Article

BIM-Based Dynamic Construction Safety Rule Checking Using Ontology and Natural Language Processing

Qiyu Shen [1], Songfei Wu [1], Yichuan Deng [1,2,*], Hui Deng [1] and Jack C. P. Cheng [3]

1   School of Civil Engineering and Transportation, South China University of Technology,
    Guangzhou 510641, China; qushen@scut.edu.cn (Q.S.); scut_wsf@163.com (S.W.); hdeng@scut.edu.cn (H.D.)
2   State Key Laboratory of Subtropical Building Science, South China University of Technology,
    Guangzhou 510641, China
3   Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology,
    Hong Kong SAR, China; cejcheng@ust.hk
*   Correspondence: ctycdeng@scut.edu.cn; Tel.: +86-020-87111030

Abstract: Real-time identification and prevention of safety risks in dynamic construction activities are demanded by construction safety managers to cope with the growing complexity of the construction site. Most of the studies on BIM-based construction safety inspection and prevention use data from the planning and design stage. Meanwhile, safety managers still need to spend a lot of time gathering reports about construction safety risks in certain periods or areas from inferred results in BIM. Therefore, this paper proposed an automatic safety risk identification and prevention mechanism for the construction process by integrating a safety rule library based on ontology technology and Natural Language Processing. An automatic inspection mechanism integrating BIM and safety rules is constructed, and a presentation mechanism of intelligent detection results based on Natural Language Processing is designed. The construction process safety rule checking system was developed, and the effectiveness of the system was verified by a case study. The outcome of this paper contributes to the development and application of ontology in construction safety research, and the NLP-based safety rule checking result presentation will benefit safety inspectors and construction managers in practice.

Keywords: construction process; safety compliance checking; ontology; BIM; NLP

1. Introduction

As construction accidents continue to cause casualties and economic losses, the efficiency of safety management and the reduction of the occurrence of accidents remain an urgent problem to be solved. The current way of safety management heavily relies on safety managers' understanding of safety regulations and their experience. With the growth of scale and complexity of projects, the tasks of safety inspection and risk prevention are time-consuming and error prone. Analyzing the causes and formation mechanisms of safety accidents is the prerequisite for preventing safety risks [1]. At present, the accident causation theory is often used to analyze the causes of accidents. The theory of the cause of construction safety accidents refers to the analysis of typical accidents, thus as to extract the mechanism of accident occurrence and to establish the accident causation model. This can provide a scientific basis for the qualitative and quantitative analysis, prediction, and prevention of accidents, as well as the improvement of safety management [2]. The accident causation theory has experienced three stages, namely the early accident causation theory, the accident causation theory after World War II, and the modern system safety theory [3]. The early accident causation theory includes accident frequency tendency theory, Heinrich accident causation sequence theory [4], Frank Bode's modern accident causation theory [5], and the accident causation theory after World War II includes the trajectory cross theory and accidental release of energy theory. For building structures with increasingly complex structures and sizes, the sources of safety accidents are many and complex, including

unsafe behaviors of construction workers, errors of safety management personnel, unsafe conditions of construction machinery and tools on-site, and unsafe conditions. Protective measures and environment. The modern system theory is a theory specifically aimed at complex accident systems. In order to comprehensively manage the hidden dangers and problems of the construction site, it is necessary to sort out the sources of various safety problems during the construction process in order to establish a comprehensive safety prevention system. This paper adopts the modern system safety theory, which believes that the most important factor that causes accidents is the various hazards in the production process. The theory also believes that the main work of safety management is to identify, evaluate, and control hazard sources.

Therefore, the automatic identification of construction safety risks and their corresponding solutions has gained more and more attention from researchers worldwide. However, current studies on the issue mainly focus on the use of BIM (Building Information Modelling) technology in safety inspection and risk prevention from the design stage. While the limited range of safety risks in construction are covered in current studies, such as openings or edges, it is believed that real-time identification and prevention of safety risks in dynamic construction activities will provide a more desirable solution to the ever-growing complexity of the construction site. In addition, the current research is unable to form a system of reusable safety knowledge in the form of standardized provisions for use in subsequent safety management projects. There still remains a gap in the establishment of reusable construction safety management knowledge and its application in the dynamic construction site.

In recent years, advanced information technologies such as BIM, ontology, NLP (Natural Language Processing) are introduced in the construction industry to improve production efficiency. Ontology, which can be interpreted as a knowledge description model, was originally derived from the field of philosophical concepts. A complete ontology consists of at least five parts: concept or class, relation, function, axiom, and instance. An instance is a concrete instance of a concept [6]. Ontology has been widely used in information retrieval, semantic web, knowledge management, digital library, and other fields. As an important knowledge representation model, ontology technology has been widely studied in the field of construction safety management, especially in information reasoning and extraction. Wang and Boukamp (2011) [7] established an ontology for analysis of potential job hazards, including construction activities and specific operation steps, and established a set of ontology reasoning mechanisms to realize the safety inspection of construction activities, with which the response rate of construction companies to construction activity risks was improved. To identify specific safety risks and automatically generate safety measures, Lu (2015) et al. [8] established five categories of CSCOntology (construction safety checking ontology), including the line of work, task, precursor information, hazard, and solution. The rules regarding high falls as promulgated by OSHA (Occupational Safety and Health Administration) were first translated into SWRL (Semantic Web Rule Language) rules that can be processed by computers, the JESS inference machine was then used to realize the inference for a given instance of ontology and generate corresponding solutions. However, the information needed for detection still needs to be created and entered manually, implying an insufficient degree of automation. Kim et al. (2016) developed a BIM-based scaffolding automatic planning system. The system used simulation engines and BIM to simulate daily construction progress and identify potential hazards according to safety specifications. Zhang (2015) et al. [9] proposed an automatic safety planning system combining BIM and ontology. This system established the mapping between BIM components and concepts in the ontology and then used SWRL rules in the ontology to carry out a safety inspection on automatically extracted masonry components in BIM and give a visual results report. In 2016, Ding et al. [10] applied ontology technology and semantic network technology in the safety risk management of deep foundation pit construction. There are several shortfalls in the current research of ontology applications in construction safety management. Firstly, the information for reasoning using ontology either comes from

manual input or is extracted from a static BIM model in the phase, both of which cannot represent the dynamic nature of construction safety management. Secondly, a limited range of construction safety management is covered in current research, with emphasis on falls from height and border identification. Lastly, the results of safety checking are presented in the form of databases or tools for ontology reasoning, which imposes restrictions on construction managers who are not familiar with such tools.

Natural Language Processing is an important direction in the field of computer science and artificial intelligence, which helps computers understand the real meaning of human natural language. NLP is mainly used for information extraction and retrieval from documents written in natural language. At present, there have been increasing applications of NLP in construction safety management. One typical study was reported in Zhang and EI-Gohary (2015) [11], which used established semantic mapping rules and conflict resolution rules to convert some chapters of the International Building Code 2009 into logical sentences that can be used for safety code inspection. In this study, NLP was used for syntactic analysis, and ontology was used for semantic analysis. In order to automatically extract relevant information from the BIM model for automatic specification checking, Zhang and EI-Gohary (2016) [12] proposed the use of NLP to extract concepts related to consistency checking from the specification. The concepts were then mapped to the IFC (Industry Foundation Classes) hierarchical relationship to automatically match the information in the specification and the BIM model. In another study, Zhang and EI-Gohary (2017) [13] developed an NLP-based system to automatically extract and transform specification information and design information in BIM, which were then used in consistent reasoning. Lin (2015) [14] et al. proposed a method for intelligent retrieval and presentation of BIM cloud data based on NLP. The number of materials used on the construction site and the amount of progress completed can be presented in an intelligent way through inquiry.

In light of recent advancements in safety rule checking and result presentation, although progress was made recently towards intelligent construction safety management, there were some limitations that need to be resolved. Firstly, most of the studies were on BIM-based construction safety inspection and prevention use data from the planning and design stage. There were relatively few studies on real-time safety risk inspection and prevention of complex and changeable construction sites. Secondly, there were few studies using BIM as the information source for a safety inspection, which reported that a lot of manual intervention was still needed to name BIM model components in order to realize the mapping relationship between safety rules and BIM components. Lastly, while BIM has shown its potential in inferring construction safety risks and prevention measures, safety managers still need to spend a lot of time gathering reports about construction safety risk in certain periods or areas from these inferred results. There are few studies that focus on how to intelligently present construction safety reports thus that safety managers can quickly obtain safety issues for a real-time safety inspection.

This paper contributes to the body of knowledge of construction safety management by the establishment of a BIM-based construction process safety risk inspection system integrating ontology and Natural Language Processing. In this study, BIM is used to contain the dynamic construction process and serves as an efficient information retrieval hub. Combined with ontology and Natural Language Processing, the information in BIM is used to identify and prevent construction safety risks, thus as to achieve day-to-day dynamic and comprehensive safety management. Two research questions are addressed in this paper: (1) building an ontology-based computer-recognizable safety regulation knowledge base that helps to automatically identify and prevent safety risks in everyday construction activities; (2) comprehensive and easy-to-use ways of providing construction managers with real-time safety risk identification results and prevention measures. The work presented in this paper not only effectively reduces the workload of managers but also improves the accuracy of safety risk inspections and ultimately achieves the goal of reducing safety accidents and improving construction safety management.

2. Methodology

2.1. Methodology Overview

In this paper, a BIM-based dynamic construction safety rule checking framework using ontology and Natural Language Processing was proposed to overcome the shortcomings of current research. The methodology of the framework development included 3 major steps, namely the development of construction safety management ontology, the establishment of a safety rule library for the construction process and prosing a safety risk identification and retrieval mechanism. The framework can be illustrated in Figure 1.
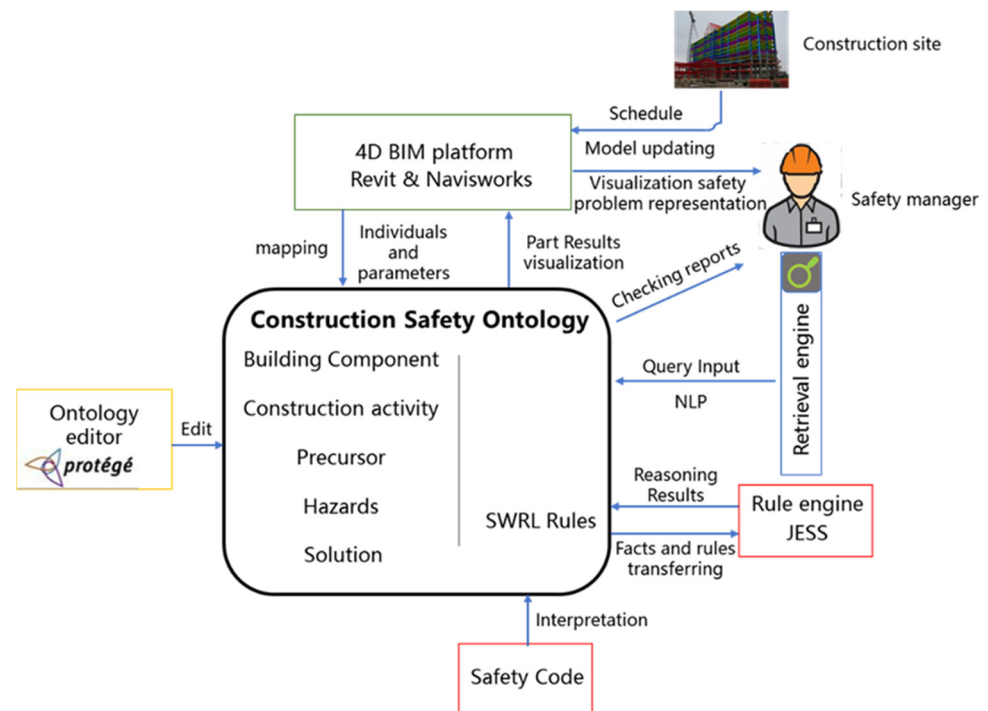
Figure 1. Illustration of the framework.

2.2. Development of Ontology

The construction safety management ontology integrates the relevant knowledge of the construction process, various types of safety accident knowledge related to the construction process, and the precursor information and corresponding safety solutions of the accident. The improved 7-step method was applied to complete the construction of the safety ontology. The standard 7 steps of ontology development were: (1) determine the domain and scope of the ontology; (2) consider reusing existing ontologies; (3) enumerate important terms in the ontology; (4) define the classes and the class hierarchy; (5) define the properties of classes—slots; (6) define the facets of the slots; and (7) create instances. The development of the safety ontology specifically includes 5 processes.

2.2.1. Determine the Scope of the Ontology

Since the establishment of the ontology is a complex and elaborate process, the first step in building the ontology is to determine the professional domain and scope of the established ontology, that is, to define the role of the ontology, the domain of the application, the user, and the maintenance of the ontology. In this paper, the developed ontology will focus on the civil construction process and renovation process of the building, as well as the corresponding safety issues and solutions. The development of the ontology will include the review of the concepts and interactions of the construction process, precursor, hazard, and solution based on relevant construction specifications and safety specifications.

2.2.2. Reusing the Existing Ontology

The existing ontology related to the scope of the study should be reused in order to reduce the difficulty of developing the ontology. Jiao [15] established a subway construction safety risk ontology knowledge database, and Zeng [16] established a knowledge database of subway construction safety accidents, both of which discussed the subway construction process from the perspective of accident causes, the fundamental aspects of the subway construction process, and the risk management ontology corresponding to the solution. Zhang [17] established a risk management ontology of construction engineering under a BIM environment to realize the semantic reasoning and retrieval of risk accidents. With reference to the structure and establishment process of these ontologies, this paper established an ontology that can effectively manage the safety risks of construction sites in view of the dynamic characteristics of the construction process in the construction field.

2.2.3. Hierarchy and Properties of Classes

This paper used a top-down approach to enumerate and define classes. The ontology classes constructed in this section mainly include 4 categories: construction process; precursor; hazard; and solution, each of which includes different sub-classes, hierarchy of these sub-classes, and the relationship between them (Table 1). In this paper, the WBS (Work Breakdown Structure) decomposition results included C1 (civil engineering), C2 (decoration engineering), and C3 (facility and pipeline). C1 (civil engineering) included 4 categories, and C2 (decorative engineering) included 7 categories.

Table 1. Category concept definition.

Name	Definition
Construction Process	The activities of construction
Precursor	Status or condition that may cause a safety incident
Hazard	Safety accidents caused by precursory information such as falling, collapse, and object striking
Solution	Preventing safety accidents or actions

The precursor information included 4 parts: structural member; material; equipment; and; environment, and these 4 precursor information classes were each divided into 3 different sub-classes. An example specification of the precursor information is shown in Table 2.

Table 2. Example of precursor definition.

Construction Process	Precursor	Location	Cause
C1.3 Formwork, scaffolding; C1.4 The stacking of masonry materials and construction machinery and tools in the masonry engineering; C2.1 Veneer engineering C2.2 Veneer engineering machinery stacking	Protective thickness at the top of the protective shed	Pedestrian access shelter	The thickness of the top of the pedestrian access shelter is not enough to resist falling objects from high places
	Geometry size of the protective shed	Pedestrian access shelter	The geometric dimensions of the pedestrian access protection shelter do not meet the requirements of the specification
	Safe net	Main building	The safety net around the main building is not erected or the erection quality/density does not meet the specification requirements
	Scaffolding position	Main building	The stacking position of the scaffolding is too close to the unprotected opening or the edge
	Tool stacking position	Building main body/scaffold board	The stacking position of tools is too close to the unprotected hole or the edge, or placed on the scaffold
C1.3 Tower crane installation and dismantling in the main structure project	Tower crane installation/removal	Tower crane	Tower crane installation and demolition have an intersection with construction pedestrians within a certain safe distance

The hazard component included 5 parts: fall hazard; struck; collapse; lifting injury; electric shock; fire; and explosion.

Solutions included staff protection system and safety control of material and equipment.

2.2.4. Definition of Class Attributes

The definition of class attributes in this paper included 2 aspects: the definition of the object property and the definition of data type property. The object attributes define the relationship between classes and classes, while the data attributes define the relationship between the classes and data.

Object property: the ontology established in this paper included 4 major categories, and the logical relationship between the 4 categories can be expressed by 3 relational terms. Some construction processes or precursors generated by unsafe construction state can be expressed by "Has_precursor"; the precursor may cause construction safety accidents (hazard), thus "Cause_hazard" can be used; and construction accidents (hazard) occur, which need corresponding countermeasures (solution) to prevent or solve, thus "Has_solution" can be used to express the relationship.

The formal expression of these attributes can be expressed by Domain U and Range U in the property's toolbar of Protégé 5.5.0 from Stanford University in the United States, an open-source ontology editing software. Each construction process may have several kinds of precursory information, but each specific precursory information instance can only belong to a construction process, and the setting of the attributes of objectivity can be seen in Table 3. The process of attribute construction in Protégé is shown in Figure 2.

Table 3. Object property definition.

Object Attribute	Domain	Range	Function Characteristics
Has_precursor	Construction Process	Precursor	Inverse Functional
Cause_hazard	Precursor	Hazard	Inverse Functional
Has_solution	Hazard	Solution	Inverse Functional
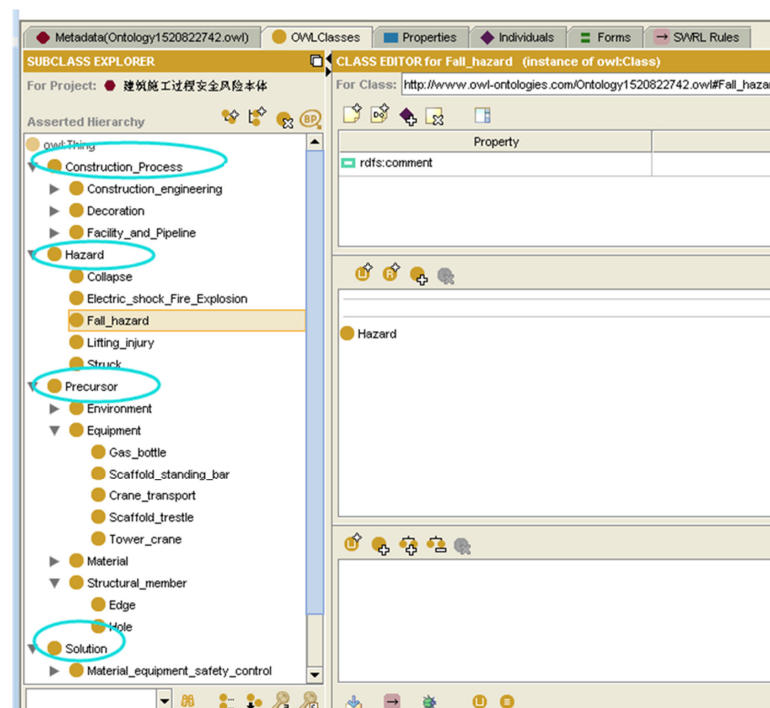
Figure 2. Object relationship setting.

The relationship between classes of the same level was defined above, and the relationship between different levels was an inclusion relation. According to the predefined definition of Protégé software from Stanford University in the United States, "is_a" represents the attribution relationship between the class and the inclusion class. For example, "Block_Masonry" of brick masonry projects belongs to a larger category of construction processes. Thus far, the definition of object type attributes of different classes at the same level was completed. Figure 3 presents an OntoGraf view of the class hierarchy.
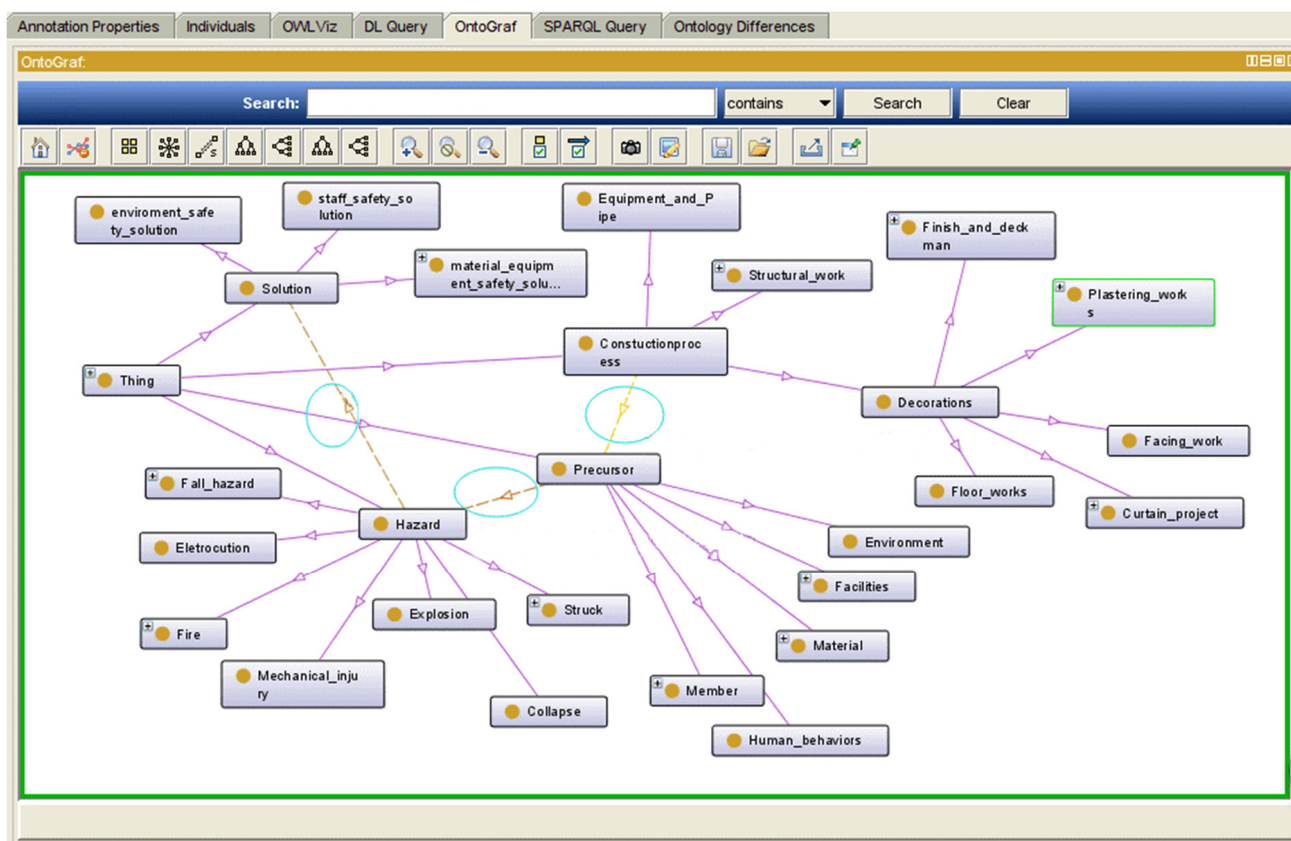
Figure 3. Class and class hierarchy relationship OntoGraf view.

Datatype Property: this paper defines data type attributes according to the characteristics of the 4 major classes, and the characteristics of the sub-classes contained in each class were relatively different; thus, for convenient construction, this paper summarizes the characteristics of each sub-class to the main class.

The datatype properties of the construction activities were summarized as follows: (1) the basic information includes the construction activity number, construction activity location, construction content, and start and end time; (2) the precursory information that may be included in the construction process, such as the components, materials, equipment, environment, specific attribute names, value types and function characteristics, as shown in Table 4.

Table 4. Construction activity data type attribute definition.

Data Type Attribute Names	Value Types	Function Characteristics
Construction_ID	int	Functional
Construction_Location	float	Functional
Construction_Task	string	Functional
Start_time	string	Functional
Ending_time	string	Functional
Construction_member	string	Functional
Construction_material	string	Functional
Construction_equipment	string	Functional
Construction_environment	string	Functional

The datatype properties of the precursory information were summarized as follows: the number of construction activities, the number of precursory information entries, and the description of precursory information, including the components, materials, equipment, and environment. Different types of precursory information have different characteristics; thus, different types of precursory information have different data attributes. For example, when considering the precursor information of the hole, the information regarding whether there was a cover (has_cover), whether there was a guardrail (has_guardrail), and the direction of the hole (has_direction) should be included. For each of the 4 types of values and functions, the characteristics are shown in Table 5.

Table 5. Precursor information data type attribute definition.

Data Type Attribute Names	Value Types	Function Characteristics
Precursor_ID	int	Functional
Belong_to_Constuction	int	Functional
Precursor_Description	string	Functional
Precursor_member	string	Functional
Precursor_material	string	Functional
Precursor_equipment	string	Functional
Precursor_environment	string	Functional

The datatype properties of the construction risk were summarized as follows: construction risk number, accident location, correspondence to the precursor information number, and accident occurrence time, as shown in Table 6.

Table 6. Construction risk data type attribute definition.

Data Type Attribute Names	Value Types	Function Characteristics
Risk_Number	int	Functional
Cor_to_Precursor	int	Functional
Risk_Description	string	Functional
Risk_Location	string	Functional
Risk_Task	string	Functional
Risk_Time	string	Functional

The datatype properties of the safety measures were summarized as follows: safety measure quantity, risk accident number, description of measures, safety measures for personnel, measures for materials, and measures for equipment, as shown in Table 7.

Table 7. Safety measure data type attribute definition.

Data Type Attribute Names	Value Types	Function Characteristics
Solution_ID	int	Functional
Risk_ID	int	Functional
Measure_Description	string	Functional
Solution_Person	string	Functional
Solution_Material	string	Functional
Solution_Equipment	string	Functional

With the object and datatype properties of the ontology class established above, the ontology was then stored in Protégé. Figure 4 shows the data type attributes of the construction process number (Construction_ID) in the construction process class.
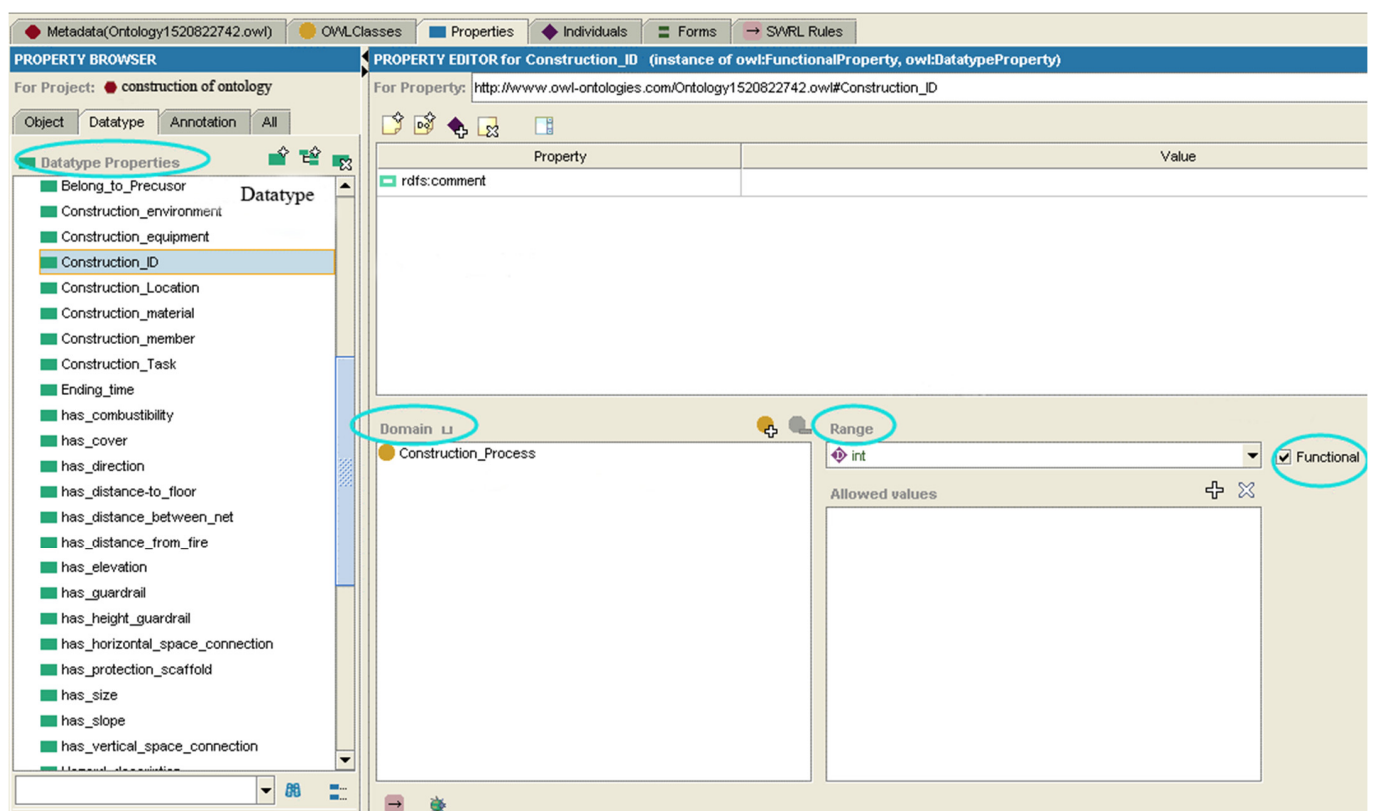
Figure 4. Various data type attribute definition interfaces.

2.3. Establishment of a Safety Rule Library for the Construction Process

2.3.1. Parameterization of Safety Rules

The safety rules of the construction process refer to the regular patterns or logical relationships that use certain pre-set parameters to determine whether certain precursors will lead to safety accidents in the construction process. The composition of safety rules generally includes 3 parts: the subject of the accident; the judgment conditions (parameters); and the safety risk. A general safety rule checking process starts by determining the subject of the accident through the inspected object and then continues to verify the inspected object through the distinguishing conditions to determine whether it has violated safety rules. The rule interpretation refers to the conversion of the normative terms expressed in natural language into a form that can be recognized by the computer, which forms the basis for automatic safety rule checking. Therefore, in order to fully identify the safety accidents in the construction process, the interpretation of each safety rule should consider the comprehensiveness of the discrimination conditions. The safety rule library formed

by these safety rules in a unified form can become an important tool to help the safety managers at the construction site to identify potential safety hazards and promptly give safety precautions, avoiding complicated manual search and regulation checks.

The interpretation of safety rules needs to consider 2 key issues: the identifiability of the subject of the safety risk and the discrimination conditions. The premise of whether safety rules can be applied digitally is to find the "object/subject" that may cause accidents, and the key to the effective implementation of the judgment condition of the accident subject is the detail of the subject's attribute parameters. In addition, the safety rules were norm clauses compiled by experts of different professions and described in natural language, which makes the subject of the safety risk and the judgment conditions vague and abstract; moreover, for the same kind of safety risk, different norms and clauses have different focuses or overlap; thus the interpretation of safety rules needs to consider the expression and comprehensiveness of the clauses. Therefore, the process of interpreting safety rules was to clarify the subject of the accident. Moreover, for the same kind of risk subject, we should comprehensively collect the safety clauses describing the risk in the safety regulations with different expressions and then sort out the logical relationships among them according to the characteristics of the clauses to realize the operability or usability of safety rules.

In this paper, the establishment of safety rules can be preliminarily divided into 2 categories according to the different accident subjects and different information retrieval processes:

(1) The subjects of safety risks involve the components produced in the construction process, such as floors and balconies;

(2) The subjects of safety risks occur in dynamic construction operations or processes, such as welding operations.

A complete set of safety rule interpretations should include information about the subject, location, attributes, and parameters of the accident and solutions. In this paper, related safety standards in China [18–23] were combined with the ontology defined in Section 3.1 to establish the safety rules used in this study. The result of safety rule parameter establishment for a hole is shown in Table 8.

Table 8. Safety rule parameter establishment for a hole.

Accident Subject	Location	Attributes	Parameter Information	Treatment
Hole	Reserved hole (Floor, roof) platform	Landscape	L < 25 cm; No protection	Solid cover
		Landscape	$25 \leq L < 50$ cm; No protection	Cover
		Landscape	$50 \leq L < 150$ cm; No protection	Steel mesh
		Landscape	L > 150 cm; No protection	Protective fence, safety net
	Wall (Window)	Vertical	H < 80 cm and V > 2 m; Unprotected railing	1.2 m high temporary railing
	Elevator wellhead	Landscape	$W < 2 \times F$ and W < 10 m; No gate	Set up a safety net
	Other holes	Landscape	Same as reserved holes	Same as reserved holes

L: the length of the hole; H: the lower edge of the hole to the floor or the bottom; V: the side drop; W: the safety net spacing in the elevator wellhead; F: the floor height.

2.3.2. Building a Safety Rule Library Based on Ontology

(1) Semantic Web Rule Language (SWRL).

The semantic web rule language (SWRL) is a language that presents rules in a semantic way. SWRL integrates the ontology description language OWL Lite, OWL DL (description logic), and Datalog RuleML. SWRL is human-readable with the advantages of the rule editing function combined with the elements defined in the ontology to improve knowledge expression and reasoning [24]. The SWRL consists of 4 parts: Imp, Atom, Variable, and Building, as shown in Figure 5 [25].
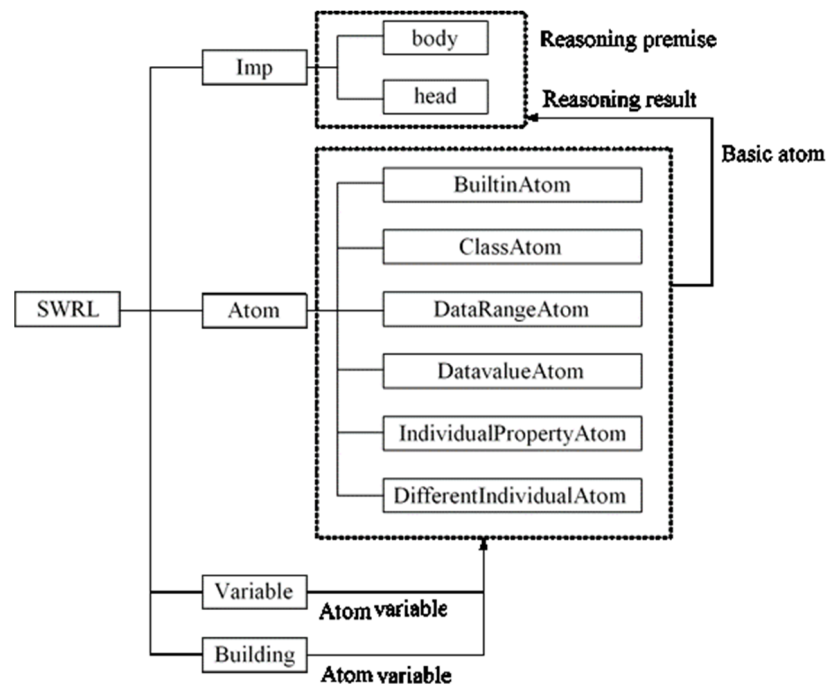
Figure 5. SWRL framework.

With the help of the developed safety management ontology classes and attributes, along with its rich arithmetic operations, the SWRL rules on safety management were established. To better connect the ontology and SWRL rules, the SWRLTab plug-in provided in the Protégé ontology editing software seamlessly connects the SWRL rule environment and the ontology editing environment, further improving the convenience of SWRL rule editing.

(2) Establishment of the SWRL safety rule database.

The ontology established in this paper included 4 categories: construction activities; precursor information; risk accidents; and solutions. The links formed by these 4 categories lay the premise for the establishment of SWRL rules. The elements (Atom) used in a complete SWRL rule for accident subjects, attributes, discriminant parameters, and prevention measures came from the 4 major classes of ontology, and the specific discriminant conditions or parameters were obtained in the interpretation of the safety rules. Taking the unprotected wall holes as an example, they were caused by the construction work brick masonry in the paving ash masonry and other construction steps for the outer window and other reserved holes. If the bottom edge of the hole was more than 2 m wide and the distance from the bottom of the floor was more than 80 cm without proper protection, accidents such as falling from a height may occur. The solution was to add a 1.2 m high protective railing around the hole in accordance with the specifications. The corresponding rule in SWRL database can be expressed as shown in Figure 6. The results of the complete set of safety rules for fall from height in the Protégé plugin SWRLTab developed in this study are shown in Figure 7. This paper also builds an SWRL safety rule base containing 5 major incident identifications in accordance with this process, as shown in Figure 8.
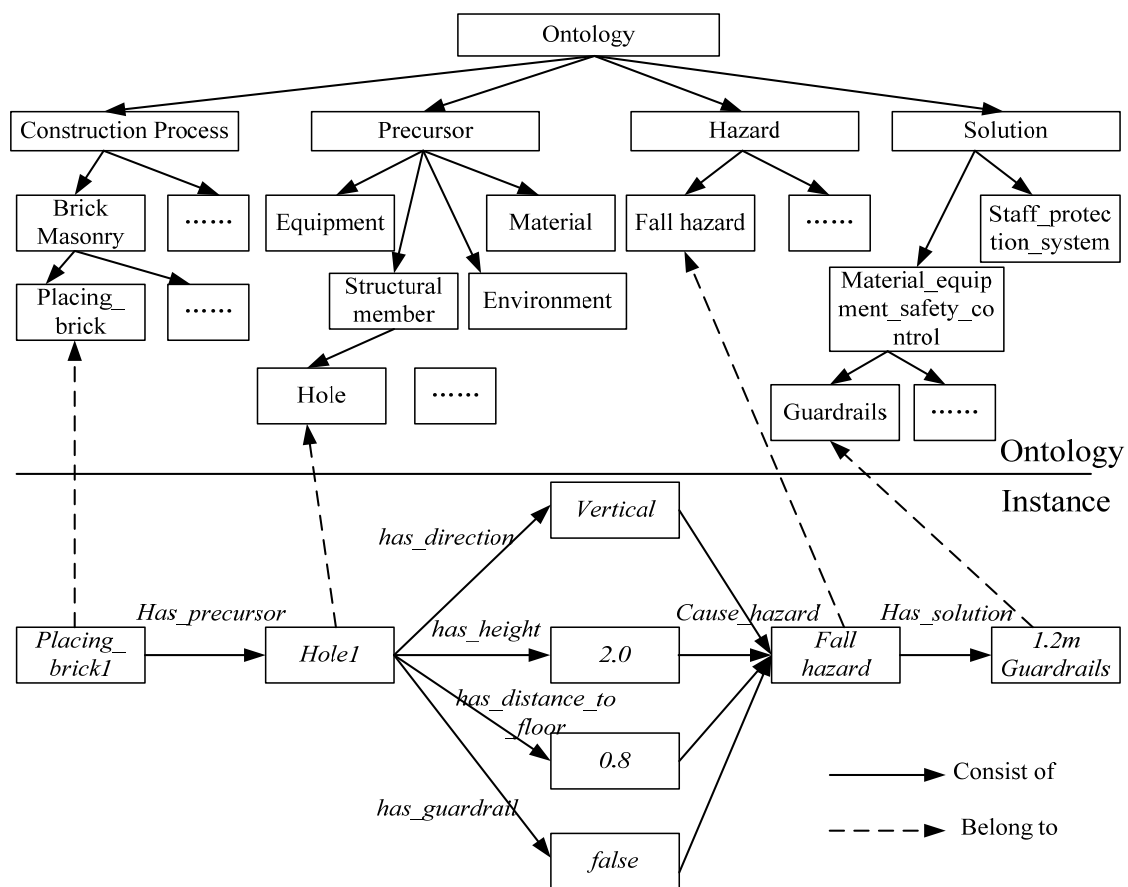
Figure 6. Vertical hole rule frame.

Figure 7. Vertical hole and cutting fire instance rules.

Figure 8. SWRL safety rule library for major safety risks.

2.4. Safety Risk Identification and Retrieval Mechanism

Safety rules were composed of accident subjects, judgment conditions, risks, and solutions, etc. The implementation of safety rules inspection needs to link the subjects of accidents in construction activities and extract discriminant parameters for verification. As a data source, the BIM model has its parameterized characteristics; thus that the information of each component and the relationship between each other can be directly attached to the component, which provides better convenience for the extraction of relevant information for a safety inspection. In addition, due to the characteristics of parameterization, modification or model building can be achieved directly by adjusting the parameters. This convenient way can save time and labor for safety managers on construction sites.

The safety risks in the construction process are dynamic and may only exist in a certain stage or process. For example, when the construction of the floor slab is completed, the unprotected periphery of the building forms edges that may cause the risk of falling from a height; and after the building of the walls eliminates the risk of the edges, the vertical opening of the windows on the masonry creates a new risk of falling from the height of the opening. Therefore, the process of safety inspection is not only the inspection of static safety risks but also includes more practical dynamic risk identification. BIM can realize the dynamic simulation of the construction process according to the given model and progress information (4D BIM) and serve as an intuitive process information hub. The dynamic risk inspection in the everyday construction process is only made possible by automatic identification and prevention of safety risks through the integration of BIM, ontology, and NLP. In light of the complexity of information related to safety risk, it was also necessary that safety managers can obtain access to the information in a convenient manner. Information related to safety management checked by the safety rules, such as accident subject/component attribute information, location information, time, over-limit parameters, etc., should be extracted as a safety problem report as required for the use of the safety managers.

2.4.1. Safety Risk Identification

The safety risk identification contains the following processes: first, use the BIM software tool to build the construction plan and schedule into a virtual construction process, and assign the necessary parameters and attribute information to it according to the characteristics of the construction components. Then, the ontology established a link with the detected object or construction process in BIM, automatically established an instance name in the safety ontology class according to the subject exposed to safety risks and extracted the required discrimination parameters to the ontology instance. Next, use the established SWRL rules to complete the inspection of safety risks and identify unsafe components and their parameters. Finally, all safety risks were summarized and output as an inspection report and displayed in the visual 3D model according to the inquiry made by construction safety managers. Considering that buildings are becoming more complex, and the content of the safety inspection report is complicated or cumbersome, safety managers still need to spend a lot of time finding safety management issues at different stages. For this reason, this paper uses the developed natural language search engine to achieve intelligence retrieval, which provides a convenient way for safety managers to deal with safety issues under their jurisdiction.

When the inspected object is a specific building component, the information of the component in the BIM model includes component name, attribute parameters, and other ancillary information, and the attribute parameters include direct attribute parameters such as elevation coordinate parameters, geometric information, material parameters, etc., and indirect parameters such as the topological relationship with the surrounding space. In building components such as stairs and elevator shafts, the names of building components have a corresponding relationship with the subject of the accident. This provides a preliminary way to link components with safety rules. After the connection was generated, multiple instance names were created by the scanning of the linked object, and

the corresponding parameters or attributes were extracted to the instance to complete the instance creation. When the detected object was an abstract construction process, the link between the accident and the construction process was established through the name of the construction process, which was then used to detect related attributes or parameters of components or materials within a certain distance of the process, thus as to create an instance for safety rule checking. The process is shown in Figure 9.
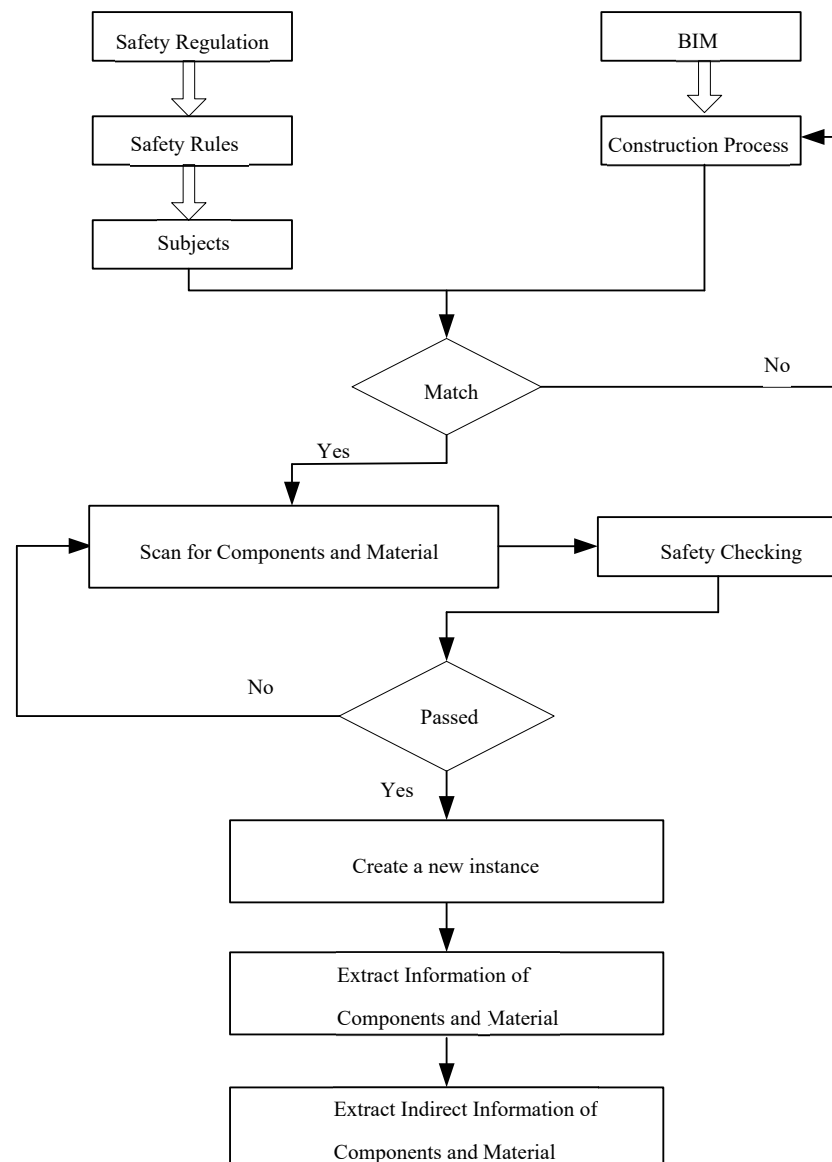

Figure 9. Linking building components and construction processes to safety rules.

When given specific domain rules and facts, the rule-based reasoning engine can be used to complete the corresponding reasoning tasks. However, the safety rules expressed by the SWRL rule language and the fact knowledge expressed in the form of ontology cannot be understood by JESS; thus, it was necessary to convert the SWRL rules and facts into the knowledge format supported by JESS to realize the reasoning. When the JESS inference engine is activated, the inference process begins. Pattern matching was used to determine what rules were executed and when to execute, after which the Agenda executes the commands in the activated rules. Then, the execution engine completes the execution of the rules. The processes are illustrated in Figure 10, which are then implemented in the case study.
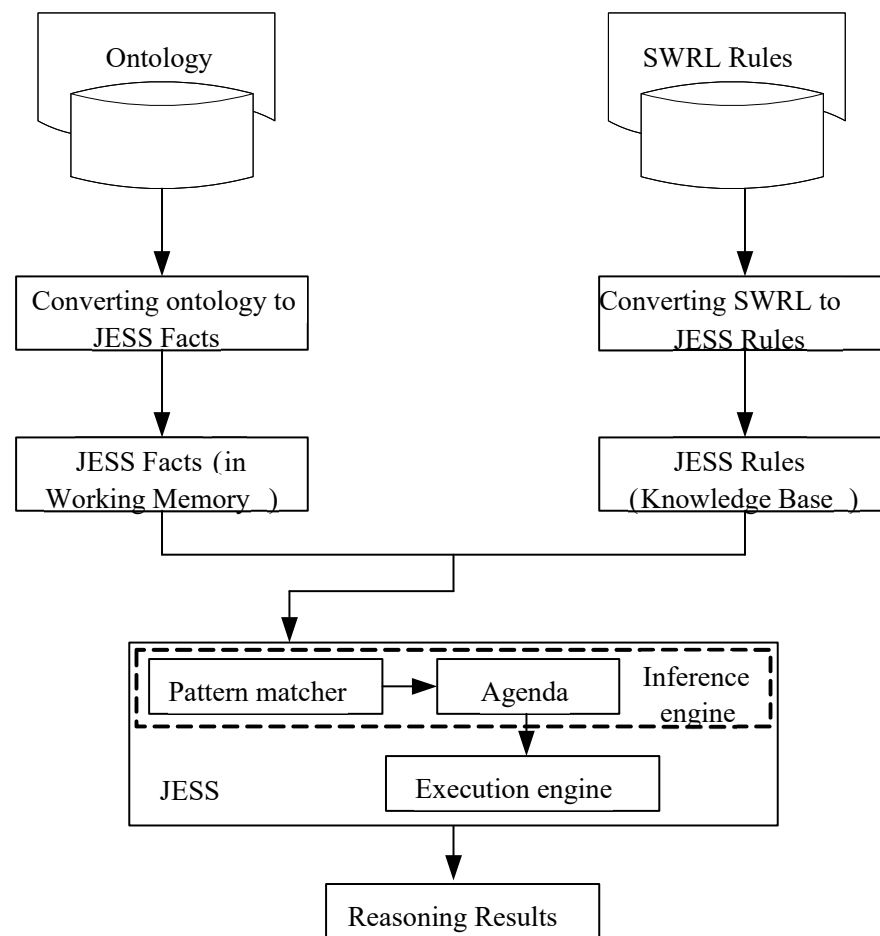
Figure 10. The reasoning processes.

2.4.2. Retrieval Mechanism

In contrast to traditional keyword-based retrieval technology, NLP-based retrieval technology was used to understand the semantics of the user's retrieval statement, which results in more accurate retrieval and allows safety managers to obtain construction safety information in a convenient and efficient manner. The retrieval mechanism designed in this paper included word segmentation, part-of-speech tagging, parsing, target keywords, and constraint sequence expansion.

(1) Word segmentation, in which the natural sentence from the user is divided into individual words after identifying each word and punctuation mark;

(2) Part of speech tagging, by marking the part of speech of each word based on statistical data or other methods, such as the Penn Treebank POS tag set;

(3) Parsing, designed to capture the relationship between words, forms a syntactic tree of queries through the Stanford parser. The NLP tool was used to obtain the relationship of different words in the sentence to form a syntactic structure tree. The results of the syntax analysis are shown in Figure 11, where IN, ADJ, NN, NNS, NP, and PP represent prepositions, adjectives, singular nouns or uncountable nouns, plural nouns, noun phrases, and prepositional phrases, respectively.
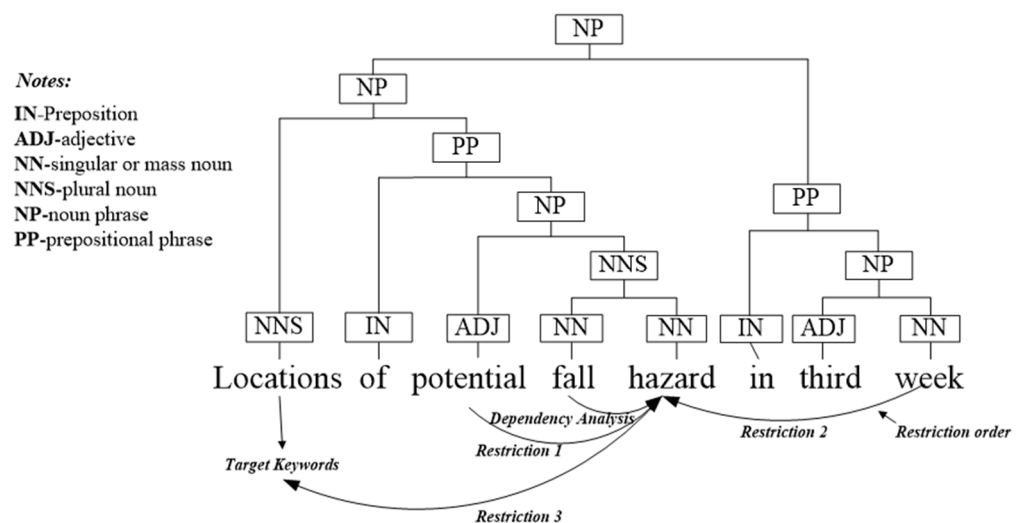
Figure 11. Syntactic structure tree.

The purpose of obtaining the syntactic tree was to find the words that actually represent the user's intent to search (also known as "target word") and the constraint relationship, while the keywords that contain the user's search purpose often exist in the noun phrase. Firstly, starting from the root node of the tree, create a line containing only nouns or noun phrases. The noun phrase closest to the root node was the target keyword, and the non-noun node (preposition phrase or adjective phrase with the same root node) was the constraint keyword of the target keyword. Secondly, for the subtree of the structure tree (the subtree root node is a noun or noun phrase), the leaf whose noun (noun phrase) was in the leaf node represents the target keyword of the subtree, and a similar case holds for the non-identical parent node. The noun node (prepositional phrase or adjective phrase) was the constraint word of the target keyword of the subtree. Then, a reverse analysis of the subtree to the root node was used to obtain a constraint sequence for the target keyword. According to the syntax tree analysis in the example sentence, the keywords "locations," "week," and the case of "fall" and "hazard" for the 2 keywords with the same root node were nouns. Finally, dependency grammar analysis was required to determine which word was a keyword and which word was a constraint word. After completing the analysis of this situation, the keywords of the whole sentence were "locations," "week," "hazard," and according to the order of constraints, "locations" was the keyword of the target (representing the purpose of the user). The user wants to obtain the location of the risk of falling from a height, and the risk of falling from a height was based on the third week of the project schedule. Therefore, the other keywords were all the constraints of this keyword and have a sequence (Restriction 1, Restriction 2).

(4) Keyword extension and mapping: since the query statement is expressed in natural language; the keywords or constraints of non-professional expression cannot be used for database retrieval. Therefore, this section designs the concept extension and mapping using ontology. This process consists of 3 tasks: (1) formalization of the extracted keywords or constraints; (2) query expansion using semantic similarity techniques to determine the final query term; and (3) mapping to the database.

Formal standardization of extracted keywords or constraints was used because the same term may have different expressions, such as abbreviations, spellings, or singular and plural numbers, such as "locations" becoming "location." Since the diversity of expressions leads to different expressions with the same meaning, the standardization of forms also includes conceptualization. For example, the concept of the use of the expression "location" in the ontology was "coordination," and the expression of risk was "hazard" instead of "risk." Therefore, formal standardization also includes the transformation of keywords or constraint sequence words obtained by search statements into standard concepts defined in the ontology, which were all integrated in the Protégé software.

The query expansion algorithm was used to improve the recall rate and retrieve the results that more strongly satisfy the user's needs. This paper determines the concept of extension based on the degree of approximation between ontology concepts. This process is called ontology mapping, which is the process of mapping to exchange information in a semantically sound manner. The target words obtained in the previous step were used to calculate semantic similarity of other words according to the concept hierarchy of the ontology (such as the parent class and subclass), which was then used as an expansion of query. This paper chose to use the Leacock–Chodorow formula [26] (Formula (1)) to characterize the degree of conceptual semantic similarity.

$$R(C, C_i) = -lg\frac{len(C, C_i)}{2Depth} \qquad (1)$$

where $R(C, C_i)$ represents the similarity value; $len(C, C_i)$ represents the conceptual standard concept of the target keyword and the constraint keyword; C and the extended concept C_i represent the shortest distance in the construction safety ontology structure; $Depth$ is the maximum depth of the construction safety ontology; and the coefficient 2 was used to ensure a positive value of similarity.

To remove the less similar concepts, this paper sets the similarity threshold to filter the concepts of lower similarity value, and the remaining expansion concepts and the initial standard concept were used as the final search keywords to complete matching with the Database. The final search keyword was used to match the database and sort the search results. Each target keyword and constraint sequence, extended keyword, and constraint sequence form a query vector and a record formed by the database record to calculate the similarity value, and the result was presented according to the similarity value. This paper used the Cosine space method to determine the size of the similarity value, as shown in Formula (2).

$$V = \frac{(X \cdot Y)}{\| X \| \| Y \|} \qquad (2)$$

where

V: cosine of similar distance.

X: the feature vector formed by the target keyword or the extended keyword and its constraint sequence.

Y: the feature vector of the database data record.

2.4.3. Report of Retrieval Results

At present, the most common data presentation methods included 2-dimensional charts, network diagrams, documents, pictures, and animations, and this paper selected the document report + visual map in BIM to present the search results. The safety inspection report designed in this paper included the construction process, construction procedure, planning time, construction location (area, floor, etc., location map), type of accident, over-limit parameters, and recommended solutions. This information needs to be obtained from the database, and the information was presented according to the intention of the search phrase from the user.

3. Validation

3.1. Framework of the Proposed Safety Checking System

This paper chose the student dormitory building (C3 building) of the commuter college of a university as a case to demonstrate the framework. The student dormitory building (C3) covers an area of approximately 827.2 square meters. It extends seven floors above the ground and one floor below the ground. Each story height was 3.3 m, and its basement height was 3.6 m. The three-dimensional model of the dormitory building, which was generated in the Revit software (license number: 559-06926929), is shown in Figure 12 below.
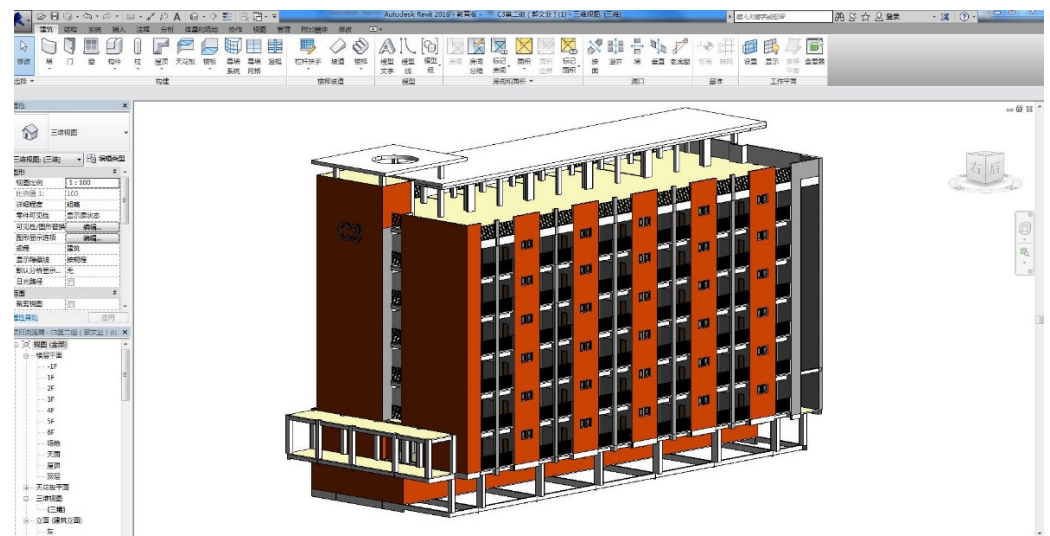
Figure 12. BIM model of a student dormitory building.

To realize the dynamic safety inspection process, this paper first used Microsoft Project (license number: W2JJW-4KYDP-2YMKW-FX36H-QYVD8) to develop a progress plan of simulated construction, as shown in Figure 13. The concrete pouring process was divided into structure construction and decoration construction. The schedule began on 1 July 2017, when concrete was poured into the basement beams and slabs. On 23 September 2017, the installation of the top doors and railings was completed. The total duration was 85 days.
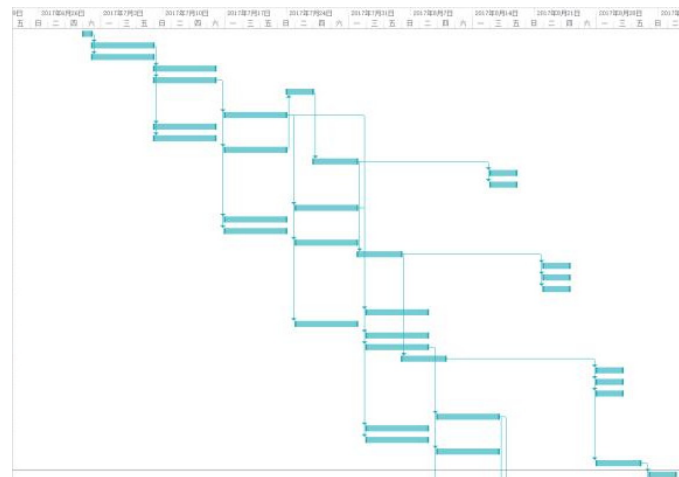

Figure 13. Construction schedule of the student dormitory building.

This paper then integrated the Revit information model of student dormitory building and MS project schedule into Navisworks software (license number: 559-06926929) and realized the dynamic simulation process by dividing the construction process and connecting time information of components in the construction process, as shown in Figures 14–17.
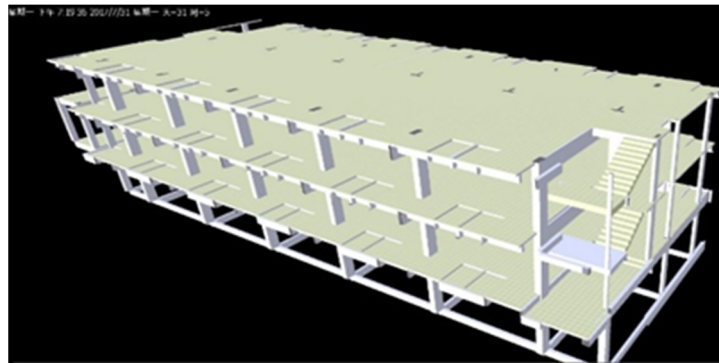
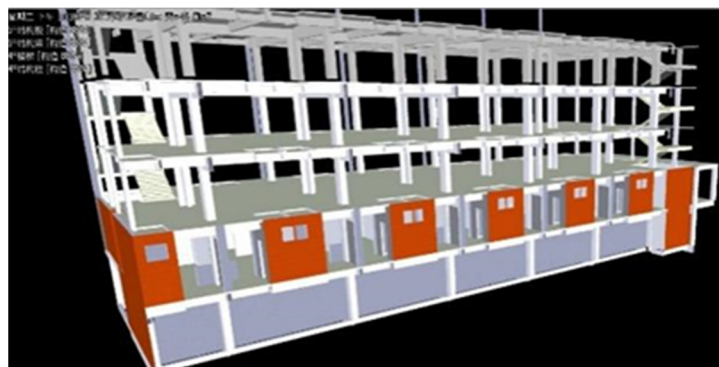**Figure 14.** Project progress in week 4.

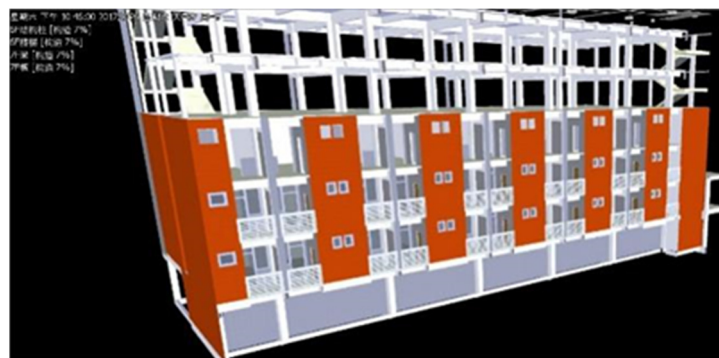**Figure 15.** Project progress in week 7.

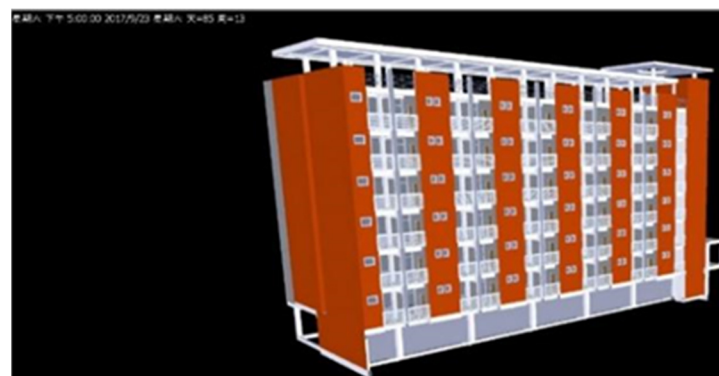**Figure 16.** Progress of the project in week 8.

Figure 17. Progress of the project in week 13.

The following is an example of the process and results of the performed safety inspection on holes. According to the construction schedule, the time period for the top floor to complete the floor and stairs was 10 weeks (26 August to 2 September 2017), while the maintenance wall of the top floor took 12 weeks (15 September to 19 September), and the opening formed by the stairs and slabs had not undergone corresponding protection measures during the period of 10 to 12 weeks (Figure 18). This condition may lead to an accident involving falling from a high place.

The automatic identification process was performed by scanning the construction process files containing time information by the developed framework, which established the connection with the top-level components according to the detection rules of the openings and extracted the relevant parameters. The results from the reasoning are shown in Figure 19. Under the implementation of a reasoning engine, the safety parameters can be identified automatically, and the risks and solutions can be identified (Figure 20).
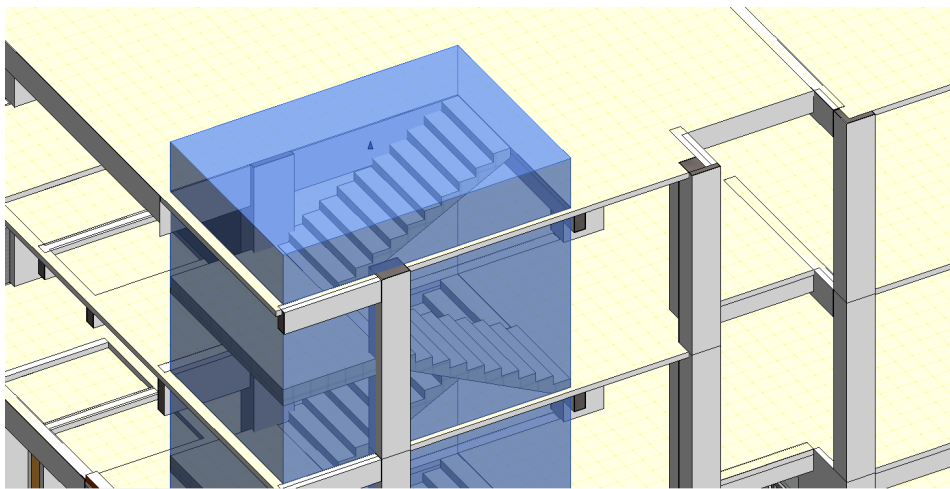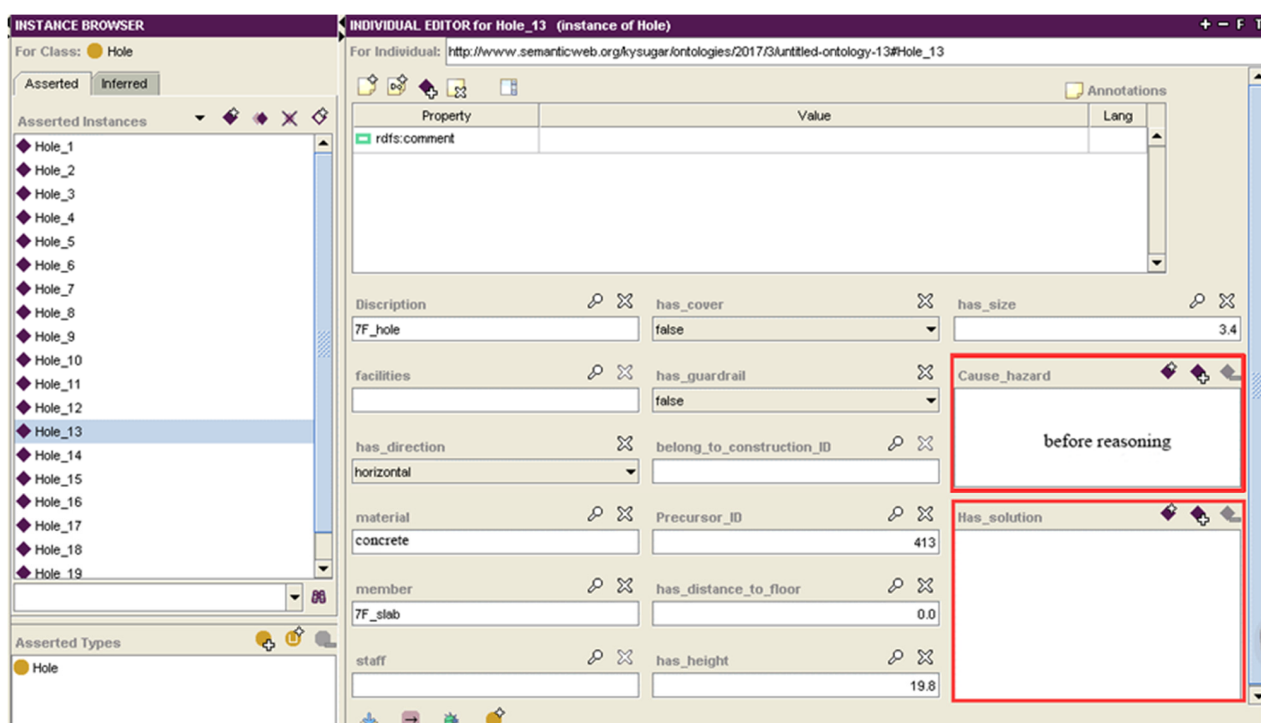

Figure 18. Top floor stair opening.

Figure 19. Extraction of the top floor stair opening information to the ontology instance.
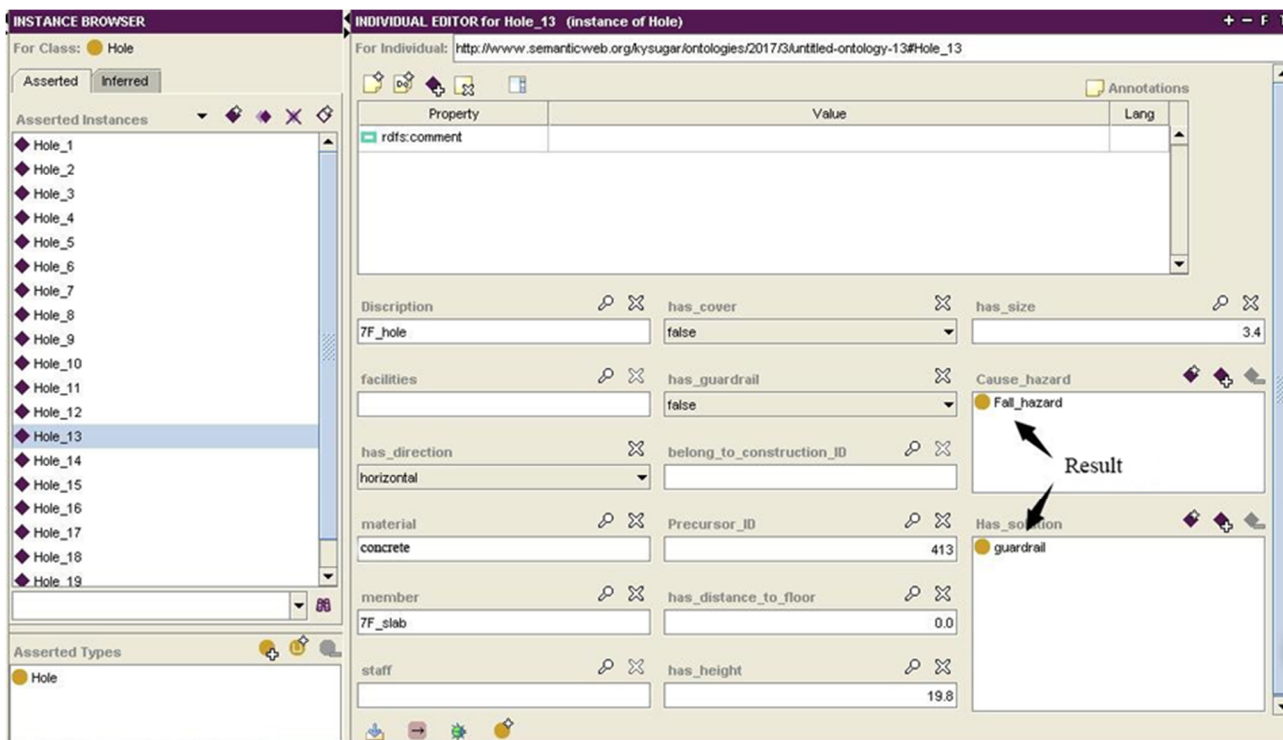
Figure 20. Risk reasoning results of the top floor stair opening.

The NLP-based query system to present the identification results of related risks according to the user's requirements was then illustrated. The login interface and the query interface are shown in Figures 21 and 22, respectively. This paper takes the query statement "I need to know about any potential hazards on the third floor" as an example to show the query results (Figures 23 and 24) and query reports (Figure 25). The project manager of the case study stated that the developed plugin was user-friendly, and the results could help improve the construction safety management since they were rational according to their experiences and requirements. Similar to many other industries, people tend to use a system that can make their queries easily without previous knowledge of semantic technologies [27].
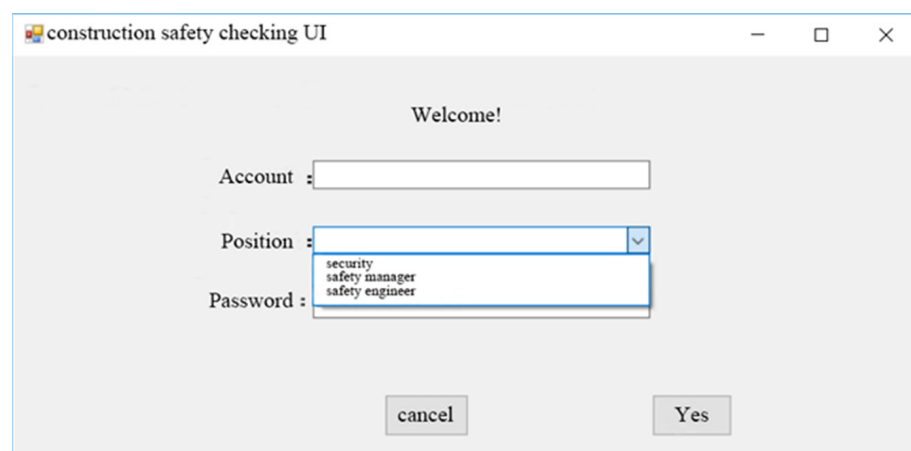

Figure 21. Landing interface of the developed safety risk inspection system.

Figure 22. Landing interface of the safety risk inspection system in the construction process.

Figure 23. Safety risk query result for potential risks on the third floor.

Figure 24. Safety risk query result and the suggested safety measures.

Figure 25. Query report.

3.2. Experiment Results

In the experiment, the developed system presents the safety inspection results of the third floor. It was shown that the system successfully identified the safety risks in the construction process of the third floor, such as the stairway opening and window opening, as well as the multiple adjacent edges. The reasoning accuracy by virtue of the ontology was close to 100%. For the evaluation of retrieval effectiveness, this paper adopted two indicators: recall rate and precision rate. Suppose A was the correct number of results retrieved by a certain retrieval sentence, B was the number of incorrect results retrieved by a certain retrieval sentence, and C was relevant results that were not retrieved by a certain search sentence; thus, the precision rate can be expressed by A/(A + B), and the recall rate can be expressed by A/(A + C).

However, these two indicators have a mutual restriction relationship; that was, if the recall rate was required to be high, the search result would contain more wrong results; thus, the accuracy rate would be reduced, and if the accuracy rate was high, the search result will miss a certain correct result. For this reason, an F-value metric that takes the two into account was introduced to consider the retrieval level of retrieval technology. The Formula (3) is as follows.

$$F_\beta = \left(1 + \beta^2\right) \cdot \frac{Pre \cdot Re}{\beta^2 \cdot Pre + Re} \tag{3}$$

where *Pre* stands for accuracy;

Re stands for recall rate; and

β Represents the weight of recall relative to accuracy;

This article randomly tested about 50 natural language search sentences and made statistics on accuracy and recall. The results are shown in Figure 26. According to the chart, the accuracy rate of the retrieval system was higher than 60% when the recall rate was lower than 50%.
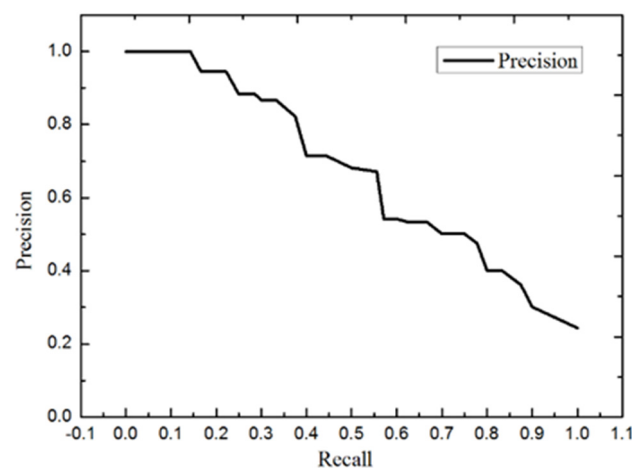
Figure 26. Retrieval recalls and precision curves.

4. Conclusions and Future Work

This paper built an automatic mechanism for identifying and preventing safety risks during the construction process by integrating the safety rules library and BIM technology with advantages in visualization and parameterization. Moreover, an intelligent presentation method based on the Natural Language Processing technology was proposed to intelligently present identified results from a developed system.

In the building of the ontology, the hierarchical construction activities of residential buildings were broken down by using WBS. Moreover, then the precursory information of four categories, including structural components, materials, equipment, and environment, were established by combining construction activities and types of safety accidents. Then, domain ontology was built depending on the seven-step method, which consists of con-

struction activities, precursor, hazard, and solution. Next, on the basis of the established precursor and relevant codes, the parameterized forms for five kinds of accident safety rules were summarized and converted into SWRL format. To realize the automatic safety rules checking, automatic checking mechanism of integrating BIM, and safety rules was proposed. Meanwhile, an intelligent results presentation based on Natural Language Processing was designed. Finally, the construction process safety rule checking system was developed, and the effectiveness of the system was verified by a case study.

Two innovative methods were proposed in this paper: (1) ontology technology and BIM were put forward for safety rule checking for the dynamic construction process of the residential building; (2) the intelligent presentation mechanism based on the Natural Language Processing was presented. It not only provided ideas for the reuse and sharing of the safety knowledge but also proposed a solution for immediate and intelligent presentation of the safety problems on dynamic problems, which made a certain contribution to the improvement of efficiency of construction safety. The NLP-based presentation method for construction safety rule checking results will facilitate the access of construction safety managers to the relevant information, which will then improve the efficiency of construction safety management in practice. However, there are certain limitations of the paper that needs to be addressed in future work: (1) this paper only aimed at the construction process of civil residential buildings to establish the construction safety ontology, it is not comprehensive and extensive enough; (2) this paper selects only common safety regulations and clauses.

Author Contributions: S.W. and Y.D. were in charge of the conceptualization, methodology and formal analysis under the supervision of Q.S., H.D. and J.C.P.C. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Yan, W. *Construction Hazards Management Research Based on BIM*; Shanghai Institute of Technology: Shanghai, China, 2016.
2. Luo, C.; Xie, X. Comparison study of accident-causing theories. *J. Saf. Sci. Technol.* **2007**, *3*, 111–115. Available online: https://en.cnki.com.cn/Article_en/CJFDTotal-LDBK200705028.htm (accessed on 3 November 2021).
3. Zhi, P. *A Research in the Safety Control. of Reverse Construction*; Tianjin University: Tianjin, China, 2005.
4. Liu, G.; Lei, L. Analysis of heinrich accident-causing theory and the factors of safety ideology. *Saf. Environ. Eng.* **2013**, *20*, 138–142. Available online: https://en.cnki.com.cn/Article_en/CJFDTotal-KTAQ201301032.htm (accessed on 3 November 2021).
5. Bird, F. A handbook of personnel management practice: Michael Armstrong, 2nd edition, Kogan Page, London, 1984, £9.95 paperback. *Int. J. Hosp. Manag.* **1985**, *4*, 48. [CrossRef]
6. Li, J.; Su, X.; Qian, P. The methodology of developing domain ontology. *Comput. Agric.* **2003**, *12*, 7–10. Available online: http://en.cnki.com.cn/Article_en/CJFDTOTAL-JSJN200307002.htm (accessed on 3 November 2021).
7. Wang, H.; Boukamp, F. Ontology-based representation and reasoning framework for supporting job hazard analysis. *J. Comput. Civ. Eng.* **2011**, *25*, 442–456. [CrossRef]
8. Lu, Y.; Li, Q.; Zhou, Z.; Deng, Y. Ontology-based knowledge modeling for automated construction safety checking. *Saf. Sci.* **2015**, *79*, 11–18. [CrossRef]
9. Zhang, S.; Boukamp, F.; Teizer, J. Ontology-based semantic modeling of construction safety knowledge: Towards automated safety planning for job hazard analysis (JHA). *J. Comput. Civ. Eng.* **2015**, *52*, 29–41. [CrossRef]

10. Ding, L.; Zhong, B.T.; Luo, H. Construction risk knowledge management in BIM using ontology and semantic web technology. *Saf. Sci.* **2016**, *87*, 202–213. [CrossRef]

11. Zhang, J.; EI-Gohary, N. Automated information transformation for automated regulatory compliance checking in construction. *J. Comput. Civ. Eng.* **2015**, *29*, B4015001. [CrossRef]

12. Zhang, J.; EI-Gohary, N. Extending building information models semiautomatically using semantic natural language processing techniques. *J. Comput. Civ. Eng.* **2016**, *30*, C4016004. [CrossRef]

13. Zhang, J.; EI-Gohary, N. Integrating semantic NLP and logic reasoning into a unified system for fully-automated code checking. *Autom. Constr.* **2017**, *73*, 45–57. [CrossRef]

14. Lin, J.; Hu, Z.; Zhang, J.; Yu, F. A natural-language-based approach to intelligent data retrieval and representation for cloud BIM. *Comput. Aided Civ. Infrastruct. Eng.* **2016**, *31*, 18–33. [CrossRef]

15. Jiao, H. *Modeling and Application of Ontology-Based Knowledge Base for Risks of Subway Construction*; Southeast University: Nanjing, China, 2015.

16. Zeng, W. *Pre-Control Study of Building Construction Safety Accident Based on DFS*; Southeast University: Nanjing, China, 2015.

17. Zhang, J. *Research on Ontology-Based Retrieval of Construction Domain under BIM Environment*; Dalian University of Technology: Dalian, China, 2013.

18. Ministry of Housing and Urban-Rural Development of the People's Republic of China. *Construction Safety Inspection Standards (JGJ 59-2011)*; China Building Industry Press: Beijing, China, 2012.

19. Ministry of Housing and Urban-Rural Development of the People's Republic of China. *Technical Specifications for Safe Operation at Height during Construction (JGJ 80-91)*; China Planning Press: Beijing, China, 1992.

20. Ministry of Housing and Urban-Rural Development of the People's Republic of China. *Technical Regulations for Safety in the Use of Construction Machinery (JGJ 33-2012)*; China Building Industry Press: Beijing, China, 2013.

21. Ministry of Housing and Urban-Rural Development of the People's Republic of China. *Technical Specifications for Temporary Electricity Safety at construction Sites (JGJ 46-2005)*; China Building Industry Press: Beijing, China, 2010.

22. Ministry of Housing and Urban-Rural Development of the People's Republic of China, General Administration of quality supervision, inspection and quarantine of the People's Republic of China. *Technical Specifications for Fire Safety on Construction Site of Construction Projects (gb50720-2011)*; China Planning Press: Beijing, China, 2011.

23. *GB8624-2012*; State General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, China National Standardization Administration Committee Classification of Combustion Performance of Building Materials and Products. China Standard Publishing House: Beijing, China, 2012.

24. Yu, S. *Research on SWRL-Based Semantic Relevant Discovery Based Ontology Mapping Approach*; Nanjing University of Aeronautics and Astronautics: Nanjing, China, 2009.

25. Horrocks, I. SWRL—A Semantic Web Rule Language Combining OWL and RuleML. 2003. Available online: http://www.daml.org/2003/11/swrl (accessed on 3 November 2021).

26. Leacock, C.; Chodorow, M. *Combining Local Context and Wordnet Similarity for Word Sense Identification*, 1st ed.; MIT Press: Cambridge, MA, USA, 1998.

27. Rosales-Huamaní, J.A.; Castillo-Sequera, J.L.; Montalvan-Figueroa, J.; Andrade-Choque, J. A Prototype of Speech Interface Based on the Google Cloud Platform to Access a Semantic Website. *Symmetry* **2018**, *10*, 268. [CrossRef]