Article

Architecture and Application of Traffic Safety Management Knowledge Graph Based on Neo4j

Danling Yuan *, Keping Zhou and Chun Yang

School of Resources and Safety Engineering, Central South University, Changsha 410083, China;
kepingzhou@csu.edu.cn (K.Z.); chunyang@csu.edu.cn (C.Y.)
* Correspondence: 205511066@csu.edu.cn; Tel.: +86-18707494130

Abstract: A large amount of traffic safety information has been generated. This will further promote the sustainable development of transport. However, its content, form, and structure are complex and scattered, lacking effective information integration and a comprehensive framework. Combined with the concept of safety analysis, a traffic safety management knowledge graph was designed for structured data, which include 54 types of node entities and 14 types of relationship entities. Six types of information were collected and imported, including illegal acts, vehicle failure, emergency response, legal norms, organization information, and road-related information. Ultimately, a knowledge query function was realized using Cypher, and an automatic Q&A function was created based on rule matching. A traffic accident knowledge graph was constructed for unstructured data, with people and institutions involved, vehicles involved, and accidents as the core, including 21 types of node entities and 22 types of relationship entities. Comparing the node entity extraction performance of Bert, Bert-CRF, Bert-BiLSTM, and Bert-BiLSTM-CRF models, Bert BiLSTM-CRF performs the best. The Bert model was used for relationship entity extraction. The traffic accident knowledge graph can structurally display accident information and support a query function to facilitate safety analysis.

Keywords: traffic; safety; knowledge graph; management; Neo4j

1. Introduction

Transport has a significant impact on energy consumption and greenhouse gas emissions. Traffic safety management is beneficial to reducing the impact of transport on energy consumption and the environment. At the same time, new technologies such as big data, the Internet, and artificial intelligence are deeply integrated with the transportation industry, providing unprecedented development opportunities for sustainable transport. It is necessary to construct a new type of traffic safety ecosystem based on traffic safety management knowledge graphs, which can enhance management capability, resist disasters and risks, and improve the efficiency of transportation networks.

The method of storing information using computers has gradually developed. In the 1960s, semantic networks appeared, which express human knowledge structure in the form of networks. M. Ross Quillian proposed a theory of computer formatting, organizing, and using human memory for semantics and concepts [1]. In addition, he proposed the Teachable Language Comprehender, aiming at understanding English texts [2]. From the 1970s to the 1980s, an object-oriented database was established to handle the core features of the data. Stanford University developed the MYCIN system, which helps diagnose possible factors in patients' infections based on rules and recommend the best treatment plan for patients [3]. Japan launched the "Five Generation Machine Plan" to achieve problem solving and reasoning, knowledge base management, and the creation of intelligent interfaces. Prolog emerged for non-numerical computing, which is an efficient descriptive language. The functions of reasoning and problem solving are gradually being achieved. Since the 1990s, knowledge expression based on databases has become an

inevitable demand in an information society. Virtual knowledge bases include CYC [4], WordNet [5], How Net [6], etc. How Net is a large-scale language knowledge base annotated by Dong Zhendong and Dong Qiang, mainly targeting Chinese vocabulary and concepts. A knowledge graph is a useful tool to save data in a more understandable way. A knowledge graph is essentially a semantic network that is used to reveal the relationships between all things. It describes entities and their relationships in the objective world in a structured way and expresses information in a form closer to how human cognition processes the world. Based on knowledge graphs, functions such as search optimization, knowledge Q&A, text classification, and sentiment analysis can be achieved, reducing the reading burden in daily work. They provide reliable and comprehensive information for application in professional fields. Famous knowledge graphs include Wikipedia, DBpedia, Freebase, Google Knowledge Graph, YAGO, etc. Knowledge graphs have been applied in multiple fields such as construction [7,8], chemistry [9,10], medicine [11,12], and transport [13]. In terms of traffic safety, the most important sources of texts are traffic accident reports. For traffic accidents, L.Y. Zhang [14] stores structured traffic accident case data in Neo4j, achieving accident profiling, classification, statistical, and associated path query functions. C. Liu et al. [15] extracted information from UK railway accident reports, constructed a knowledge graph to elucidate the risk factors and propagation characteristics that affect railway safety, and applied them to railway risk identification and assessment. J.T. Liu [16] et al. proposed a method to explore railway operation accidents based on knowledge graphs, which reveals the potential laws underlying accidents and hazards in heterogeneous networks.

Security models have become more complex and need to realize multi-dimension analysis. However, in practical work, they are difficult to apply. The performances of models often depends on personal understanding of safety management work. Designing and constructing knowledge graphs can better solve practical problems and integrate information. It has become a new development requirement to establish a highly systematic database centered on the work content of security managers and decision makers, which can improve the efficiency of security management.

2. Overview of Knowledge Graph Construction

The construction sequence of a knowledge graph can be top-down, bottom-up, or a combination of the two. Its hierarchical structure can be divided into the schema layer, data layer, and technical layer. The schema layer is the core of knowledge map construction, which focuses on the professional knowledge. As far as the road traffic safety knowledge graph is concerned, it is the integration and system construction of knowledge in the fields of road traffic and safety. The data layer contains multi-source data. Based on the data structure, data can be structured, semi-structured, and unstructured. It also can be divided into four types according to the form: text, images, video, and sound. The technical layer is the key to the effectiveness of the knowledge graph. Q. Liu [17] et al. construct a technical framework for each round of the knowledge graph iteration process, as shown in Figure 1.

Named-entity recognition and relationship extraction are key links in unstructured data processing. Named-entity recognition refers to the recognition of entities with specific meanings in texts. Traditional recognition tasks mainly include institution name, place name, person name, organization name, proper noun, etc. Entity extraction methods can be divided into rule-matching methods, statistical model methods, and deep learning methods. Goyal, A. et al. [18] summarize the developments made in named-entity recognition research. For low-resource named-entity recognition, M. S. Zhong et al. [19] construct a dual learning framework.
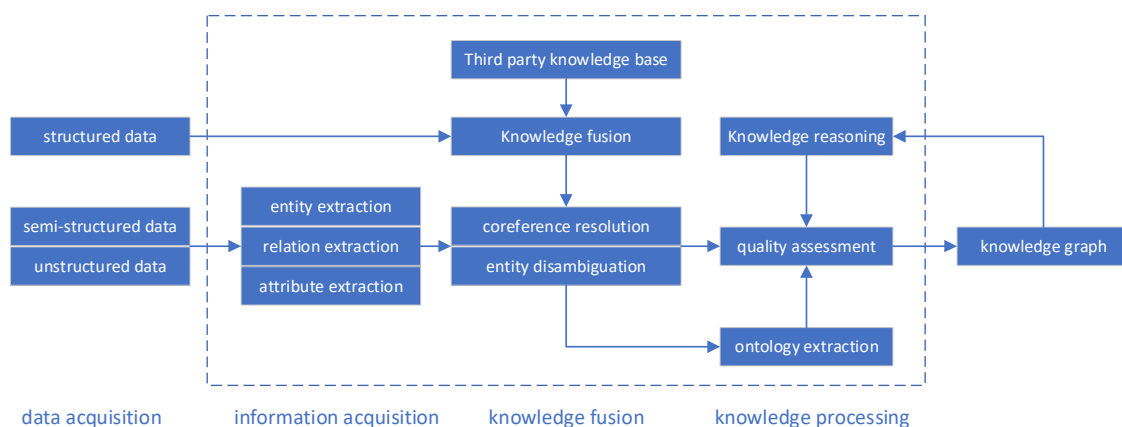
Figure 1. Process of knowledge graph construction.

Relationship extraction is the extraction of the SPO (Subject Predicate Object) triplet from a text. Relationship extraction requires identifying the subject and object in the text and determining the relationship type between them. Relationship extraction methods mainly have two categories: pipeline extraction and joint extraction. Pipeline extraction completes entity recognition and relationship classification separately. Different datasets and models can be used for learning, but errors in entity recognition accumulate into the relationship extraction task. It can be carried out by rule matching, semi-supervised learning, supervised learning, and other methods. Q. Liu et al. [20] summarize the relation extraction methods based on machine learning and deep learning. Joint extraction means that entities and relationships are extracted simultaneously, which can be based on parameter sharing or joint decoding. B. Qiao et al. [21] propose a joint extraction model based on BERT. T. T. Hang et al. [22] realize joint extraction by labeling a source–target entity.

Neo4j is a commonly used tool for storing knowledge graphs. It is a NoSQL database implemented in Java that supports ACID transactions, cluster, backup, and failover services. Neo4j builds a graph database from bottom to top that stores data in a freely adjacent graph structure. Compared with a relational database, its query method is linear and more agile. Nodes and relationships are fundamental elements of a graph. Nodes can contain multiple labels and attributes. Relationships can set direction while including labels and attributes, but they cannot connect empty values. Neo4j supports the use of Cypher statements for queries, allowing for modification, addition, and subtraction of data. As a declarative query language, Cypher has more expressive power.

3. Design and Construction of Traffic Safety Management Knowledge Graph

3.1. Design of Traffic Safety Management Knowledge Graph

According to the method of data processing, Section 3 mainly focuses on structured data disclosed by government apparatus, enterprises, and media and directly creates entities and their relationships in Neo4j using py2neo, without information extracting. Building a knowledge graph centered on traffic safety management work promotes traffic safety's sustainable development. Create four information modules: vehicle, road facility, emergency rescue, and management. Construct a total of 54 types of node entities and 14 types of relationship entities, as shown in Figure 2. To better serve traffic safety management work and ensure that the overall information system meets security requirements, the key design parameters are as follows:

(1) The setting of the entity label 'Organization Type' divides 21 types of organizations into 6 categories based on 6 nodes about institutional functions. It further refines and summarizes the information. Record important attributes including phone number, address, and operation scope for each organization. The 6 types of organizations are explained as follows.
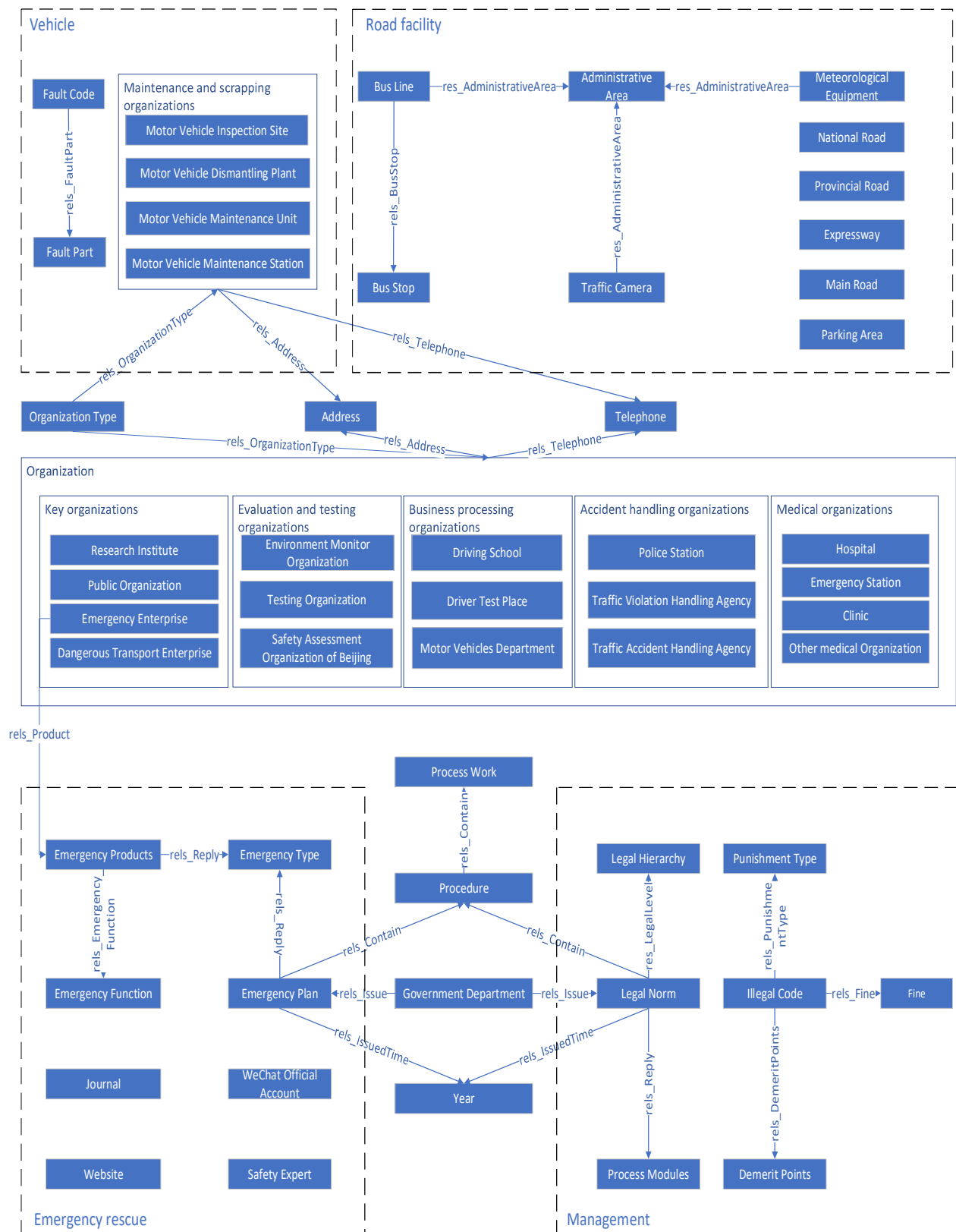
Figure 2. Knowledge graph framework of structured data.

Business processing organizations: (1) Driving schools are training organizations that teach traffic regulations and rules, driving safety knowledge, and driving skills. (2) Driving test locations are important to traffic safety and award driving qualifications. (3) Mo-

tor vehicle departments are mainly responsible for business related to motor vehicles and drivers, such as the registration, modification, and transfer of motor vehicles, as well as the issuing, replacement, and renewal of motor vehicle driving licenses.

Maintenance and scrapping organizations: (1) Motor vehicle inspection sites perform safety checks in different cycles based on the specifications and structures of the inspected vehicles. (2) Motor-vehicle-dismantling plants mainly recycle and dismantle various scrapped motor vehicles, sell useful materials, and process the paperwork of disposal. (3) Motor vehicle maintenance units and motor vehicle maintenance stations are mainly responsible for vehicle repair.

Accident-handling organizations: (1) Traffic-accident-handling agencies are operated by the traffic division of each administrative region, and they process business pertaining to accident and violation handling. (2) Traffic-violation-handling agencies mainly are the law enforcement stations of each administrative region's traffic division, responsible for off-site law enforcement and information consultation. (3) Police stations are responsible for investigating and collecting evidence and assisting traffic police in handling accidents.

Medical organizations: (1) Hospitals and clinics are key organizations for traffic accident rescue. (2) Other medical organizations include medical laboratories, health service centers, physical examination centers, and medical information centers. (3) Emergency stations mainly include first aid centers, first aid stations, and emergency medical rescue centers.

Evaluation and testing organizations: (1) Safety assessment organizations of Beijing provide services such as traffic safety facility quality inspection, geological disaster assessment, safety management consulting, and safety sign design. (2) Environmental monitoring organizations are responsible for ecological monitoring, protection, and awareness, as well as environmental information management, emission management, and hazardous solid waste management. (3) Testing organizations mainly include relevant national institutions.

Key organizations: (1) Research institutes provide theoretical and technical guidance for the development of traffic safety. (2) Public organizations are professional associations related to emergency rescue. (3) Dangerous transport enterprises are enterprises that transport hazardous chemicals as stipulated by law. (4) Emergency enterprises are important contacts in procurement work.

(2) The setting of the entity label 'Process Modules' can divide 63 laws and regulations into 7 process modules based on 7 nodes about managed objects. It can realize an effective query by facilitating searching for managed objects to find relevant laws. The managed objects of the 7 process modules are related to the legal texts as follows:

Vehicle management: The legal texts regulate the entire lifecycle of a vehicle, including registration, operation management, scrapping, and recycling. Staff management: The legal texts take the People's Police Law of the People's Republic of China as the core and provide detailed regulations on the work, supervision, internal affairs, dress, police rank system, and standardized management of public security administrations.

Motor vehicle driver management: The legal texts regulate the training and management of drivers, the issuing and usage of drivers' licenses, and the qualification management of drivers of specific vehicles.

Accident handling: The legal texts mainly discuss the handling procedures, work standards, normalized post and information processing of traffic accidents. In addition, they specify standards for on-site handling, accident identification, and results determination.

Law enforcement supervision: For traffic participants, outside institutions, and internal institutions, the legal texts are the basis for supervising law enforcement officers and protecting rights.

Duty: The legal texts are centered around traffic violations. There are detailed regulations for law enforcement personnel, typical illegal behaviors, and illegal handling.

System: The legal texts include the Road Traffic Safety Law of the People's Republic of China and the Regulation on the Implementation of the Road Traffic Safety Law of the

People's Republic of China. These two laws are the core of the legal text system for road traffic safety, defining multiple aspects of the direction of traffic safety management.

(3)   The setting of the entity label 'Procedure' contributes to clarifying the context of documents by setting the chapter or section title information of laws, regulations, and emergency plans as nodes. Nodes labeled as 'Process work' are the detailed contents corresponding to the titles.

(4)   The entity label 'Emergency Type' includes seven nodes: gale disaster, snow disaster, rainstorm disaster, fire accident, geological disaster, ecological disaster, and traffic emergency. It mainly refers to the emergency plans made public by government departments. The setting can connect emergency plans to emergency products and create the information chain 'Process Work—Procedure—Emergency Plan—Emergency Type—Emergency Products—Emergency Enterprise'.

(5)   The setting of the entity label 'Administrative Area' realizes the linkage of bus stops, bus lines, meteorological equipment, and traffic cameras. It mainly refers to the information on the website of Beijing Municipal Commission of Transport.

(6)   Set the entity label 'Illegal Code' to provide a basis for judging illegal behaviors. The illegal codes refer to the Road Traffic Management Information Code Part 31: Codes for Violations Categories, the Chinese industry standard. Set the violation, punishment basis, coercive measure, original law, and expanded knowledge as node attributes. The reference standards for the attributes are the Road Traffic Safety Law of the People's Republic of China and the Regulations on the Procedures for Handling Road Traffic Safety Violations. To assist law enforcement, connect nodes labeled as 'Illegal Code' with nodes labeled as 'Demerit Points', 'Punishment Type', or 'Fine'. Demerit points refer to the Measures for Scoring Management of Road Traffic Safety Violations. Punishment types refer to Law of the People's Republic of China on Administrative Penalty. Fines refer to the Road Traffic Safety Law of the People's Republic of China.

(7)   The setting of entity label 'Fault Code' clearly divides complex vehicle fault types. To provide reference, set the fault name and fault description as attributes. Connect fault code to fault part to integrate it into the body system of a vehicle. The classification of vehicle fault codes adopts the OBD, a fault classification standard, issued by the Society of Automotive Engineers.

(8)   In order to expand information, set entity labels such as 'Safety Expert', 'Journal', 'WeChat Official Account', 'Website', 'Research Institute' in the knowledge map. Provide more references for traffic safety management work.

(9)   Set the entity label 'Emergency Function', including four nodes: safety protection, safety emergency service, monitoring and warning, and emergency rescue and disposal. Classifying emergency products can facilitate product inquiry according to needs. The reference is the Product Catalog of Key Safety and Emergency Enterprises in Beijing.

3.2. Knowledge Graph Data Results

　　Taking Beijing as an example, collect information. To create nodes and relationships in bulk, divide the information into six parts based on its structural features and content, including illegal acts, vehicle failure, emergency response, legal norms, organization information, and road-related information.

(1)   The construction results for illegal acts are shown in Table 1, which includes the information on 452 types of illegal acts. The setting can be found in article (6) of Section 3.1. The traffic violation code consists of 5 digits. There are 6 types of demerit points: 0, 1, 2, 3, 6, and 12. The basic types of administrative penalties are divided into 7 categories: warnings, fines, confiscation of illegal gains, orders to suspend production or business, temporary suspension or revocation of permits or licenses, administrative detention, and other administrative penalties stipulated by laws and regulations. The

16 nodes labeled as 'Punishment Type' here include nodes representing combinations of basic penalty types.

Table 1. Entities about handling of illegal acts.

Type	Label	Property	Example	Number of Entities Involved
Node entity	Legal Code	Name	20010	452
		Violation	Article 70 (1) of the Road Traffic Safety Law of the People's Republic of China	
		Punishment Basis	Article 99 of the Road Traffic Safety Law of the People's Republic of China	
		Coercive Measure	Detain car.	
		Original Law	Article 70 (1): If a traffic accident occurs on the road, the driver of the vehicle shall immediately stop and protect the scene, and the driver of the vehicle shall immediately rescue the injured person and report promptly to the traffic police or the traffic management department of the public security organ on duty. Where the injured person changes location, their former position shall be marked.	
		Expanded Knowledge	Escape from a traffic accident refers to the illegal behavior of a party who, after a traffic accident, fails to protect the scene of the accident and intentionally escapes from it without reporting to the public security organs to escape responsibility.	
	Demerit Points	Name	0	6
	Punishment Type	Name	Fines, detention	16
	Fine	Name	500 yuan	20
Relation entity	rels_DemeritPoints (from Illegal Code to Demerit Points)	Name	rels_DemeritPoints	452
	rels_PunishmentType (from Illegal Code to Punishment Type)	Name	rels_PunishmentType	452
	rels_Fine (from Illegal Code to Fine)	Name	rels_Fine	452

(2) There are 5419 types of vehicle fault. The construction results for relevant entities and relationships are shown in Table 2. The setting can be found in article (7) of Section 3.1. The OBD fault code consists of one letter and four digits. "P" represents the powertrain codes, "C" represents the chassis codes, "B" represents the body codes, and "U" represents the network codes. According to the OBD fault codes, the fault parts include 15 categories, which are specific fault parts of four systems.

Table 2. Entities about vehicle failure.

Type	Label	Property	Example	Number of Entities Involved
Node entity	Fault Code	Name	B2692	5419
		Chinese Definition	前排乘客座椅安全带扣开关电路故障	
		English Definition	Front Passenger's Seat Belt Buckle Switch Circuit Fault	
		Fault Description	The function of the seat belt buckle switch/sensor is to monitor whether the seat belt is fastened.	
	Fault Part	Name	Body	15
Relation entity	rels_Part (from Fault Code to Fault Part)	Name	rels_Part	5419

(3)   Emergency response mainly includes the information about emergency plans, safety experts, key organizations, emergency rescue products, and safety media. The construction results are shown in Table 3.

Table 3. Entities about emergency.

Type	Label	Property	Example	Number of Entities Involved
Node entity	Emergency Plan	Name	Highway Traffic Emergency Plan	13
	Government Department	Name	Ministry of Transport	5
	Year	Name	2009	7
	Emergency Type	Name	Traffic emergency	7
		Introduction	Traffic emergencies include damage or potential damage to transportation facilities, traffic interruptions, and traffic jams. In addition, they include transportation support for natural disasters, accidents, public health incidents, and social security incidents that occur suddenly, as well as emergency measures to evacuate or rescue personnel.	
	Procedure	Name	Emergency organization_Highway Traffic Emergency Plan	102
		First Serial Number	Chapter 2	
	Process Work	Name	Emergency leading group_Highway Traffic Emergency Plan	392
		Content	(omission)	
	Safety Expert	Name	Mr. Jia	464
		Sex	Male	
		Work Unit	National Center for Occupational Safety and Health, Beijing, China	

Type	Label	Property	Example	Number of Entities Involved
		Title	Senior engineer	
		Education Background	Postgraduate	
		Major	Policy research	
	Emergency Enterprise	Name	AKD Communication Technology Co., Ltd., Beijing, China	111
		Introduction	AKD Communication Technology Co., Ltd. was registered on 17 December 2001 in Beijing Zhongguancun Science and Technology Park with a registered capital of 39.36 million yuan. It is a high-tech enterprise specializing in the research and development, production, and sales of satellite communication antennas and terminal systems. This company provides customized satellite communication products. At the same time, this company also undertakes various satellite communication system integration projects.	
		Email	slwang@akdtech.com.cn	
		Fax	010-63726086	
		Website	www.akdtech.com.cn	
	Dangerous Transport Enterprise	Name	Beijing Ande Jingcheng Transportation Co., Ltd., Beijing, China	229
		Operation Scope	Transporting dangerous goods belonging to the third category	
	Public Organization	Name	China Society of Emergency Management, Beijing, China	13
		Introduction	China Society of Emergency Management is a primary organization under the leadership of the Party School of the Central Committee of C.P.C. It is a national, academic, and public welfare organization composed of experts; scholars; practitioners; relevant professional organizations, enterprises, and institutions; non-governmental organizations engaged in theoretical research, teaching, and training; and consulting services for emergency management. It is committed to developing and applying modern emergency management concepts, methods, and technologies and enhancing the ability of the entirety of society to prevent and respond to various emergencies.	

Table 3. *Cont.*

Type	Label	Property	Example	Number of Entities Involved
	Research Institute	Name	China Academy of Safety Science and Technology, Beijing, China	38
		Introduction	China Academy of Safety Science and Technology is a comprehensive scientific research institution for public welfare affiliated with the Ministry of Emergency Management of the People's Republic of China. It focuses on fundamental, comprehensive, and forward-looking scientific research in the field of safety, as well as addressing major technical issues such as accident prevention, monitoring, early warning, and emergency rescue.	
	Emergency Products	Name	Mobile emergency communication vehicle	269
		Introduction	Mobile emergency communication vehicles are suitable for handling emergency situations related to mobile communication. In case of large-scale meetings, gatherings, or communication-related accidents, they can rapidly increase communication capacity.	
		Product Type	Emergency communication and command product	
	Emergency Function	Name	Emergency rescue and disposal	4
	Journal	Name	China Safety Science Journal	74
	WeChat Official Account	Name	Beijing Emergency	98
		Introduction	Interprets authoritative security policies. Transmits important security information. Disseminates practical security knowledge.	
	Website	Name	Safety Management Network	19
		Website	http://www.safehoo.com/, accessed on 12 December 2022	
		Organizer	Orient Imagination Technology Co., Ltd., Beijing, China	
Relation entity	rels_IssuedTime (from Emergency Plan to Year)	Name	rels_IssuedTime	13
	rels_Reply (from Emergency Plan to Emergency Type)	Name	rels_Reply	13
	rels_Contain (from Emergency Plan to Procedure, from Procedure to Process Work)	Name	rels_Contain	494

Table 3. *Cont.*

Type	Label	Property	Example	Number of Entities Involved
	rels_Issue (from Government Department to Emergency Plan)	Name	rels_Issue	13
	rels_Product (from Emergency Enterprise to Emergency Products)	Name	rels_Product	269
	rels_EmergencyFunction (from Emergency Products to Emergency Function)	Name	rels_EmergencyFunction	269
	rels_Reply (from Emergency Products to Emergency Type)	Name	rels_Reply	269

There are a total of 13 emergency plans, including 5 national emergency plans and 8 Beijing emergency plans. Based on unexpected situations, create 7 nodes labeled as 'Emergency Type' to divide emergency plans. The settings can be found in articles (3) and (4) of Section 3.1.

According to the Science and Technology Work Manual of Beijing Emergency Management available on the official website, collect information about emergency management experts. Set the sex, work unit, title, education background, and major of each safety expert as attributes, so that experts in the corresponding fields can be consulted when different accidents happen. The setting can be found in article (8) of Section 3.1.

Key organizations include research institutes, public organizations, dangerous transport enterprises, and emergency enterprises. The setting can be found in article (1) of Section 3.1. The attribute 'Operation Scope' of dangerous transport enterprises defines the hazard class of dangerous goods, which refers to the Classification and Code of Dangerous Goods.

Emergency rescue product information is crucial for both prevention and emergency response. The data come from the Product Catalog of Key Safety and Emergency Enterprises in Beijing. To achieve better linkage of emergency information, emergency products are classified by connecting them with nodes labeled as 'Emergency Type'. Simple and clear function information is conducive to the search for emergency products. The emergency functions are divided into four types, as shown in article (9) of Section 3.1. The attributes of emergency products include name, introduction, and product type.

(4) There are a total of 63 legal norms imported into the knowledge graph. The construction results are shown in Table 4. According to the legal hierarchy, they can be divided into laws, departmental rules, and administrative regulations, with a total of 3, 52, and 8 documents, respectively. The reference is the Legislation Law of the People's Republic of China. Additionally, the 63 laws and regulations are divided into 7 categories based on managed objects, as shown in articles (2) and (3) of Section 3.1.

Table 4. Entities about legal norms.

Type	Label	Property	Example	Number of Entities Involved
Node entity	Legal Norm	Name	Road Traffic Safety Law of the People's Republic of China	63
	Legal Hierarchy	Name	Law	3
	Year	Name	2021	20
	Government Department	Name	The National People's Congress	12
	Procedure	Name	Vehicles and Drivers_ Road Traffic Safety Law of the People's Republic of China	640
		First Serial Number	Chapter II	
	Process Work	Name	Article 12 In case any of the following circumstances occur, the corresponding registration shall be made: (1) The ownership of a motor vehicle is transferred; (2) Any registered content of a motor vehicle is modified; (3) A motor vehicle is mortgaged; (4) A motor vehicle is discarded as unserviceable.	3845
		Secondary Serial Number	Article 12	
	Process Modules	Name	Duty	7
Relation entity	rels_IssuedTime (from Legal Norm to Year)	Name	rels_IssuedTime	63
	rels_Reply (from Legal Norm to Process Modules)	Name	rels_Reply	63
	rels_Contain (from Legal Norm to Procedure, from Procedure to Process Work)	Name	rels_Contain	4456
	rels_Issue (from Government Department to Legal Norm)	Name	rels_Issue	63
	res_LegalHierarchy (from Legal Norm to LegalHierarchy)	Name	res_LegalHierarchy	63

(5) A total of 12,946 organization nodes are created, and the detailed construction results are shown in Table 5. Telephone and address are basic information for most organizations, as detailed in Table 6. The setting can be found in article (1) of Section 3.1. The attribute 'Company Type' of the nodes labeled 'Motor Vehicle Maintenance Station' is explained as follows: According to the national standard Certification Requirements for Motor Vehicle Maintenance and Repair Enterprises, automobile maintenance enterprises are divided into three categories. The first kind of enterprises are engaged in

major vehicle repair, routine vehicle repair, special vehicle repair, and vehicle maintenance. The second kind of enterprises are engaged in primary and secondary maintenance and routine vehicle repair. The third kind of enterprises are only engaged in special vehicle repair or maintenance.

Table 5. Entities about organizations.

Type	Label	Property	Example	Number of Entities Involved
Node entity	Driving School	Name	Jingtian Driving School	104
		Postal Code	100022	
		Work Time	8:00–17:00 (from Monday to Sunday), excluding national holidays	
		Bus Line	Take bus 30, 34, or 801 to Qijiafen and get off.	
	Driving Test Location	Name	Laoshan Examination Hall	29
		Work Time	8:00–12:00, 13:00–17:00 (from Monday to Sunday), excluding national holidays	
		Operation Scope	Holds examinations for motor vehicle drivers.	
	Motor Vehicles Department	Name	Chaoyang Traffic Division Vehicle Management Station	29
		Work Time	8:30–18:00 (from Monday to Friday), 9:00–16:00 (Saturday and Sunday), excluding national holidays	
		Operation Scope	The verification, modification, and replacement of motor vehicle driving licenses and other business	
	Motor Vehicle Inspection Site	Name	Hongyuannankou Motor Vehicle Inspection Station	36
		Work Time	8:00–12:00, 13:00–17:00 (from Monday to Saturday)	
		Operation Scope	Large, medium, small, and micro passenger vehicles; light and micro trucks; heavy, medium, light, and micro special work vehicles; and light trailers (except for diesel engines)	
	Motor-Vehicle-Dismantling Plant	Name	Beijing Motor Vehicle Dismantling Co., Ltd., Beijing, China	7
		Work Time	8:00–16:00 (from Monday to Friday), excluding national holidays	
		Operation Scope	Mainly handles the following business: dismantling scrapped cars, leasing prop cars, and processing small metal objects.	
	Motor Vehicle Maintenance Unit	Name	Beijing Kedi Automobile Service Company	1591
		Juridical Person	Mr. Sun	
	Motor Vehicle Maintenance Station	Name	Beijing Shangdixinchen Automobile Sales and Service Co., Ltd.	6189
		Company Type	Class I motor vehicle maintenance station	
		Employer Address	1 Sports University West Road, Haidian District, Beijing	

Type	Label	Property	Example	Number of Entities Involved
	Traffic-Accident-Handling Agency	Name	Dongcheng Traffic Division Temple of Heaven Branch	57
		Work Time	8:30–18:00 (from Monday to Friday), 9:00–16:00 (Saturday and Sunday), excluding national statutory holidays	
		Operation Scope	Handles the extension of the Beijing Entrance Permit. Maintains traffic facilities, traffic order, and traffic safety and smoothness. Handles vehicle traffic accidents and violations. Creates publicity for traffic laws and regulations.	
	Traffic-Violation-Handling Agency	Name	Pinggu Traffic Division Law Enforcement Station	46
		Work Time	8:30–18:00 (from Monday to Friday), 9:00–16:00 (Saturday and Sunday), excluding national statutory holidays	
		Operation Scope	Handles vehicle violations.	
	Police Station	Name	Chaoyang Gate Police Station	351
	Hospital	Name	Beijing Jishuitan Hospital	724
		Postal Code	100035	
		District Code	110102	
		District Name	Xicheng District	
		Hospital Level	Grade III	
		Hospital Rank	A	
		Big Class	A. hospital	
		Middle Class	A1. comprehensive hospital	
		Small Class	A100. comprehensive hospital	
		Economic Type	State-owned	
		Management Type	Non-profit medical institution	
	Clinic	Name	Beijing Chongren Clinic of Traditional Chinese Medicine	4032
		Postal Code	100035	
		District Code	110102	
		District Name	Xicheng Didtrict	
		Small Class	D212. Chinese medicine clinic	
		Economic Type	Stock cooperation	
		Management Type	For-profit medical institution	
	Other Medical Organization	Name	Beijing Adikang Medical Laboratory	124
		Postal Code	100176	
		District Code	110115	
		District Name	Daxing District	
		Economic Type	Other	
		Management Type	For-profit medical institution	

Table 5. *Cont.*

Type	Label	Property	Example	Number of Entities Involved
	Emergency Station	Name	Beijing Chaoyang District Chuiyangliu Aid Station	12
		Postal Code	100022	
		District Code	110105	
		District Name	Chaoyang District	
		Economic Type	State-owned	
		Management Type	Non-profit medical institution	
	Safety Assessment Organization of Beijing	Name	Beijing Long'an Kanghua Production Safety Research Center	33
		Introduction	The company has strong technical strength and rich management experience in safety evaluation, safety training, and safety production technology consulting. The company mainly provides services in the fields of safety films and television, safety management training, and safety evaluation.	
		Website	http://184255.71ab.com/, accessed on 12 December 2022.	
	Environment-Monitoring Organization	Name	Chinese Research Academy of Environmental Sciences	16
	Testing Organization	Name	National Safety Production Beijing Labor Protection Equipment Testing and Inspection Center	4
		Introduction	The National Safety Production Beijing Labor Protection Equipment Testing and Inspection Center is the first professional testing institution for labor protection equipment established in China. It has been engaged in various inspections, research, and service work for labor protection equipment for a long time, and it is a national inspection institution for public welfare.	

Table 6. Entities about telephone and address.

Type	Label	Property	Example	Number of Entities Involved
Node entity	Telephone	Name	84081551	2107
	Address	Name	121 Chaonei South Street	7273
	Organization Type	Name	Medical Organization	6

Table 6. *Cont.*

Type	Label	Property	Example	Number of Entities Involved
Relation entity	res_Telephone (from Organization to Telephone)	Name	res_Telephone	7499
	res_Address (from Organization to Address)	Name	res_Address	2407
	res_OrganizationType (from Organization Type to Organization)	Name	res_OrganizationType	12,946

The attributes 'Hospital Level' and 'Hospital Rank' of the nodes labeled 'Hospital' are explained as follows: According to Rules to be in Charge of Hospitals by Grade, the Chinese hospital classification system divides hospitals into primary (grade I), secondary (grade II), and tertiary (grade III) levels. Grade I and grade II hospitals are further divided into A, B, and C, while grade III hospitals are further divided into AAA, A, B, and C. The attribute 'Hospital Level' refers to the grade of the hospital. The attribute 'Hospital Rank' refers to the sub-classification, be it AAA, A, B, or C.

Attributes of medical organizations are explained as follows: The reference is the national health industry standard Classification and Codes for Health Institutions. According to the nature of the service provided, medical institutions are divided into big class, middle class, and small class. According to management type, medical institutions are divided into three categories: non-profit medical institutions, for-profit medical institutions, and other medical institutions. The economic type refers to the *Classification and Code for Economic Categories*.

(6) Road-related information is mainly about roads, buses and facilities. The results are shown in Table 7. The setting can be found in article (5) of Section 3.1.

Table 7. Entities about roads.

Type	Label	Property	Example	Number of Entities Involved
Node entity	Parking Area	Name	Daoxiangyuan Qiaonan Parking Lot	502
		Park Type	The park area under the overpass	
		Park Address	Daoxiangyuan Bridge, Haidian District	
		Park Business Unit	Beijing Static Traffic Anda Parking Management Co., Ltd., Beijing, China	
	Bus Line	Name	101	1219
	Bus Stop	Name	Baiwanzhuang West Entrance	8971
	Administrative Area	Name	Dongcheng District	16
		Area	41.84 km^2	
		Government Residence	Jingshan Subdistrict	

Table 7. *Cont.*

Type	Label	Property	Example	Number of Entities Involved
		Jurisdiction	Donghua Gate Subdistrict, Jingshan Subdistrict, Jiaodaokou Subdistrict, Anding Gate Subdistrict, Beixin Bridge Subdistrict, Dongsi Street, Chaoyang Gate Subdistrict, Jianguo Gate Subdistrict, Dongzhi Gate Subdistrict, Hepingli Street, Qianmen Subdistrict, Chongwenmenwai Subdistrict, Donghuashi Subdistrict, Longtan Street, Tiyuguan Road Subdistrict Street, Tiantan Street, Yongdingmenwai Subdistrict	
	National Road	Name	Beijing Harbin Expressway	177
		Route Code	G1	
	Provincial Road	Name	Jingliang Road	443
		Route Code	S315	
	Expressway	Name	East Fourth Ring South Road (main road)	80
		Start Point	Yaowa Lake Bridge	
		End Point	Shibalidian Bridge	
		Length	4766.9 m	
	Main Road	Name	East Fourth Ring South Road (outer-ring auxiliary road)	359
		Start Point	Yaowa Lake Bridge	
		End Point	Shibalidian Bridge	
		Length	4959.2 m	
	Meteorological Equipment	Name	3110116001	39
		Location Point	G111 Jingjia Road K84 Houshanpu Village	
	Traffic Camera	Name	Wuxing Intersection of National Road G103	31,504
		Equipment Direction	East to West	
		North–South Intersection Road	Wuxing Road	
		East–West Intersection Road	National Road G103	
Relation entity	res_AdministrativeArea (from Bus Line, Meteorological Equipment, Traffic Camera to Administrative Area)	Name	res_AdministrativeArea	25,417
	res_BusStop (from Bus Line to Bus Stop)	Name	res_BusStop	31,911

According to administrative level, highways can be divided into national roads and provincial roads. The main roads belong to the urban areas, and their levels are higher than those of secondary roads but lower than those of expressways, according to the Design for Specification of Urban Road Engineering. Create four types of nodes, labeled as "National Highway", "Provincial Road", "Expressway", and "Main Road". According to

the public information from the Beijing Municipal Commission of Transport, the attribute "Park Type" of nodes labeled "Parking Area" includes three types: public service, the park area off road, and the park area under the overpass.

Bus lines' names starting with letters are related to administrative areas. For example, bus lines in Changping District have names starting with "C", and bus lines in Fangshan District have names starting with "F". In order to better reflect the relationship between bus lines and administrative areas, create 16 nodes labeled "Administrative Area" to represent 16 administrative areas of Beijing, and add jurisdiction information to their attributes.

Facilities mainly include meteorological equipment and traffic cameras. Highway meteorological equipment can detect weather indicators such as temperature, wind direction, and precipitation and provide effective suggestions for transportation and travel. Set attributes including information about intersecting roads and directions for nodes labeled 'Traffic Camera'.

3.3. Query Implementation

3.3.1. Query Implementation Based on Cypher

Cypher is a declarative graph query language that draws on SQL language. Queries are composed of different statements. Based on the characteristics of the database in this article, the main applications are as follows: (1) Node query based on ID. (2) Query nodes for specific entity labels, attributes, and relationships. (3) Querying the number of nodes. (4) Node query based on the existence of the node attribute. (5) Implementing node query based on specific format of the node attribute. (6) Querying the existence of relationships. (7) Querying the relationship type between entities. (8) Multi-level relationship query. (9) Projection query.

For example, one query that can be investigated by setting a specific format is which motor vehicle inspection sites have light trailer inspection qualification, and the query statement is 'MATCH (n:Motor Vehicle Inspection Site) WHERE n. OperationScope CONTAINS 'light trailer''. If the projection query target is to obtain the names, introductions, and product types of emergency products produced by Beijing CCOM Communications Technology Limited Company, the query statement is 'MATCH (n:Emergency Enterprise {Name:'Beijing CCOM Communications Technology Limited Company'})-[r:res_Product] →(m:Emergency Product) With n, collect (m{.Name,.Introduction,.ProductType}) AS ms RETURN n {.Name,ms:ms}'. The result is a dictionary containing two keys: 'Name' and 'ms'. The value corresponding to "Name" is the company name, and the value corresponding to 'ms' is a dictionary of the attributes of name, introduction, and product type, which are stored in a list.

3.3.2. Query Implementation Based on Rule Matching—Taking Vehicle Failure as an Example

To achieve the query function, set 11 problem types based on rule matching to meet the needs of describing vehicle failure, as shown in Table 8. The problem type 'FaultPart - FaultCode' represents the query of all fault codes of a certain fault part. Set six types of question vocabularies. The specific steps are as follows:

(1) Build five types of dictionaries of query objects, including fault code, Chinese name, English name, fault description, and fault part. Build six types of question dictionaries, as shown in Table 9. Choose the AC automaton, a multi-pattern-matching algorithm. Add five types of dictionaries to the actree dictionary using the actree module of the ahocorasick library.

(2) Decompose the input questions using the actree dictionary, and match the types of the words that are the query objects.

(3) Loop the question vocabularies through the input questions for matching.

(4) Construct 11 rules to classify the questions and determine the types of them based on the matching results of query objects and question words.

(5) Modify the question sentences, that is, convert them into Cypher query templates. Then, modify the returned query results, that is, set the answer templates of the chat robot. Implement the final query function as shown in Figure 3.

Figure 3. Query function implementation.

Table 8. Question type.

Order Number	Question Type
1	FaultPart_FaultCode
2	ChineseDefinition_FaultCode
3	FaultCode_FaultPart
4	ChineseDefinition_FaultPart
5	FaultCode_ChineseDefinition
6	FaultPart_ChineseDefinition
7	FaultCode_EnglishDefinition
8	ChineseDefinition_EnglishDefinition
9	FaultCode_FaultDescription
10	ChineseDefinition_FaultDescription
11	FaultSystem

Table 9. Question vocabulary.

Class	Question Vocabulary
system division	fault type, fault system, fault classification, fault part, division, segmentation, etc.
fault code	code, fault code, etc.
fault part	where, part, section, fault part, occur, etc.
Chinese definition	Chinese, name, fault name, performance, fault type, etc.
English definition	English, name, English name, etc.
fault description	phenomenon, performance, how, detail, fault description, content, detailed description, meaning, cause, etc.

4. Design and Construction of Traffic Accident Knowledge Graph

4.1. Characteristics of Traffic Accident Report Writing

According to the method of data processing, Section 4 discusses extracting node entities and relationship entities from unstructured data. Firstly, annotate node entities and relationship entities in unstructured data according to the design. Secondly, train the extraction models of node entities and relationship entities using deep learning methods. Like structured data, use py2neo to create entities in Neo4j for data storage.

Traffic accidents have a huge impact on vehicles, transportation facilities, and personnel, which seriously hinders the sustainable development of transport. Compared to news, traffic accident reports provide more detailed descriptions and clearer perspectives for safety analysis. The writing characteristics of traffic accident reports are analyzed as follows:

(1) Introduction

The first sentence of a traffic accident report clearly outlines information on the time, location, nature, casualties, and economic cost of the traffic accident. Then, it explains how the government organizes its investigation.

(2) Basic information

This mainly refers to the basic information on the drivers, vehicles, roads, and organizations related to traffic accidents. The driver part introduces the driver's vehicle, sex, age, household registration, driving license qualification, class of vehicle for which their driver's license is valid, transportation certificate, etc. The vehicle part introduces the license plate number, maker and model, occupants, load situation, nature of the vehicle's use, road-worthiness certificate, transportation permits, driving speed, etc. The road part introduces specific locations, directions, slopes, widths, curves, traffic markings, road signs, etc. In addition, a brief introduction is given to the weather and visibility conditions. The organizations part usually introduces the operational status of the freight company.

(3) Accident process and emergency response status

The process part introduces the vehicle behavior and driver behavior in chronological order and describes the accident scene and consequence. The emergency response part mainly focuses on the actions of government departments dealing with traffic accidents, such as medical and health institutions, safety supervision departments, public security departments, and transportation departments.

(4) Reason for and nature of the accident

The causes of the accident are divided into two parts: direct causes and indirect causes. In most cases, the direct cause is related to personal behavior and the indirect cause is related to safety management of enterprises or institutions.

(5) Suggestions for handling the persons and units responsible for the accident

List the persons or units responsible, and describe laws violated, dereliction or illegal behavior, punishment suggestions, or results.

(6) Suggestions for accident prevention and rectification

The suggestions are to work on the weak points of traffic safety management.

4.2. Entity and Relationship Settings for Traffic Accidents

Based on the results from analysis of the writing characteristics, set node entities and relationship entities to extract main safety elements of traffic accident reports. The traffic accident report knowledge graph centers on the analysis of accidents, vehicles involved, and people or institutions involved to integrate information, as shown in Figure 4. It includes 21 types of entities and 22 types of relationships. The settings are as follows:

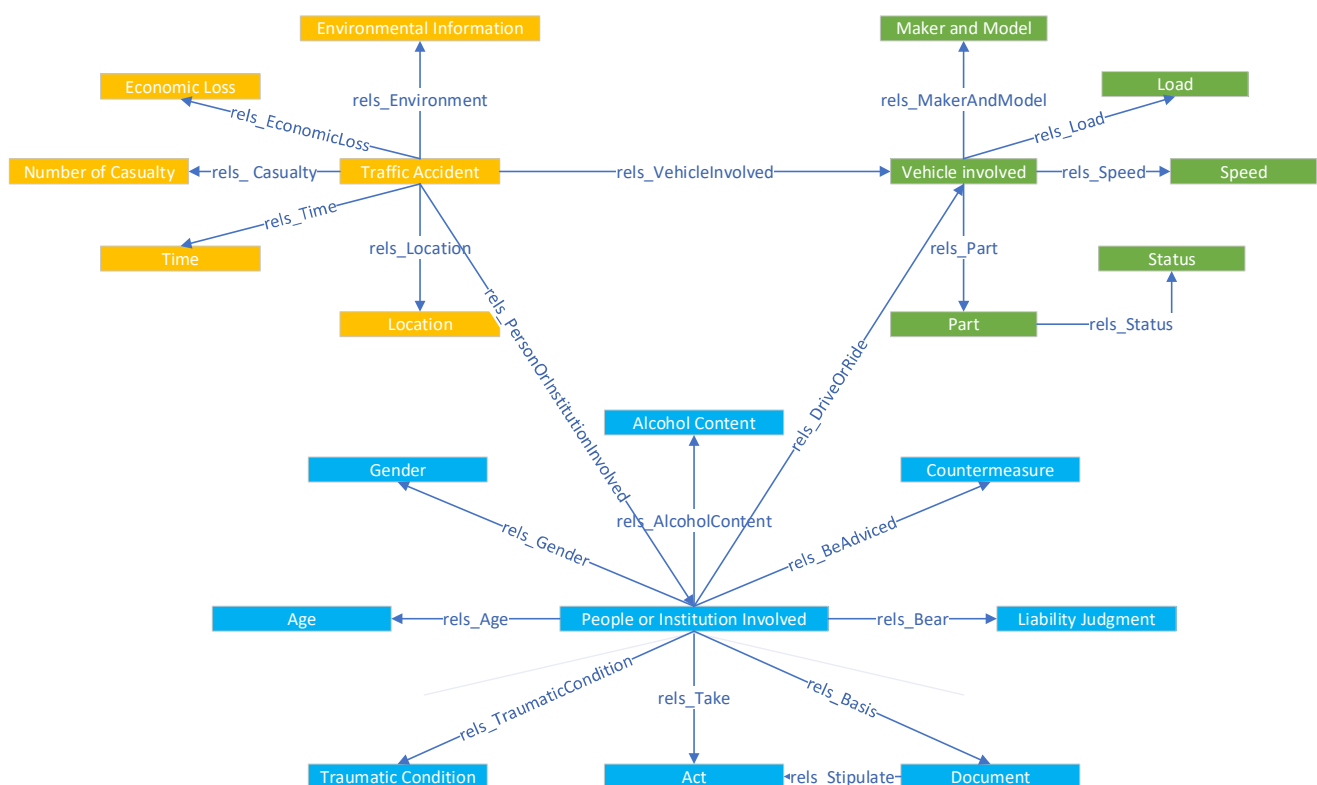(1) Set node entities and relationship entities centered on accidents.

Figure 4. Knowledge graph framework of traffic accident.

The extraction of accident entities is the basis for subsequent multiple relationship extraction. To increase the number of entities, label accident triggers, which are the verbs such as 'crush', 'scratch', 'roll', 'overturn', 'fall', and 'fire'. Secondly, it is necessary to clearly outline the overall situation of the accident. Set entity labels of basic traffic accident information, including time, location, vehicles involved, persons or institutions involved, environmental information, number of casualties, and economic loss. The description of time and location varies in the level of refinement. The numerical information is about the number of casualties and economic loss. Environmental information is mainly about weather, road and transportation facility.

(2)　Set node entities and relationship entities centered on the vehicles involved

Label the vehicles involved by the description of their vehicle type and license plate. Take vehicles involved as the core, and set other entities based on two aspects: basic information and technical status. The entity labels of basic information include the vehicle's maker and model, load, and speed. The model contains the specification and structure information of the vehicle. The load status and driving speed are important information to determine whether the driving is legal. The technical status of the vehicles involved is the key element of safety analysis. The technical status can be expressed through two types of entities: body parts and body status.

(3)　Set node entities and relationship entities centered on the people or institutions involved

Mainly extract the names of the people or institutions involved. Pay attention to the basic information of the people, including age, gender, traumatic condition, and alcohol content. Information about unsafe behaviors and responsibilities is the key to analyzing the causes of accidents. Therefore, set entity labels including act, document, liability judgement, and countermeasure.

4.3. Entity and Relationship Extraction Results

Crawl traffic accident report data from safety information websites and official websites of emergency management bureaus. Store each traffic accident report in Excel. In this case, the data size is 758 KB. Use "Baidu Brain EasyData" to annotate entities and relationships. The entity recognition part adopts the BMES annotation system.

4.3.1. Named-Entity Recognition Results

The ratio of training set, test set, and validation set is 8:1:1. The dropout learning probability is 0.1. Train data using the Bert model, Bert-CRF model, Bert-BiLSTM model, and Bert-BiLSTM-CRF model. The training is divided into a total of 15 Epochs. Batch size is 8. Learning rate is 3×10^{-5}.

Use the Bert model to predict entity labels on annotated data. The maximum sentence length is 512. Each word is represented by a 768-dimensional vector. Use multiple-attention mechanism with 12 head counts.

On the basis of the Bert model, add CRF model to calculate the transfer probability between labels, thereby enhancing the constraints of label annotation rules. On the basis of the Bert model, add BiLSTM model, converting vectors into 128 dimensions for training, to enhance the learning of the previous and subsequent texts. On the basis of the Bert model, add BiLSTM-CRF model to integrate the advantages of the three models.

Adopt precision, recall rate, and F1 as evaluation indicators. The calculation formula is as follows:

$$xPrecision = \frac{\sum_{i=1}^{n} \text{The number of correctly predicted labels in sentence}_i}{\sum_{i=1}^{n} \text{The number of predicted labels in sentence}_i} \tag{1}$$

$$Recall = \frac{\sum_{i=1}^{n} \text{The number of correctly predicted labels in sentence}_i}{\sum_{i=1}^{n} \text{The number of labels that should be included in sentence}_i} \tag{2}$$

$$F_1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \tag{3}$$

The training results of each model on the test set are shown in Table 10. According to accuracy and F_1 of entity prediction, the order of model performance is Bert-BiLSTM-CRF > Bert-BiLSTM > Bert-CRF > Bert. The recall of entity prediction is ranked as follows: Bert-BiLSTM-CRF > Bert-CRF > Bert-BiLSTM > Bert. Overall, Bert-BiLSTM-CRF has the best prediction performance.

Table 10. Training results of Bert, Bert-CRF, Bert-BiLSTM, and Bert-BiLSTM-CRF.

Model	Precision	Recall	F1
Bert	84.87%	89.96%	87.34%
Bert-CRF	85.98%	90.71%	88.28%
Bert-BiLSTM	86.64%	90.33%	88.45%
Bert-BiLSTM-CRF	86.93%	91.20%	89.01%

4.3.2. Relationship Extraction Results

Compared with entity extraction, it is easier to achieve better results for relationship extraction. Use Bert model to extract relationships. In the case of multiple relationships, Bert can extract the relationship from each entity pair separately.

(1) Data processing: (1) In terms of encoding, use special symbols to identify the positions of subjects and objects, and incorporate relationship label to improve learning effectiveness. (2) Data augmentation: Due to the lack of annotation for some relationships, data augmentation was performed on 8 types of relationships: 'rels_AlcoholContent', 'rels_Age', 'rels_TraumaticCondition', 'rels_Speed', 'rels_Maker AndModel', 'rels_Load', 'rels_EconomicLoss', and 'rels_Stipulate'. Use Easy Data Augmentation, including synonym replacement, random insertion, random exchange, and random deletion

methods. (3) Data shuffling: To prevent similar text from being continuously input, shuffle the data to make them more uniform.

(2) Training parameters: The specific parameters are shown in Table 11.

Table 11. Parameters of Bert.

Parameters	Value
Max sentence length	512
Hidden size	768
Number of hidden layers	12
Number of epochs	2
Learning rate	5e-5
Batch size	16
Dropout rate	0.5

(3) The training results of the test set are shown in Table 12. Due to the uneven sample size, use macro-average to calculate the mean of each type result, which is beneficial for evaluating the performance of small data size.

Table 12. Training results of Bert.

Average	Precision	Recall	F1
macro-average	0.9850	0.9657	0.9744

$$recision = \frac{\sum_{i=1}^{n} The\ number\ of\ correctly\ predicted\ relationships}{\sum_{i=1}^{n} The\ number\ of\ predicted\ relationships} \tag{4}$$

$$Recall = \frac{\sum_{i=1}^{n} The\ number\ of\ correctly\ predicted\ relationships}{\sum_{i=1}^{n} The\ number\ of\ relationships\ that\ should\ be\ included} \tag{5}$$

$$F_1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \tag{6}$$

4.4. Application of Traffic Accident Knowledge Graph

Import traffic accident entities and relationships into the knowledge graph. Create a total of 9318 nodes and 14,273 relationships. The traffic accident knowledge graph can structurally display accident information and condense accident reports based on the core nodes labeled as accident, people or institutions involved, and vehicle involved. It reduces the burden of reading and effectively integrates key safety elements for safety analysis.

The knowledge graph centered on accidents is shown in Figure 5. The accident is characterized by the word 'scratch' and displayed as the core to show the time, location, people and institutions involved, vehicle information, and their relationships. Depict the accident succinctly and clearly.

In addition to structurally depicting accidents for safety analysis, the traffic accident knowledge graph can achieve more convenient query and statistic functions. Query the nodes about the accident elements and their specific relationships to meet work needs. For example, if the query target is the driver of a specific vehicle, the query statement is 'MATCH (n:Vehicle Involved)←[r:res_DrivingOrRiding]-(m:Person or Institution Involved) WHERE n.name='Anhui Province KG**' RETURN m'.
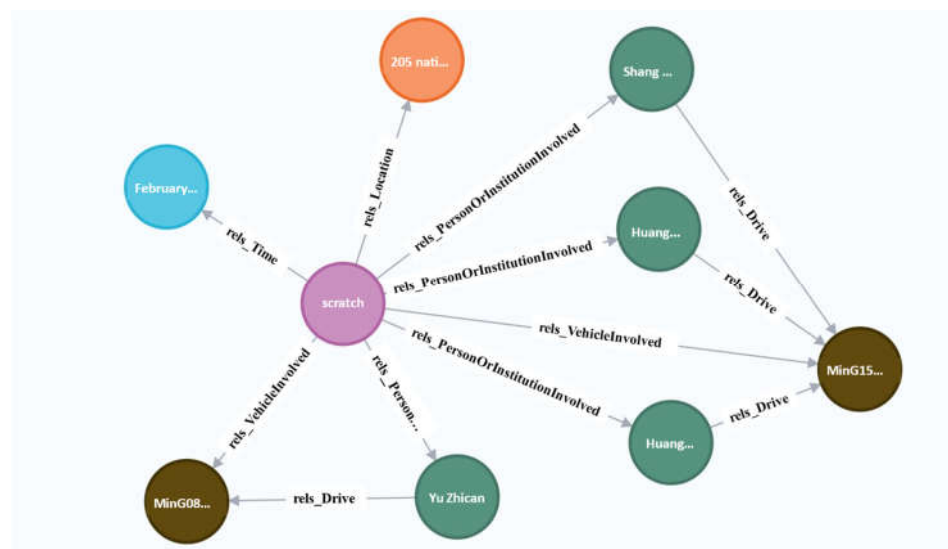
Figure 5. Knowledge graph taking traffic accidents as the core.

Furthermore, the graph can calculate accident statistic indicators, such as the number of a certain gender, accidents in a certain year, or accidents in a certain province. If the query target is the number of accidents that occurred in 2017, the query statement is 'MATCH (n:Time) WHERE n.name=~'.*2017.*' RETURN COUNT (n)'.

A knowledge graph can be used to search for specific safety points, such as unsafe behavior, and countermeasures. For example, the goal is to search for overspeed. The label is 'Act', containing the content 'overspeed', and the query statement is 'MATCH (n:behavior) WHERE n.name=~'.*overspeed.*' RETURN n'. If the goal is to search for safety training, the label is 'Countermeasure', including the content 'safety training', and the query statement is 'MATCH (n:Countermeasure) WHERE n.name=~'.*Training.*' RETURN n'.

5. Conclusions

As data accumulate in the information age, the structured and systematic display of knowledge is beneficial for sustainable and healthy development of transport. Knowledge graphs can reflect safety analysis ideas by depicting entities, attributes, and relationships, allowing them to be better applied in work. According to the method of data processing, a traffic safety management knowledge graph and a traffic accident knowledge graph were constructed.

For structured data disclosed by government apparatus, enterprises, and media, a total of 54 types of node entities and 14 types of relationship entities were designed, to serve traffic safety management work. According to the design, py2neo was used to build the traffic safety management knowledge graph in Neo4j, which achieves information linkage and information search service for specific work.

Unstructured data mainly include traffic accident reports. In accordance with the writing characteristics, a traffic accident knowledge graph was constructed with people and institutions involved, vehicles involved, and accidents as the core, including 21 types of node entities and 22 types of relationship entities. Deep learning methods were used to train node entity extraction models. The entity extraction result for traffic accident reports shows that the Bert-BiLSTM-CRF model performs the best. Relationship extraction was implemented using the Bert model. The traffic accident knowledge graph was built by py2neo in Neo4j to achieve structured display and query functions for traffic accidents. The results show that the traffic safety management knowledge graph is a good tool for the sustainable development of transport.

Safety knowledge graphs can be applied to more fields. They can combine safety management models and safety data, which is beneficial to storing safety information tailored

to industry characteristics. Safety elements from a large amount of texts are captured by the model with security analysis thinking, and graphs can be formed displaying structured information, which is helpful for sustainable development of safety management.

Author Contributions: Conceptualization, D.Y. and K.Z.; methodology, D.Y. and K.Z.; software, D.Y.; validation, D.Y., K.Z. and C.Y.; formal analysis, D.Y. and K.Z.; investigation, D.Y.; resources, D.Y.; data curation, D.Y.; writing—original draft preparation, D.Y.; writing—review and editing, D.Y. and C.Y.; visualization, D.Y.; supervision, K.Z.; project administration, K.Z. All authors have read and agreed to the published version of the manuscript.

References

1. Quillian, M.R. Word concepts: A theory and simulation of some basic semantic capabilities. *Syst. Res. Behav. Sci.* **1967**, *12*, 410–430. [CrossRef] [PubMed]
2. Quillian, M.R. The teachable language comprehender: A simulation program and theory of language. *Commun. ACM* **1969**, *12*, 459–476. [CrossRef]
3. Buchanan, B.; Duda, R.; Chen, Y.; Xu, Y. Rule-based expert systems principles. *Comput. Sci.* **1986**, *22*, 23–37.
4. Guha, R.V.; Lenat, D.B. CYC: A mid-term report. *Appl. Artif. Intell.* **1991**, *5*, 45–86. [CrossRef]
5. Miller, G.A.; Fellbaum, C. WordNet Then and Now. *Lang. Resour. Eval.* **2007**, *41*, 209–214. [CrossRef]
6. Zeng, X.; Yang, C.; Tu, C.; Liu, Z.; Sun, M. Chinese LIWC lexicon expansion via hierarchical classification of word embeddings with sememe attention. In Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018.
7. Jiang, Y.; Gao, X.; Su, W.; Li, J. Systematic Knowledge Management of Construction Safety Standards Based on Knowledge Graphs: A Case Study in China. *Int. J. Environ. Res. Public Health* **2021**, *18*, 10692. [CrossRef] [PubMed]
8. Pedro, A.; Pham-Hang, A.-T.; Nguyen, P.T.; Pham, H.C. Data-Driven Construction Safety Information Sharing System Based on Linked Data, Ontologies, and Knowledge Graph Technologies. *Int. J. Environ. Res. Public Health* **2022**, *19*, 794. [CrossRef]
9. Mao, S.; Zhao, Y.; Chen, J.; Wang, B.; Tang, Y. Development of process safety knowledge graph: A Case study on delayed coking process. *Comput. Chem. Eng.* **2020**, *143*, 107094. [CrossRef]
10. Zheng, X.; Wang, B.; Zhao, Y.; Mao, S.; Tang, Y. A knowledge graph method for hazardous chemical management: Ontology design and entity identification. *Neurocomputing* **2020**, *430*, 104–111. [CrossRef]
11. Nicholson, D.; Greene, C.S. Constructing knowledge graphs and their biomedical applications. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 1414–1428. [CrossRef]
12. Yu, T.; Li, J.; Yu, Q.; Tian, Y.; Shun, X.; Xu, L.; Zhu, L.; Gao, H. Knowledge graph for TCM health preservation: Design, construction, and applications. *Artif. Intell. Med.* **2017**, *77*, 48–52. [CrossRef]
13. Zhang, Q.; Wen, Y.Q.; Han, D.; Zhang, F.; Xiao, C.S. Construction of knowledge graph of maritime dangerous goods based on IMDG code. *J. Eng.* **2020**, *2020*, 361–365. [CrossRef]
14. Zhang, L.; Zhang, M.; Tang, J.; Ma, J.; Duan, X.; Sun, J.; Hu, X.; Xu, S. Analysis of Traffic Accident Based on Knowledge Graph. *J. Adv. Transp.* **2022**, *2022*, 3915467. [CrossRef]
15. Liu, C.; Yang, S. Using text mining to establish knowledge graph from accident/incident reports in risk assessment. *Expert Syst. Appl.* **2022**, *207*, 117991. [CrossRef]
16. Liu, J.; Schmid, F.; Li, K.; Zheng, W. A knowledge graph-based approach for exploring railway operational accidents. *Reliab. Eng. Syst. Saf.* **2020**, *207*, 107352. [CrossRef]
17. Liu, Q.; Li, Y.; Duan, H.; Liu, Y.; Qin, Z.G. Knowledge Graph Construction Techniques. *J. Comput. Res. Dev.* **2016**, *53*, 582.
18. Goyal, A.; Gupta, V.; Kumar, M. Recent Named Entity Recognition and Classification techniques: A systematic review. *Comput. Sci. Rev.* **2018**, *29*, 21–43. [CrossRef]
19. Zhong, M.; Liu, G.; Xiong, J.; Zuo, J. DualNER: A Trigger-Based Dual Learning Framework for Low-Resource Named Entity Recognition. *IEEE Intell. Syst.* **2022**, *37*, 79–87. [CrossRef]
20. Tuo, M.M.; Yang, W.Z. Review of entity relation extraction. *J. Intell. Fuzzy Syst.* **2023**, *44*, 7391–7405. [CrossRef]

21. Qiao, B.; Zou, Z.; Huang, Y.; Fang, K.; Zhu, X.; Chen, Y. A joint model for entity and relation extraction based on BERT. *Neural Comput. Appl.* **2021**, *34*, 3471–3481. [CrossRef]

22. Hang, T.; Feng, J.; Wu, Y.; Yan, L.; Wang, Y. Joint extraction of entities and overlapping relations using source-target entity labeling. *Expert Syst. Appl.* **2021**, *177*, 114853. [CrossRef]