of changes in one factor on another factor among hierarchical concepts, but the direction of this influence (positive or negative) depends on specific circumstances and relationships.

Instance Knowledge refers to the specific states and parameters summarizing the underlying concepts of landslide events, serving as concrete objects of a particular entity. These pieces of knowledge encompass specific geomorphological features, stratigraphic geological information, climatic vegetation information, etc., constituting a comprehensive collection of information about landslide disasters or hazard points.

3. Expression of Landslide Knowledge Based on Ontology

3.1. Landslide knowledge sources

The sources of landslide knowledge used to construct the landslide ontology must ensure the accuracy, authority, and authenticity of the information, meeting the high standards required for precise data processing. The knowledge in this paper mainly comes from five aspects: standard specifications, books and textbooks, journal papers, experimental data, and investigation reports. The acquisition of knowledge relies on a combination of manual extraction and machine learning, utilizing two mainstream knowledge extraction models, CRF and BiLSTM-CRF, to automatically extract landslide entities and attribute information from text^[10].

3.2. Landslide knowledge expression

Given the characteristics of landslide knowledge, this paper adopts Protégé software to establish the landslide ontology, realizing the storage, querying, and visualization of landslide knowledge. Protégé provides rich model architectures and functionalities, suitable for various forms of knowledge representation^[11]. By defining class properties, object properties, and data property descriptions of landslides, specific features and basic information are described, establishing the connection between landslide instances and conceptual knowledge. This forms the framework of landslide ontology expression, facilitating computers to understand, infer, and process structured knowledge in the landslide domain, as shown in Fig 2.

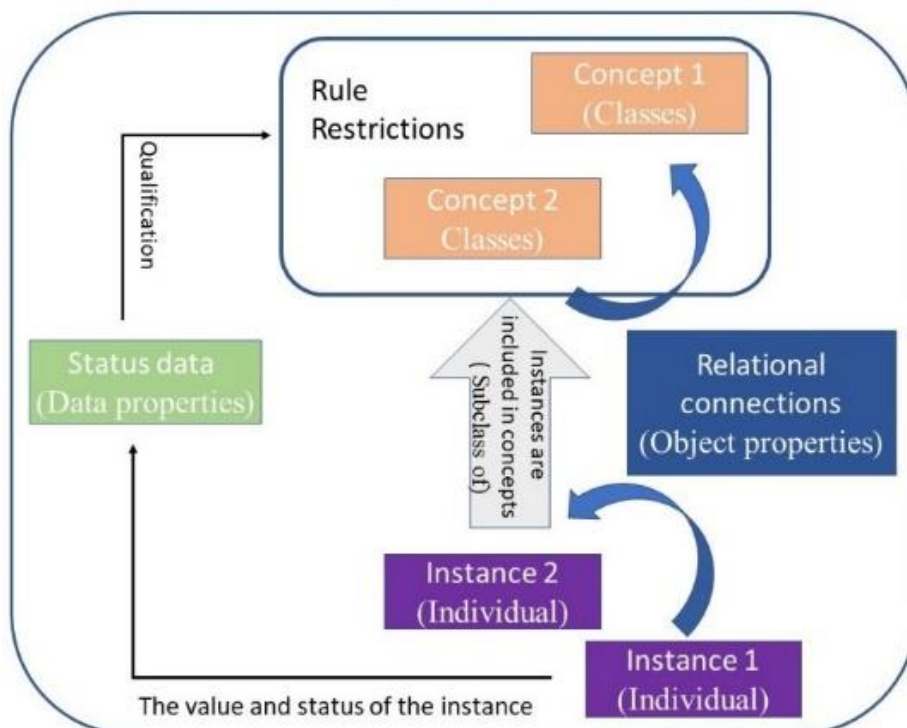


Fig. 2 Ontology expression framework

The concept knowledge of landslides is expressed in a top-down structure, following the logical rules of landslide evolution. It is divided into three main secondary concepts according to the landslide knowledge framework: basic landslide information, landslide mechanisms, and landslide

deformation mechanisms. Under these three secondary concepts, further sub-concepts are divided based on landslide relationship knowledge and rule knowledge. Taking landslide mechanisms as an example, it expands into seven tertiary concepts based on the landslide breeding environment and evolution rules: litho-geomorphology, geological structure, the role of water, surrounding environment, and physical parameters. Under each tertiary concept, more specific sub-concepts are divided based on the relationships between different concepts. The top-down concept structure aims to comprehensively cover landslide knowledge while ensuring its scalability for subsequent updates and supplements. The expression of some landslide concept knowledge is shown in Fig 3, where blue arrows represent inclusion relationships.

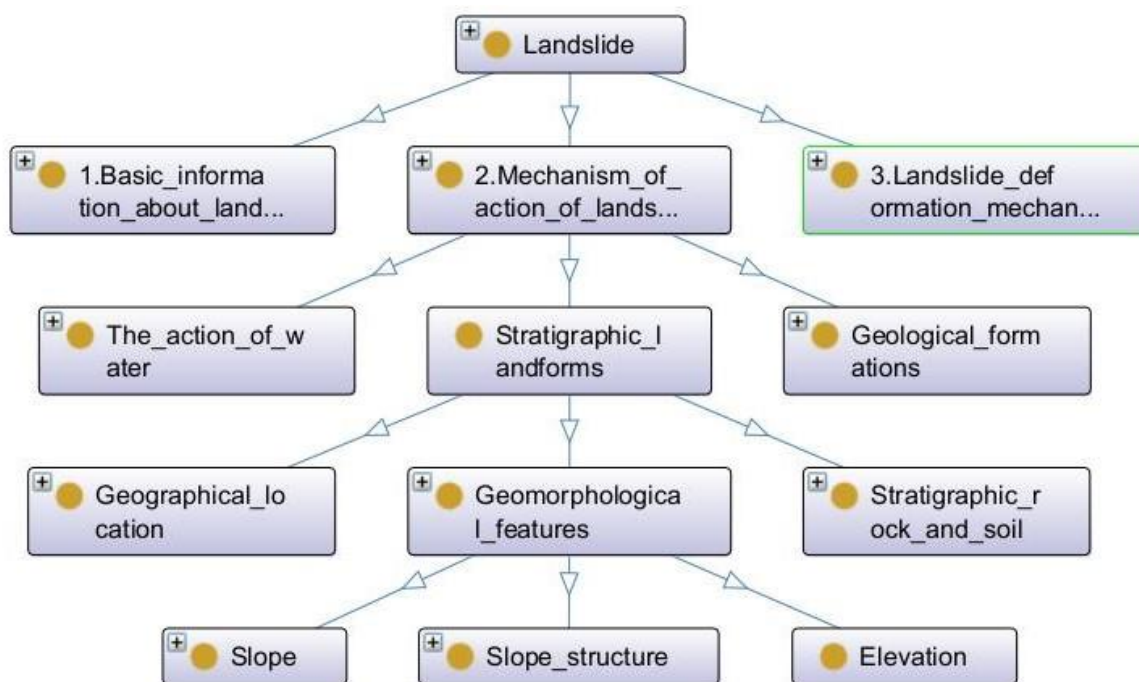


Fig. 3 Conceptual knowledge expression

The relationships between upper-level concepts of landslides are relatively simple, primarily serving a supportive role. However, there exist complex relationship knowledge and rule knowledge among lower-level landslide concepts of different categories (Table 1). Centered around the fourth-level concepts of "stratigraphy" and "geomorphological features," further expansion into lower-level concepts forms a dense network of rules. For instance, "climate" factors can influence "vegetation coverage" while simultaneously positively promoting "weathering erosion degree." Through the rules between different concepts, various factors' correlations constitute the real and complex core structure of the landslide ontology, achieving the visualization of the evolution mechanism of landslide destruction chains for already occurred or potentially risky landslide events.

The conceptual knowledge and rule knowledge are based on landslide theory, but a complete landslide ontology structure requires support from instance knowledge. The actual factor states of specific landslide events belong to upper-level landslide concepts as shown in Fig 4, while also corresponding to corresponding instance types. Taking the Xiangjiatang landslide in Fengjie County as an example, the specific "slope shape" is "three-level platforms + gentle terrain," belonging to a terraced slope; the "weathering erosion degree" is "debris rock weathering accumulation soil," exhibiting a strong weathering state. The colored dashed lines in the Fig 4 represent the states of different factors of landslide events, while solid lines represent the subordinate relationship. Rules can be used to analyze the destruction chains of specific landslide events and gain a deeper understanding of landslide destruction mechanisms.

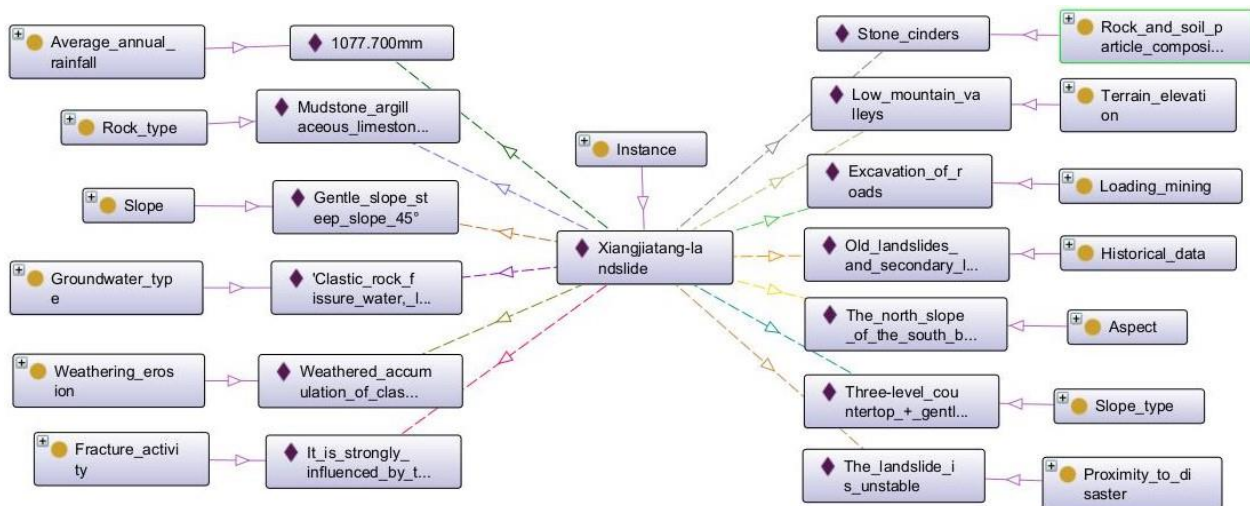


Fig. 4 Example knowledge representation

The construction of the landslide ontology organizes landslide knowledge into a large network structure, which is not only convenient for knowledge storage, display, and retrieval but also forms a comprehensive and complete visual organic whole. The landslide ontology established in this paper includes 320 categories, 16 attribute relationships, 1410 rule associations, 413 instances, 30 data types, and nearly a hundred annotation explanations.

4. Ontology-based landslide risk assessment

4.1. Evaluation indicators and weights

The process of landslide disaster occurrence and sliding can be divided into three stages: initial deformation, uniform deformation, and accelerated deformation^[12]. Precise quantification fitting of landslide deformation data imposes high demands on personnel and equipment. Based on the constructed landslide ontology, the factor superposition model with weight contribution method is adopted to leverage the advantages of high knowledge density of the landslide ontology. This method can break through the limitations of conventional superposition model in the number of index factors and compensate for the subjective selection of index factors.

The selection of landslide evaluation indicators must follow certain principles^[13]. In this study, focusing on Fengjie County in the Three Gorges Reservoir Area as the research area, from the constructed landslide ontology, the main rule intersection nodes of lower-level concepts are selected as evaluation indicators. Considering the principles of indicator selection and the difficulty of obtaining actual data, a total of 28 indicator categories and 126 indicator states are selected as adaptive key indicators for landslide hazard assessment in Fengjie County from 92 categories of breeding factors and 379 factor states in the landslide ontology, including slope morphology, lithology, structural geology, surface and subsurface water, human activities, and environmental factors.

Traditionally, the weighting of single-digit indicators mainly comes from expert scoring, statistical analysis, and experimental data. Considering the numerous types and large quantity of ontology indicators, reliable and accurate factor weights are obtained through the combination of three weighting sources, utilizing the knowledge associations of the constructed landslide ontology. As shown in Table 2, by setting different weights, the differences in the impact of factors on landslide development are reflected. The higher the weight, the greater the impact of the factor on landslide development.

Table 2. Examples of some evaluation indicators and weights

Perspective	Index C_i	Index Weight w	Index Status C_{ij}	Index Status Weigh r
Environmental factors	Signs of deformation	0.111	The signs are clear	1.000
			The signs are more obvious	0.600
			There are signs	0.300
			The signs are not obvious	0.100
			No signs	0.000
Slope morphology	Slope height	0.044	≥ 50 meters	1.000
			50-30 meters	0.700
			30-10 meters	0.400
			< 10 meters	0.100
Structural geology	Rock mass structure	0.028	Bulk structure	1.000
			Fragmented structures	0.500
			Laminar structure	0.200
			Block-like structure	0.100
Surface/Groundwater	Groundwater type	0.028	The upper layer of the fracture pore is stagnant	1.000
			Fracture pore diving	0.500
			Other hydrogeological structures	0.300

4.2. Implementation of Risk Assessment

In landslide hazard assessment, specific landslide risk assessment is set as the target layer; the main aspects leading to landslide disasters are considered as the criterion layer; factors such as slope gradient, lithology, altitude, etc., are taken as the indicator layer, and the states of each indicator serve as the basis for judgment, forming the hierarchical structure of landslide hazard assessment, as shown in Fig 5.

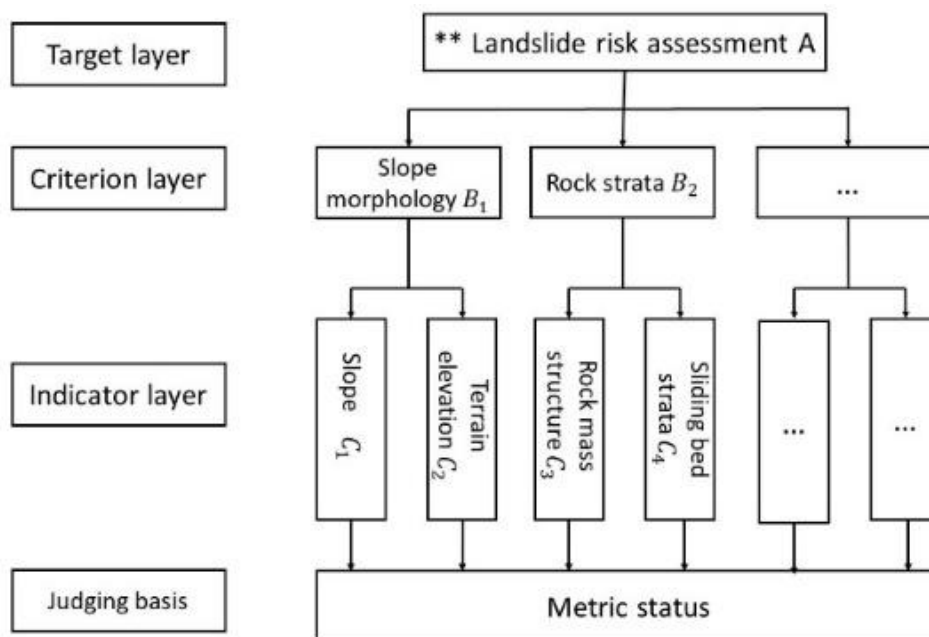


Fig. 5 Hierarchical structure of landslide ontology risk assessment

After selecting the criterion layer and the indicator layer, determining the weights of each evaluation factor, and according to the principle of multi-factor superposition, the landslide hazard calculation model is established as follows:

$$H = \sum_{i=1}^{28} w_{C_i} r_{C_{ij}}, (i = 1, 2, \dots, 28; j = 1, 2, \dots) \tag{1}$$

In the equation: H —represents the hazard degree of a single landslide, w_{C_i} —represents the weight of the landslide indicator, $r_{C_{ij}}$ —represents the weight of the landslide indicator state.

In Protege, users are allowed to define logical rules to transform the landslide hazard calculation model rules into SWRL rules. Utilizing the Pellet reasoning engine, the ontology rules formula (1) are inputted, and combined with ontology decision rules as shown in Table 3, to perform reasoning operations, achieving automatic inference and assessment of landslide hazard. Referring to the actual situation and computational results, this paper divides into five levels with a tolerance of 0.20.

Table 3. Landslide risk classification and ontology decision-making

Hazard Level	Gradation of Hazard Level	Characteristics of Disaster	Ontology Decision Rules
0.0-0.20	Extreme Low Hazard Level	Essentially stable, temporarily not posing disaster conditions.	
0.20-0.40	Low Hazard Level	Localized creeping may occur, currently not causing disasters or significant harm.	(Hazard Level some xsd:decimal[>Lower Limit])
0.40-0.60	Moderate Hazard Level	In the initial phase of destruction, resulting in minimal disasters and significant harm, easily manageable.	and (Hazard Level some xsd:decimal[<= Upper Limit])
0.60-0.80	High Hazard Level	Internal deterioration is in the developing stage, capable of causing major disasters, manageable.	
0.80-1.00	Extreme High Hazard Level	In the imminent sliding stage or overall accelerating destruction stage, posing enormous harm and extremely difficult to manage.	

4.3. Evaluation results and analysis

This article summarizes the development characteristics of landslides in Fengjie County, Three Gorges Reservoir Area, and selects adaptive landslide evaluation indicators accordingly. Taking the landslides in Xiangjiangtang, Hefeng Township, and Guihuajing as examples, landslide hazard assessment is conducted based on the established landslide ontology, and graded warnings are issued, as shown in Table 4.

Table 4. Landslide Hazard Assessment in Fengjie County

Landslide area	Landslide name	Hazard Level	Gradation of Hazard Level classification
Fengjie County, Three Gorges Reservoir Area	Xiangjiangtang Landslide	0.803	Extreme High Hazard Level
	Hefeng Township Landslide	0.764	High Hazard Level
	Guihuajing Landslide	0.783	High Hazard Level
	Xinpu Landslide	0.768	High Hazard Level
	Chenjiagou Landslide	0.811	Extreme High Hazard Level
	Baiyian Landslide	0.772	High Hazard Level
	Zengjiapeng Landslide	0.733	High Hazard Level

According to the landslide ontology risk assessment rules, the Xiangjiangtang Landslide has a hazard index of 0.803, belonging to the category of extremely high-risk landslides, as shown in Fig 6. Combining the rules of the ontology with the actual situation analysis, in its natural state, the overall landslide is stable, but under the influence of the river water, the front edge of the landslide may experience local deformation leading to a landslide. After the operation of the Three Gorges Reservoir, the front part of the landslide is submerged, the groundwater level rises, the water content of the landslide body increases, and the stability of the Xiangjiangtang Landslide decreases, putting it in a state of instability.

The evaluation results were verified, according to the literature^[14]: Below the elevation of 135m, the Xiangjiangtang Landslide shows severe cracking and deformation of the sliding body and buildings, with frequent local slip sliding of the shallow sliding body above. Stability calculations for the secondary landslide show a stability coefficient of 1.03, which decreases to below 1 after the water level rises in the Three Gorges Reservoir. In the section of the secondary landslide body, the landslide is in an unstable state after the operation of the Three Gorges Reservoir. The hazard assessment results are basically consistent with the conclusions of previous researchers.

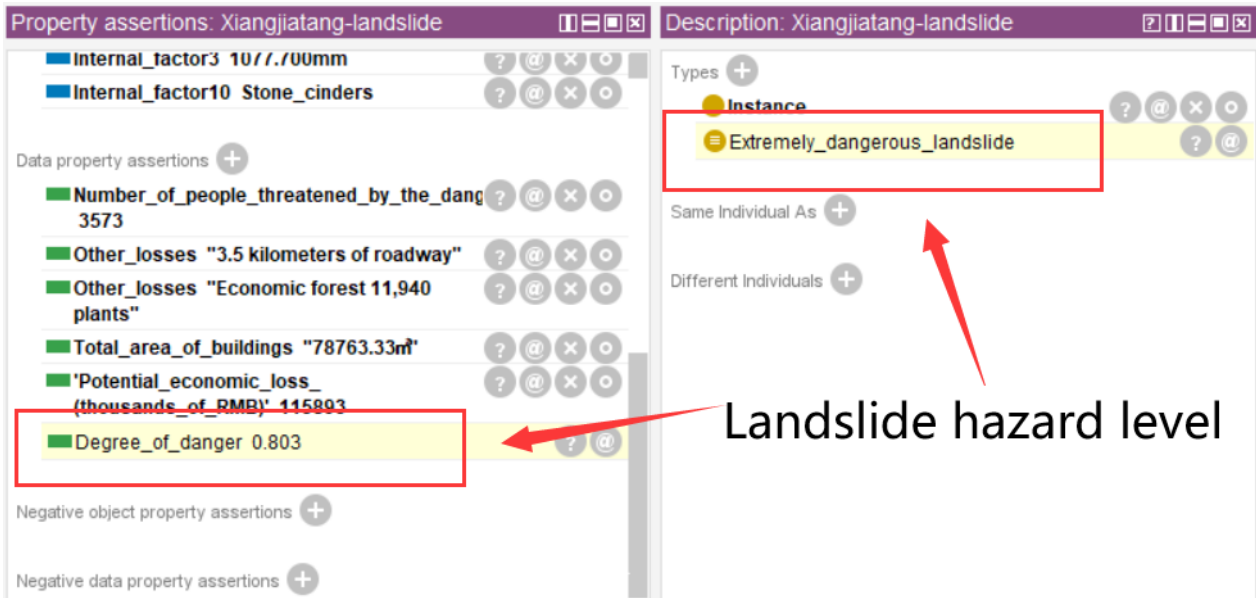


Fig. 6 Xiangjiangtang Landslide Ontology Inference Evaluation Results

5. Conclusion

This study has achieved the visualization and management of landslide knowledge by conducting in-depth research on landslide knowledge and constructing an ontology in the landslide domain. Leveraging the knowledge-intensive nature of the ontology, landslide risk assessments were performed with satisfactory results. The establishment of the landslide ontology to a certain extent ensures the consistency of knowledge in the landslide domain, thus avoiding conflicts in results caused by inconsistent knowledge.

However, there are still some shortcomings in this research. Firstly, it is necessary to further refine and enrich the construction of the landslide ontology, including refining and enriching the landslide knowledge system, to improve the completeness and usability of the ontology. Secondly, in landslide risk assessments, more consideration needs to be given to the characteristics of different regions and types of landslides, in order to conduct more comprehensive risk assessments and predictions. Additionally, further exploration is needed on how to integrate ontology construction with practical engineering applications to validate the effectiveness and practicality of ontology construction.

References

- [1] Peng Jianbing, Wang Qiyao, Men Yuming, Xu Qiang, Zhuang Jianqi. Landslide Disaster on the Loess Plateau [M]. Beijing: Science Press, 2019: 26-33.
- [2] Cheng Gang, Wang Zhenxue, Li Gangqiang, et al. Progress and Prospect of Landslide Integrated Full-Dimensional Monitoring and Early Warning Technology [J]. Journal of Safety Science of China, 2023, 33(1): 105-114.
- [3] Xu H, Zhao Y, Huang H, et al. A comprehensive construction of the domain ontology for stratigraphy [J]. Earth Science Frontiers: English Edition, 2023.

- [4] Wang Yanni,Liu Gang.Research and Application of Ontology in the Field of Geological Disasters [J].Geography and Geo-Information Science,2011,27(06):36-40.
- [5] Li Aihua,Xu Yize,Chi Yuxue.Review on Ontology Construction and Application [J].Information Studies:Theory & Application,2023,46(11):189-195.
- [6] Xu Qiang,Cui Shenghua,Huang Wei,et al.Research on Construction Method of Landslide Knowledge Graph for Engineering Geology Field [J].Journal of Wuhan University (Information Science Edition),2023,48(10).
- [7] Zhu Qing,Zeng Haowei,Ding Yulin,et al.Review on Analysis Methods of Major Landslide Hazards [J].Acta Geodaetica et Cartographica Sinica,2019,48(12):1551-1561.
- [8] Shi Peijun.Theoretical and Practical Aspects of Disaster System Research:Four Discussions [J].Journal of Natural Disasters,2005(6):1-7.
- [9] Liu Yu,Li Yong.Method and Application Prospect of Knowledge Graph Construction for Urban Commercial Districts/Blocks Oriented to Urban Sustainable Development [J].Journal of Geo-Information Science,2023,25(12):2374-2386.
- [10] Liu Wencong.Construction and Application of Landslide Knowledge Graph [D].Hefei University of Technology,2022.
- [11] Tebes G,Olsina L,Peppino D,et al.TestTDO:A Top-Domain Software Testing Ontology [C]//XXIII Ibero-American Conference on Software Engineering (CIbSE'20),2020.
- [12] Jie Mingli,Zhao Jianjun,Ju Nengpan,et al.Study on Spatio-Temporal Evolution Laws of Landslides with Multi-Source Data:A Case Study of Huangnibazi Landslide [J].Journal of Wuhan University (Information Science Edition),2020,45(06):923-932.
- [13] Ding Mingtao,Hu Kaiheng,Zhang Ji.Theory and Practice of Landslide and Debris Flow Disaster Risk Assessment in Southwest Mountainous Areas [M].Beijing:Science Press,2022:118-125.
- [14] Jia Jianhong,Mao Chengwen,He Jinming.Identification and Engineering Practice of Landslide Disasters in the Three Gorges Reservoir Area [M].Wuhan:China University of Geosciences Press,2017:1-10,108-127.