

存档编号_____

华北水利水电大学

North China University of Water Resources and Electric Power

毕 业 设 计

题目 中国姓氏文化可视化系统
的设计与实现

学 院 信息工程学院

专 业 软件工程

姓 名 苏玉恒

学 号 202018018

指导教师 _____

完成时间 _____

教务处制

目录

North China University of Water Resources and Electric Power.....	I
目录.....	VII
摘 要.....	IX
Abstract	XI
第 1 章 绪论.....	2
1.1 研究背景及研究意义.....	2
1.2 研究现状.....	3
1.2.1 知识图谱研究现状.....	3
1.2.2 问答系统研究现状.....	5
1.3 本文主要研究内容.....	5
第 2 章 相关技术和理论基础.....	10
2.1 知识图谱构建技术概述.....	10
2.1.1 知识抽取.....	11
2.1.2 知识融合.....	11
2.1.3 知识加工.....	12
2.2 问答系统技术概述.....	12
第 3 章 中国姓氏文化知识图谱的构建.....	14
3.1 中国姓氏文化知识图谱的构建流程.....	14
3.2 数据采集.....	15
3.3 数据清洗.....	16
3.4 知识抽取.....	16
3.5 知识对齐.....	18
3.6 知识加工.....	18
3.7 知识存储.....	19
第 4 章 需求分析.....	22
4.1 可行性分析.....	22
4.2 系统需求分析.....	22
4.2.1 系统功能概述.....	22
4.2.2 系统目标用户.....	22
4.2.3 功能需求.....	22
4.2.4 数据需求.....	23
4.2.5 性能需求.....	23
4.2.6 运行需求.....	24
4.2.7 可靠性需求.....	24
第 5 章 系统的设计与实现.....	26
5.1 系统概要设计.....	26
5.2 系统详细设计与实现.....	27
5.2.1 主页.....	27
5.2.2 关系检索.....	27
5.2.3 关系全貌.....	28
5.2.4 问答系统.....	28

5.2.5 修改资料.....	29
5.3 系统展示.....	29
5.3.1 开发环境说明.....	29
5.3.2 系统功能展示.....	29
第 6 章 系统测试.....	34
6.1 功能性测试.....	34
6.2 非功能性测试.....	35
第 7 章总结与展望.....	36
7.1 总结.....	36
7.2 展望.....	36
参考文献.....	38
致 谢.....	40
附 录.....	42
外文文献原文.....	42
外文文献译文.....	64
任务书.....	72
开题报告.....	74

致 谢

行文至此，落笔为终。四年的学习生活，始于金秋，终于盛夏，简短的八个字很难去描述这一段很长的路。至此，学生生涯也接近尾声，我也曾欢声笑语，也曾跌跌撞撞，纵有千般不舍，也即将奔赴下一场山海。在人生中的青春年华，收获了最宝贵的财富，感谢所有相遇！

岁月如歌，师恩难忘。首先感谢我的毕业设计指导老师，从论文的选题到提纲设计，再到最终定稿，都离不开老师的帮助和指导。老师为人谦和，知识储备渊博，实践经验丰富，治学态度严谨。感谢您的耐心和包容。在写毕业论文的几个月中，给予我莫大的帮助，在一次次组会和交谈中，不仅鼓励我不断探索新知识，激发更好的灵感，同时还让我学会了如何以积极乐观的态度去审视世界，这些都对我今后的道路起到重要作用。最后在这里感谢曾经所有的授课老师，使我能够学一点知识、明一丝道理、寻一条大道。

山河不足重，重在遇知己。感谢我的朋友们，你们在生活和学业上给我很多的帮助，在科研遇到难题时，为我指点迷津；在忙碌之余，会陪我一起放松休闲；在心情烦闷之时，会陪我彻夜长聊。你们的陪伴与支持使我的研究生生活充实且快乐，在此衷心祝愿我的室友和朋友们前程似锦。

父母之爱子，则为之计深远。感谢我的家人，在过去的二十多年里，你们是我的港湾，替我遮风挡雨；又是我启蒙老师，教会我为人处事的道理；更是一束光，为我指亮前程，亦照亮归途。养育之恩，无以为报，我会继续努力，让你们过上更幸福的生活。

愿你们永远健康快乐，万事顺遂。

附 录

附录包含外文原文（近 5 年与毕业设计（论文）题目相关的期刊论文）、外文译文（不少于 2000 字）、任务书、开题报告、设计书、图表、程序等（附录部分装订顺序按照此顺序进行）。

外文文献原文

A Review of Knowledge Graph Research in Military Domain

Xudong Guo , Linxiu Chen , Yuan Li , Qinglin Wang , Weili Guan

Abstract. With the development of information technology as well as the continuous improvement of military informatization level, the complex and changing battlefield situation put forward a huge challenge on the military's massive data processing capability. This paper presents the technical framework related to the construction of knowledge graph in military field, summarizes different aspects of the application of knowledge graph in military field, and reveals the directions for improvement in view of the current situation of the application of knowledge graph in military field.

Keywords: Military Domain, Knowledge Graph, Ontology

1. INTRODUCTION

In the context of modern intelligent warfare, the widespread application of battlefield information technology and sensing technology has led to a dramatic expansion and interconnection of battlefield space and an explosion of data. Intelligence, reconnaissance, surveillance and massive data have "4V" characteristics: large scale (Volume), very fast change (Velocity), a variety (Variety), and value (Value). The huge amount of data makes the battlefield situation assessment and analysis have far needs exceeded the traditional computer-aided assessment and decision processing capabilities based on manual experience and cognition[1]. It is desirable to introduce powerful knowledge description and management methods into military application.

The knowledge graph originated from the concept of semantic web and was first proposed by Google in 2012[2]. Knowledge graphs integrate technical approaches

such as databases, information retrieval, data mining, natural language processing, and graph computational analysis[3]. Knowledge graphs can integrate knowledge of a subject area and provide services such as visual presentation. The application of knowledge graphs has significantly improved the information retrieval capability and demonstrated the powerful function of knowledge graphs in information integration.

As the forefront of science and technology in the military field , the construction of proprietary knowledge graphs can largely improve the efficiency of operations, mainly in the following aspects[4]: knowledge graphs can integrate a large amount of scattered and isolated intelligence; knowledge graphs can allow computers to understand the semantics of intelligence; knowledge graphs can correlate a large amount of intelligence at the semantic level; and knowledge graphs can mine new knowledge in correlated intelligence. Therefore, it is necessary to apply knowledge graphs to the military field.

A knowledge graph consists two levels: the schema layer and the data layer[5]. The schema layer is used to regulate the data layer and to formulate the rules of data types and hierarchies followed by the data layer, and it is the core of the knowledge graph. The schema layer usually uses an ontology library to manage the schema layer of the knowledge graph. The data layer stores the knowledge specific to the knowledge graph and can be seen as an instantiation of the schema layer.

In this paper, the construction of knowledge ontology in military domain is introduced, and the knowledge service of knowledge graph in military domain in three directions, and conclusion with a summary and outlook.

2. CONSTRUCTION OF KNOWLEDGE ONTOLOGY IN MILITARY FIELD

Knowledge graphs are constructed in two ways, the top-down and bottom-up approaches[6]. The top-down approach usually starts by defining a data schema for the knowledge graph, which is constructed from the topmost concept and gradually refined downward to form a well-structured taxonomic hierarchy, and then entities are added to the concept. The bottom-up approach organizes entities

外文文献译文

军事领域知识图谱研究综述

郭旭东、陈林秀、袁莉、王庆林、关伟丽

摘要随着信息技术的发展和军事信息化水平的不断提高，复杂多变的战场形势对军队海量数据处理能力提出了巨大挑战。本文提出了构建军事领域知识图谱的相关技术框架，总结了知识图谱在军事领域应用的不同方面，并针对知识图谱在军队领域应用的现状，揭示了改进方向。

关键词：军事领域，知识图谱，本体

1.简介

在现代智能战争的背景下，战场信息技术和传感技术的广泛应用导致了战场空间的急剧扩展和互联，数据爆炸。情报、侦察、监视和海量数据具有“4V”特征：规模大、变化快、多样性和价值。庞大的数据量使得战场态势评估和分析的需求远远超过了传统的基于人工经验和认知的计算机辅助评估和决策处理能力[1]。将强大的知识描述和管理方法引入军事应用是可取的。

知识图谱起源于语义网络的概念，由谷歌于2012年首次提出[2]。知识图谱集成了数据库、信息检索、数据挖掘、自然语言处理和图计算分析等技术方法[3]。知识图谱可以集成主题领域的知识，并提供视觉呈现等服务。知识图谱的应用显著提高了信息检索能力，展示了知识图谱在信息集成中的强大功能。

作为军事领域的科技前沿，专有知识图谱的构建可以在很大程度上提高作战效率，主要体现在以下几个方面[4]：知识图谱可以集成大量分散和孤立的情报；知识图谱可以让计算机理解智能的语义；知识图谱可以在语义层面上关联大量的智能；知识图谱可以挖掘相关智能中的新知识。因此，有必要将知识图谱应用于军事领域。

知识图谱由两个层次组成：模式层和数据层[5]。模式层用于规范数据层，并制定数据层所遵循的数据类型和层次结构的规则，它是知识图谱的核心。模式层通常使用本体库来管理知识图谱的模式层。数据层存储特定于知识图谱的知识，并且可以被视为模式层的实例化。

本文介绍了军事领域知识本体的构建，并从三个方面对军事领域知识图谱的知识服务进行了总结和展望。

2.军事领域知识本体的构建

知识图谱有两种构建方式，自上而下和自下而上[6]。自上而下的方法通常首先为知识图谱定义一个数据模式，该模式从最顶层的概念开始构建，并逐步向下细化，形成一个结构良好的分类层次，然后将实体添加到概念中。自下而上的方法归纳组织实体，形成底层概念，然后逐渐向上抽象，形成顶层概念。

军事领域是一个典型的特定领域，具有明确的概念层次，然而，部分专业知识存在于领域专家的脑海中，可能无法从数据中归纳和抽象地获得。由于军事数据的保密性和军事装备的特殊性，许多文件和信息对知识图谱的建立具有绝对的阻碍作用，导致数据结构较少，数据收集困难。因此，该领域的知识图谱在构建之初更适合采用自上而下的方法。在知识图谱的基本结构已经建立，其类别节点或关系的数量已经达到一定规模后，可以使用自下而上的方法来扩展图的实例和属性数据[6]。通过以上分析，鉴于军事装备知识图谱的特殊性，首先需要构建本体。

2.1. 本体论

本体论最初是一个哲学定义[7]，是对客观事物的本质描述。随着人工智能、知识图谱等领域的发展，本体的概念逐渐得到重视，并在这些领域得到广泛应用。1991年，Neches提出了本体论的定义，“构成所讨论领域词汇的基本术语和关系，以及使用这些术语和关系定义这些术语的扩展的规则”[8]。1993年，Gruber提出“本体是一种共享的、概念化的、形式化的规范”[10]，该定义被学术界广泛接受。

在知识图谱领域,本体是通过抽象和约束领域中不同实体概念、属性及其相互关系而获得的规范,是知识图谱最核心的部分,主要被描述为树结构[11]。本体是结构化知识库的概念模板,通过本体库形成的知识库不仅层次结构更强,而且冗余度更低[12]。构建本体的主要目的是为知识图谱定义一个清晰的数据模式,而正确的数据模式可以确保知识图谱中的实体具有强大的系统关系,这是构建知识图谱的基础[13]。

2.2. 本体构建原则

本体构建是构建知识模型的主要方式,遵循以下五个公认原则[13][14],即清晰与客观、完整性、连贯性、最大单调可扩展性和最小本体承诺原则。

在清晰度和客观性方面,本体论必须能够以客观且独立于上下文的定义有效地传达所定义术语的含义。在完备性上,本体中给出的术语定义是完整的,并充分表达了所描述术语的含义。

本体论定义的公理,以及用自然语言描述它们的文档,应该是一致的。

在连贯性方面,从术语中得出的推论与术语本身的含义是一致的,即它们支持与其定义一致的推理,不会产生矛盾;自然语言中定义的公理和说明的文档也应该是一致的。

在最大单调可扩展性上,向本体添加通用或专门术语不需要修改其现有概念定义和内容,并且支持基于现有概念定义新术语。

关于最小本体承诺,本体承诺应该是最小的,并在尽可能少的约束下处理建模对象。本体承诺意味着就如何以一致和兼容的方式使用共享词汇达成共识。一般来说,本体约定充分满足特定的知识共享需求,这可以通过定义约束最小的公理和仅定义通信所需的词汇来确保。

2.3. 本体构建方法

本体构建主要有三种方法,手动构建、重用现有本体(半自动构建和自动构建[15])。人工构建方法主要通过领域专家确定知识内容和关系,使本体质量和准确性较高。军事领域对知识的完整性和准确性要求很高,因此人工构建方法适用于军事知识本体的构建。采用通用本体手工构建的思想,军事本体可以应用自上而下的手工构建,首先构建军事知识的顶层本体,在此基础上开始构建每个子域的知识框架,逐步完成整个军事知识本体的构建[16]。

国内外比较有影响力的本体手册构建方法有五种[17][18]:骨架法、TOVE法、IDEF5法、本体方法和七步法。与其他方法相比,七步法比较成熟。它包括七个步骤,即确定专业知识的领域和范围,检查重用现有本体的可能性,列出领域中的重要术语,定义类和类层次结构级别,定义类属性,定义属性方面,以及创建实例[19]。然而,七步走的方法也有一些局限性;这种方法没有一个完整的生命周期[20],并且缺乏在开发后期对需求调整或功能添加做出响应的模型修订过程。

2.4. 军事领域本体结构

在军事装备领域,文献[13]将军事装备分为火炮、舰艇和飞行器等八类,并为装备实体定义了杀伤力和最大速度等属性。这种装备知识体系偏向于对军事装备的科学描述,忽视了面向军事作战需求的知识表达。刘等人[21]设计了一种基于需求的螺旋反馈方法来构建军事装备知识模型,定义了一个顶级核心概念——军事装备、四个相关概念(如作战活动)和七个通用概念(如国家),其中装备分为八大类和148个子类。该模型包括202个实体属性、6个实体关系,涵盖5800个武器类型实体、18个作战活动和86个国家。

从以上分类可以看出,军事领域不同任务的本体构建重点各不相同,但基本可以分为以下几个方面。

3. 军事领域知识图谱的构建

基于知识本体,可以构造知识图谱。军事领域知识图谱的构建和应用是一项系统工程,其构建过程分为六个阶段,可称为领域知识图谱全生命周期,其设计的关键技术过程主要包括知识表示、知识存储、知识提取、知识融合、知识进化和知识应用[22]。

3.1. 知识表示

任务书

毕业设计任务书

中国姓氏文化可视化系统的设计与实现

一、毕业设计目的

该设计旨在以直观、生动的方式展示中国姓氏文化的丰富内涵和独特魅力。通过这一系统，用户可以深入了解自己的姓氏起源、发展历程和家族故事，增强文化认同感和归属感。同时，该系统也可作为研究、教育工具，促进中外文化交流与传播，为中华文化的传承与创新注入新的活力。

二、主要内容

1. 数据收集与处理模块：负责从各种来源收集姓氏数据和相关信息，并进行必要的清洗、整理和分类。
2. 数据存储与管理模块：用于存储和管理姓氏数据和信息，提供高效的数据检索和访问功能。
3. 数据可视化模块：利用各种可视化工具和技术，将姓氏数据以直观、生动的方式呈现给用户，如动态图表、交互式地图、时间线等。
4. 查询模块：根据用户输入的姓氏，提供相关的查询结果和推荐信息，满足用户的个性化需求。
5. 系统管理模块：负责对整个系统进行管理和维护，包括数据更新与维护等。

三、重点解决的问题

- 1、任务模块的整体划分
- 2、前端展示界面的设计

3、数据的完整性

四、主要技术指标或主要参数

系统响应时间：一个好的预约系统应该能够快速响应用户的操作，提供流畅的用户体验。

并发用户数：保证一定数量用户同时访问时系统的稳定性和可用性。

可视化效果：系统采用的数据可视化技术需要提供清晰、直观、生动的可视化效果，使用户能够快速获取信息。

五、基本要求

完成软件设计、代码编写与模块测试，形成包含上述功能模块的软件原型产品；

完成 1 篇英文文献翻译；

完成不少于 15000 字的毕业论文一篇

六、其它（包括选题来源）

实践

指导教师： 2024 年 1 月 10 日