Preliminary design

1. System architecture

As is shown in Figure.1, our system architectures can be seen as a system in three layers, among which exists multiple interfaces to help different layers conenect with each other as well as transporting information. In concrete, we capture and colloect semi-structured and unstructured data from external data in websites. Then, the DBMS(Database Management System) layer store these data into different kinds of database like mysql, mongodb and neo4j in terms of data features. Next we construct the military knowledge graph in KG construction layer by using millions of entities, 10 millions of triples and extra ontology like conceptions, properties and axioms. And in the top application layer, KG can be utilized to develop some useful applications like search engine, query answering robot and data visualization, etc.

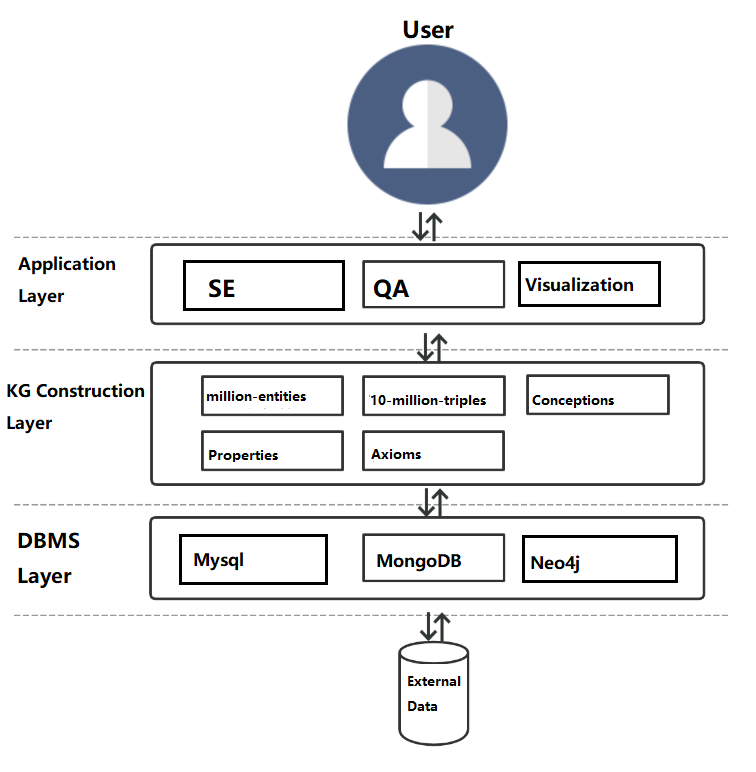


Fig.1: System architecture

2. Key modules and their functions

According to the military KG , we plan to develop three modules in our system (Fig.2), which are dependent from each other. In information retrieval module, the user can search for the information especially those in military field, like news and entities. In query answering module, a cute robot is happy to answer questions from users. In data visulization module, questions about an entity can be linked to its related entities to show more details.

Fig.2: Key modules

3. Key techniques and challenges

The research content of this project aims to realize knowledge inference and prediction of military knowledge graph by studying rule-based parallel inference technology and knowledge graph embedded representation learning technology through multi-source heterogeneous knowledge such as military knowledge graph and military detection log. In order to effectively improve the accuracy and scalability of reasoning.This project will focus on modeling of machine learning algorithms such as knowledge graph representation learning and Markov chain Monte Carlo algorithm, and research will focus on semantic model and representation learning technology.The specific difficulties of the current research are as follows:

(1) the uncertain and inconsistent reasoning in heterogeneous knowledge reasoning: on the one hand, the data source is multiple in the field of military, which leads to task concurrency and incompatible phenomenon, at the same time, uncertain and inconsistent reasoning tends to bring high complexity. How to improve the compatibility efficiency of uncertain and inconsistent reasoning is a difficulty. On the other hand, the traditional interpretable military reasoning methods, such as language description, attack graph and attack tree, are insufficient in terms of extensibility and global nature. They can only make preliminary analysis, but can't locate and grasp complex problems from a global perspective.In addition, at present, both uncertain and inconsistent reasoning are non-monotonous reasoning with great complexity. Therefore, the construction of high-performance inference interpretation model, especially the design of minimum inference process model, is a difficulty.

(2) Representation and reasoning of military framework knowledge: Knowledge map is mainly composed of framework knowledge (e.g., ontology axiom) and factual knowledge (e.g., triple).At present, most representation methods based on knowledge graphs focus on the factual knowledge in knowledge graphs rather than the framework knowledge in knowledge graphs.Framework for knowledge reasoning and prediction technology, knowledge of the relationship between semantic and characteristic is very important, military coalition reasoning being applied in the scene is very common, the semantic properties can not only help to improve the accuracy of military knowledge graph and the scope of the cover but also help to complete knowledge graph structure to evaluate its analysis.

(3) Mixed reasoning based on symbolic rules and machine learning: military knowledge usually consists of military knowledge base and military rule base.The former benefits from the data accumulation of information words, while the latter is manually edited and maintained by experienced commanders or combat experts.However, the reasoning before the two is independent of each other.If these two kinds of knowledge can be combined with constraints effectively, then more meaningful battlefield information can be mined and deduced, which is a major difficulty in the current joint model design.

In view of the above difficulties, the following key technologies are adopted to solve the problem:

(1) In view of the high complexity of uncertain and inconsistent reasoning, this subject establishes a parallel reasoning theory compatible with uncertain and inconsistent reasoning, and designs corresponding reasoning algorithms on this basis.The existing knowledge division theory and method are used to support the knowledge parallel inference technology, and the knowledge stratification theory and method are designed to support the iterative execution of knowledge inference.For minimum inference and reasoning process interpretation model of the existing problems of methods through the establishment of the facts of the input information one step, the key points in the knowledge base and inference conclusion as explanation space node, as the relationship between nodes in a single step reasoning explanation, on the edge of space, and explain the reasoning conclusion will be as the root, with the facts of the input information as leaves, interpretation tree dug up from the explanation space, according to the explanation in the tree node on the uncertainty of measurement explain the size of the tree, from explanation space heuristic search algorithm to efficiently search the minimum interpretation tree form minimum inference process,Then it provides a reasonable explanation for battlefield commanders in the face of complex battlefield environment.

(2) In view of the representive learning integrated with frame knowledge, this project is put forward based on Bert + BiLSTM + CRF frame knowledge representation method, and the integration of military knowledge graph within the context information and external text information to complete the modeling of hybrid knowledge representation and reasoning and prediction, so as to improve the accuracy of knowledge graph and the scope of cover.In addition, this project further improves the reasoning and prediction accuracy of the mixed embedding model by using the internal context information and external text information of the military knowledge map, so as to solve the problems of low accuracy and poor anti-noise ability of the present representation learning technology.