



# 基于本体的红枣领域知识智能服务研究

薛原<sup>1</sup>, 冯建中<sup>1\*</sup>, 杨浩<sup>1,2</sup>, 王剑<sup>1</sup>, 高祥伟<sup>2</sup>, 谢能付<sup>1</sup>, 曹丹<sup>3\*</sup>, 白林燕<sup>4</sup>, 陈彩铭<sup>1</sup>

1. 中国农业科学院农业信息研究所, 北京 100081; 2. 江苏海洋大学海洋技术与测绘学院, 江苏连云港 222005; 3. 中国消防救援学院, 北京 102201; 4. 中国科学院空天信息创新研究院, 北京 100094

**摘要:** 针对当前红枣领域知识体系碎片化、更新滞后、知识服务个性化不足、服务质量与效率低等问题, 为实现红枣领域知识的高效组织与智能服务, 本文开展了基于本体与知识图谱的红枣产业领域知识智能服务研究, 首先系统分析了红枣产业信息知识服务过程体系, 构建了红枣领域知识服务框架, 探索了红枣产业领域知识组织方式, 基于领域知识本体与知识图谱开展了红枣领域知识智能服务机制与服务模式研究; 进而通过对红枣产业种植经营及相关领域多源异构语料进行系统整合与处理, 构建了规范化的多维领域知识本体库, 同时利用 Bert-BiLSTM-CRF 混合模型, 构建了具有良好完备性、消歧容错与关联推理能力等优势的红枣产业领域知识图谱。在此基础上, 基于本体和知识图谱及领域知识服务协同优化原理, 研发设计了一套融合领域专用模型与大模型的红枣产业领域知识智能问答系统, 并且探讨了智能服务评估的优化路径, 以保障服务的高质量和可信度。本文对提升红枣领域知识智能服务的实用性与适用性、推动其产业数字化与智慧化转型升级, 探索可行的技术路径而提供了重要的实践参考。

**关键词:** 领域本体; 红枣; 知识图谱; 智能服务; 知识问答

## 1 引言

随着数智经济的蓬勃发展, 知识越来越成为重要的资源, 知识服务亦构成社会、经济、科技服务等的重要组成部分。知识服务理念源自 20 世纪 90 年代, 旨在通过对用户的知识需要和问题环境的分析, 将信息析取、重组、创新、集成, 并针对性地解决用户问题的高级阶段信息服务<sup>[1-2]</sup>。本质上是针对服务对象需求而导向式地提供个性化、精确的显性或隐性知识, 以供服务对象主体有效判断、决策或解决问题等<sup>[2]</sup>。其中, 在知识服务的过程中, 根据需求和需求的环境, 建构知识(尤其隐性知识)流动与转化的机制及其激励的策略措施, 以实现知识共享与重复利用的目标、发挥知识服务的最大效应。所以, 面向知识服务的知识组织体系的建构是一个解决用户所需的特定知识的过程, 涉及数据到知识的转化、知识到再生知识的增值过程, 即数据知识化、知识有序化和知识服务化<sup>[3-4]</sup>。

而知识服务的应用场景千差万别, 致使知识种类繁多, 分类体系庞杂, 知识服务形式复杂多样, 服务质量和效率常有待提高(知识服务质量标准随意性大、缺乏科学的一致性)等<sup>[5-7]</sup>。

本体概念从哲学引入到知识资源组织与表达<sup>[8]</sup>, 它用于对客观世界的存在进行系统化描述, 实现相关知识的重用和交互。因此, 利用本体抽象领域知识概念、关系、属性以及外延约束规则等, 形式化、规范化描述多种资源而内蕴支撑有效的语义层面信息组织、统一异构问题和领域各种信息知识资源共享与服务等<sup>[9-11]</sup>。

目前, 本体技术已经广泛地应用于知识工程的信息抽取、学习表达、知识组织与建模等, 满足信息知识服务的高效、精准、适应性智能化的需要。在农业知识工程与信息知识服务领域, 虽然研究构建与应用农业本体体系及各类农业领域知识本体理论方法与技术相对较晚, 但到目前国内外关于运用本体技术理论

收稿日期: 2025-09-07; 录用日期: 2025-11-28

基金项目: 河南省重点研发专项项目(241111211100); 国家科技创新 2030 新一代人工智能重大项目(2022ZD0119501); 新疆生产建设兵团(重点领域)科技攻关计划(编号: 2019AB002)。

联系方式: 薛原, E-mail: 1040697691@qq.com. 通信作者冯建中, E-mail: fengjianzhong@caas.cn; 通信作者曹丹, E-mail: caodan@cfri.edu.cn.

研究与实践（尤其在农业知识问答、知识智能搜索、知识管理与个性化服务等应用场景方面），仍是领域本体论研究的热点之一<sup>[12-15]</sup>，为学术界和业界广为重视且取得了诸多理论研究与应用实践成果。

早在2000年伊始，联合国粮农组织（FAO）实施了AOS（Agricultural Ontology Service，农业本体服务）研究计划，促成基本构建和完善了农业知识本体体系，奠定了全球农业本体及农业领域知识本体服务基础，成为农业信息知识服务发展的里程碑<sup>[16-18]</sup>。我国在农业本体论研究及技术应用，也取得了一系列卓然成果，以促进农业信息知识服务。例如，谢能付等<sup>[16]</sup>探讨了构建农业知识本体体系和农业知识领域本体以及获取农业知识的方法，辨析了构建一套完整本体的评价标准，服务于知识有效的共享与重用和Web服务的开发等。熊大红<sup>[19]</sup>开展了基于本体的农业知识共享服务模型和农业知识共享服务的激励机制研究。提出了构建农业本体的基本原则、流程和方法；面向农业知识服务体系，提出了交互式农业知识服务模型，并探讨了个性化知识信息推送和自助式知识概念服务等功能模式。刘琮等<sup>[20]</sup>提出果品领域本体库的基本构建原则和方法，探讨了大规模领域知识本体构建的有效途径。刘超<sup>[21]</sup>通过构建茶学领域本体，根据需要对相关知识进行动态映射、抽取、组合融合生成领域虚拟本体云，以实现农业领域多本体协同服务。朱越<sup>[22]</sup>通过对小麦病虫害命名实体识别技术进行研究，提出了一种基于小麦病虫害本体的推荐模型，为领域知识等有效传播提供支持；沈金梦<sup>[23]</sup>采用本体与图谱映射的方法，实现了水稻病虫害本体的构建和智能诊断系统的开发；赵嫦花等人<sup>[24]</sup>探讨了柑桔病虫害自适用本体构建方法，以助力柑桔病虫害高效预警与防治等。

另外，随着大模型技术的发展，本体技术与大模型的深度融合引起各界的广泛重视与新思考<sup>[25-26]</sup>，为农业领域的知识图谱动态构建、知识服务、智能决策支持等场景的协同机制，带来新的（如“大模型驱动+本体约束”的新型知识工程）范式与路径<sup>[27-29]</sup>。

红枣作为我国重要的特色林果经济作物，分布广泛，尤其在西部新疆地区红枣产业在农业经济结构中占有极为重要的地位。但目前红枣领域农业科技及其产业体系信息知识的组织与规范、标准化方面还很不完善，信息知识获取途径落后，农业科技信息知识传播共享困难、服务水平低<sup>[30-32]</sup>。至此，本研究围绕着当前红枣领域知识服务的需要，开展了基于本体的红

枣产业领域知识智能服务研究，而分析了红枣产业信息知识服务过程体系，构建了红枣领域知识服务基本框架，利用本体技术对红枣产业领域知识组织方式进行了探索，且基于领域知识本体开展了红枣领域知识智能服务机制与服务模式及应用实践案例研究，促进红枣及相关领域信息知识服务智能化发展，解决信息服务“孤岛”等问题，具有重要的科学与应用实践价值。

## 2 模型与方法

基于领域知识特征本体建模和迭代/优化嵌入知识图谱算法的深入协同运用，极大地赋能提升自动识别、关联、分析庞杂而海量的大数据，从中挖掘出有价值的信息，为服务对象/用户高效、精准地提供信息服务，已成为当今知识服务的核心目标之一，亦充分地展现出信息知识服务的智能化特点<sup>[33-34]</sup>。知识图谱作为一种有效可视呈现知识域的知识关系结构，具备很好的信息组织和推理能力，使得知识图谱现已成为知识领域智能服务的重要基础设施之一。

### 2.1 红枣产业领域知识服务架构

由于知识服务目的是通过特定的技术途径，按需将一般或某种领域知识内容与用户对象建立关联，使知识定向传递给目标用户对象，因此建立领域知识的服务体系，则为领域知识服务奠定基础。利用CommonKADS和多角度知识建模方法<sup>[35-37]</sup>，可构建出领域知识服务一般框架，如图1所示。领域知识（DK）是应用领域的概念、实体和内容等知识，推理知识（IK）即处理不同类型的推理，任务知识（TK）是指明了任务的目标，它们共同构成了领域知识服务的应用范畴。而问题解决方法（PSM）描述任务解决方法且确定递归分解的子任务及其执行的顺序，策略知识（SK）说明任务执行计划规则，二者构成了领域知识服务的问题解决途径。语境知识（CK）是描述影响知识服务的理解和使用的情境、背景和机构等因素的知识，并且在统一规范标准体系的规约约束下和运用智能协作模式实现领域知识智能服务目的。

红枣产业领域知识服务体系（图2），数据/信息知识获取环节是领域知识服务的基础。针对知识组织与构建环节，即数据、信息到知识的转化以及知识的有序化、再生等过程，首先利用本体建模方法确定红枣产业领域本体的核心类以及类的层次结构、核心属性要素，基于类的层次结构、对象属性将类进行相互

关联或映射，同时充分协同利用多源数据/信息和知识（如涉及的相关外部实例知识、网络信息）弥补领域信息知识的不足、更新不及时等缺点，动态构建红枣产业领域知识本体结构体系，并利用分类、归并与消除歧义、关联等途径<sup>[10]</sup>，支撑构建红枣产业领域本体应用的知识图谱和知识库，从而有效地实现异构知识的动态组织和领域知识语义结构化。而知识服务环节，则基于红枣产业领域知识库和知识图谱的推理

机制等，面向用户对象的需求和应用场景，实现领域知识服务的各类业务（包括领域元知识控制服务）。例如，大规模的知识查询、过滤、发现/挖掘等；在知识问答智能服务中，利用知识图谱的推理服务模式，解析用户自然语言问句，理解用户意图，精准作答问题，而缩短红枣生产中专业知识不足的从业人员与相关农业领域专家知识的差距，及时有效地解决生产中的问题。

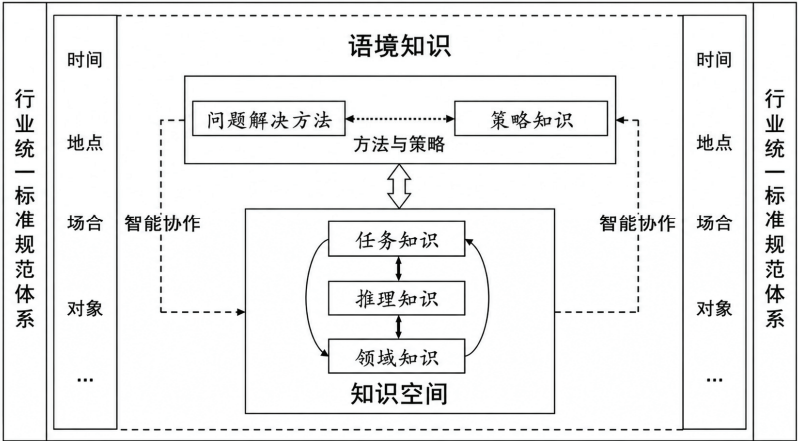


图 1 领域知识服务一般框架

Fig. 1 General framework of domain knowledge services

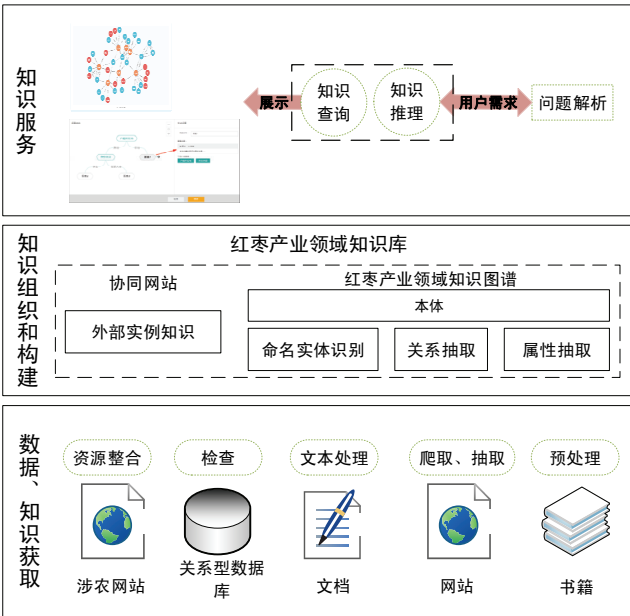


图 2 红枣产业领域知识服务体系

Fig. 2 Knowledge service system for the jujube industry

## 2.2 红枣产业领域知识智能服务传递机制

领域知识智能服务规约了领域知识服务的模式、方法及策略等趋向智能化,因此面向高质量领域知识服务的需要,构建高效、灵活的个性化领域知识智能服务的传递机制尤显重要。红枣产业领域知识服务传递机制(图3)。通过自(反)馈,用户(服务使用者/主体)准确理解领域知识、辨析抽取需求,优化传递请求,并利用服务管理(如评估、监测)和协作机制,有效实现领域知识服务的发现与检测;在服务的生成过程中,按需分析挖掘或提炼领域知识,生成

知识服务基本单元,并且(可利用知识图谱、模型自主训练等途径<sup>[38-39]</sup>实现不同服务单元的协同整合或聚合,呈现在不同的应用场景下不同用户易于接受的展现形态(及交互性界面)<sup>[40-41]</sup>,支撑个性化服务的提供,例如智能问答、问题诊断、智能推送或推荐等。其中,反馈控制机制构成了知识服务传递机制的重要组成部分,它能够有效地保证或提升服务质量和效能,以及用户的服务体验等。此外,为保证服务使用者和提供者的权益等,领域知识服务传递过程中的安全性和隐私保护等问题,也是必须要考虑的重要环节<sup>[42]</sup>。

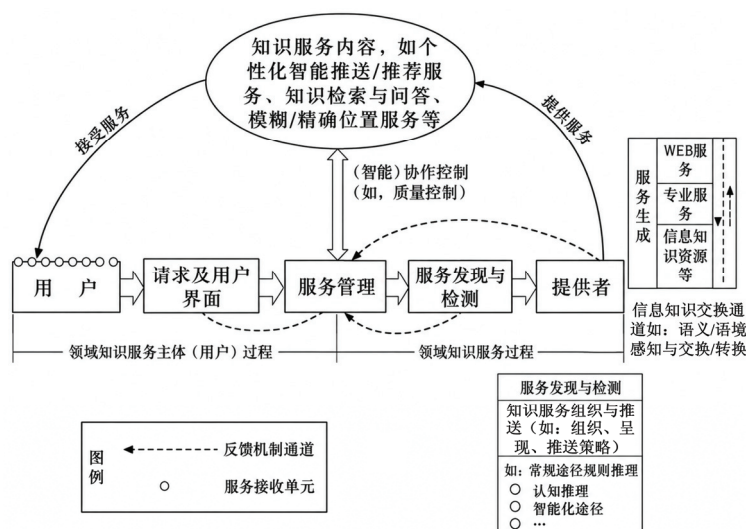


图3 红枣产业领域知识服务传递机制

Fig. 3 Knowledge delivery mechanism for the jujube industry

## 3 案例与分析

### 3.1 红枣产业领域本体组织

红枣产业本体模型组成了红枣产业领域知识图谱的模式层,它通过对语料知识结构的提炼,建立了领域知识图谱内部约束和规则,规范了实体、关系、实体属性、属性值之间的联系,是构建知识图谱的核心内容,也是实现智能服务的重要支撑。

#### 3.1.1 知识来源

本研究知识来源依据分为两大部分:(1)以张新华等人编著的《枣标准化生产》为主,李新岗主编《中国枣产业》等书籍资料;(2)以中国红枣网(<https://www.hongzaowang.com/>)为主等涉及红枣信息的门户网站。

其中,书籍作为非结构化数据源通过OCR识别后<sup>[43]</sup>,其本身具有的“标题”、“副标题”、“小节”等结构特征,通过转换处理且结合人工去噪,最终可以得到结构清晰的无噪声(半结构化)文本语料文件。《枣标准化生产》主要结构分为十部分:生产标准、苗木繁育阶段操作、建园定植阶段操作、“土、肥、水”需求及管理操作、修剪整形阶段操作、花果管理、生态枣园建设规范、病虫害管理、采收分级销售阶段事项。

中国红枣网的技术栏目关于红枣知识主要包括:病虫害防治、种植栽培、储藏保鲜、红枣加工、枣论文库等几个部分。利用Python Crawley爬虫框架对数据进行爬取和处理,以及人工审核二次去噪,得到语料片段。

### 3.1.2 红枣产业领域本体构建

本体构建方式可分为自底向上、自顶向下及二者混合三种方式<sup>[44]</sup>。自底向上通常从具体实例出发，逐步抽象形成本体概念体系（如基于数据驱动，从开放数据源中提取实体、属性和关系）；自顶向下一般先定义顶层概念框架，再逐步细化具体概念和关系，形成结构良好的概念层次树，常适合垂直领域本体构建；混合方式则结合前两者优势，但也存在开发成本高、需协调前两种方式差异等缺点。鉴于红枣产业领域知识具体情况与特性，本文采用自顶向下的本体构建方式。

红枣产业本体作为具有鲜明领域特点的本体，其构建过程与红枣生长周期密不可分，大致可分为：种植准备阶段、育苗繁育阶段、规模种植阶段、采收储存销售阶段。参考农业过程本体及相关文献<sup>[45-46]</sup>及红枣的生长发育阶段，本文将红枣产业过程本体分为苗木繁育、建园生产两个部分。且将贯穿于红枣作物生命周期的病虫害类、土肥水需求类作为独立单元。

利用 Protégé 本体构建工具，通过系统抽象与形式化定义，形成红枣产业本体共计 8 个一级类、10 个二级类、6 个三级类、14 个四级类（图 4）。其中，领域本体核心一级类包括：①枣类（jujube）。

（面向红枣及红枣产业）枣类主要描述红枣基本信息知识内容；②枣树类（jujube tree）。枣树作为枣果生长发育的植物母体，及其与生态环境等关联特性，使枣树类是一个很好的语义建模对象，主要用来描述枣树基本信息知识内容；③病虫害类（plant disease and insect pest）。病虫害描述危害红枣及枣树的病害与虫害，因此根据致病源差异，将病虫害类分为病害（disease）和虫害（insect pest）两个子类；④环境类（environment）。定义红枣生长过程中所依赖的土、肥、水等环境因素，形成系统化栽培信息知识部分；⑤地区类（region）。主要描述红枣分布区域和病虫害发生地区等；⑥生产过程操作类（production process operation）。主要描述定义在红枣生产过程中，种植业从业人员所需农事/农业操作信息知识，分为土壤操作类、肥力操作类、水分操作类、苗木繁育类、建园生产类 5 个二级类，其中根据不同作物生长阶段，苗木繁育类、建园生产类又分为苗木繁育前期、苗木繁育中期、苗木繁育后期及建园栽培、（建园）生产中期、

（建园）生产后期 6 个三级类，进一步划分为苗木选种、种子处理、播种、嫁接、播后管理、苗木出圃、定植、建园、栽后管理、整形修剪、采收分级、贮藏、运输、销售 14 个四级类，对红枣生产不同阶段及相关方面所需操作知识进行描述。⑦农药类（pesticide）。主要描述防治不同病虫害所需农药药剂；⑧政策类（policy）。主要描述涉及红枣相关政策法规。

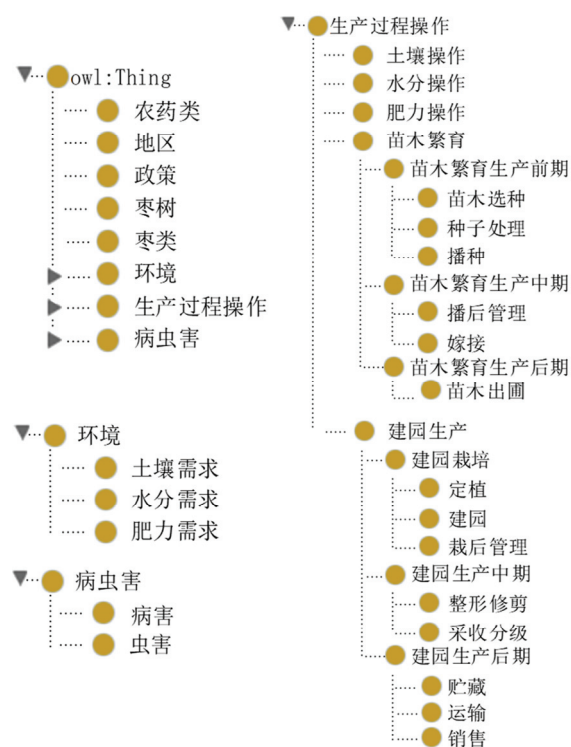


图4 枣业领域本体组织结构

Fig. 4 Ontological structure of the jujube industry domain

为明确定义、补充本体内部属性，完善、增强本体知识描述，本体内部属性可分为对象属性和数据属性两种（其中，对象属性描述本体类实例间的语义关系，为基于描述逻辑的知识推理提供结构化基础；数据属性是通过限定值域来描述本体类实例的特定特征，其值可以是文本、数值等类型，Protégé 等工具则通过集成 XSD 标准数据类型而支持实现丰富的数据属性定义与约束供使用<sup>[47]</sup>。在对语料文本进行分析后，参考《农业科学叙词表》的“用、代、属、参、分、族”叙词结构关系，本文总结定义、补充了 13 个一级对象属性（表 1）。

针对红枣产业本体的 8 个一级类，则总结定义、复用了共计 8 个一级数据属性、61 个二级数据属性。详细内容：①关于枣类，定义了 15 个二级数据属性(表 2)。②关于枣树类，复用、定义了 4 个数据属性：树体描述枣树体型，树姿描述枣树的姿态，树冠描述树冠形状；果实生长期描述枣树产果的生产周期。③关于病虫害类，复用、定义了 13 个二级数据属性(表 3)。④关于环境类，定义了 5 个土壤数据属性：PH 值、含水量、温度、湿度、通气性；6 个肥料数据属性：肥料名称、肥料种类、肥料优点、微量元素、施肥时期、施肥量；3 个水分数据属性：水质、灌溉时期、灌溉量。⑤关于地区类，定义 2 个数据属性：地点、地区名称。⑥关于生产过程操作类，定义了 4 个数据属性：操作名称、操作内容、使用场景、注意事项。⑦关于农药类，定义了 5 个数据属性：农药名称、剂型、施用方法、防治对象、作用特性。⑧关于政策类，定义了 4 个数据属性：政策内容、时间范围、政策影响、级别。

表 1 枣领域本体一级对象属性  
Table 1 Primary object properties of domain ontology for jujube

| 序号 | 对象属性名称 | 定义域 (Domain)   | 值域 (Ranges)    | 定义 (definition) |
|----|--------|----------------|----------------|-----------------|
| 1  | 调控优化   | 整形修剪           | 枣树             | 描述枣树整形修剪操作      |
| 2  | 限制     | 政策             | 枣              | 描述政策对红枣限制效果     |
| 3  | 虫害药剂   | 虫害             | 农药             | 描述用于虫害防治的农药药剂   |
| 4  | 病害药剂   | 病害             | 农药             | 描述用于病害防治的农药药剂   |
| 5  | 分布的地区  | 枣              | 地区             | 描述枣分布的地区        |
| 6  | 危害     | 虫害、病害          | 枣              | 描述虫害、病害对枣危害     |
| 7  | 发现地区   | 虫害、病害          | 地区             | 描述虫害、病害的发现地区    |
| 8  | 存在时期   | 水分操作、肥力操作、土壤操作 | 建园、整形修剪、嫁接、定值等 | 描述土肥水操作对存在的时期   |
| 9  | 完成世代在  | 虫害             | 枣              | 描述虫害完成世代所在的作物   |
| 10 | 影响     | 土、肥、水需求        | 枣              | 描述土肥水对枣的影响      |
| 11 | 有寄主    | 虫害             | 枣              | 描述虫害的寄主枣品种      |
| 11 | 有寄主    | 虫害             | 枣              | 描述虫害的寄主枣品种      |
| 12 | 有症状    | 枣树             | 病害             | 描述枣树的病害症状       |
| 13 | 销售地区   | 销售             | 地区             | 描述枣销售地区         |

表 2 枣类数据属性  
Table 2 Data properties of jujube class

| 序号 | 数据属性名称  | 定义 (definition) | 序号 | 数据属性名称   | 定义 (definition)   |
|----|---------|-----------------|----|----------|-------------------|
| 1  | 枣名称     | 描述枣名称           | 9  | 果实形状     | 描述该类枣果实形状         |
| 2  | 枣别名     | 描述枣别名           | 10 | 果实颜色     | 描述该类枣果实颜色         |
| 3  | 价格      | 描述当前某地区枣价格      | 11 | 果重       | 描述该类枣平均果重         |
| 4  | 制干率     | 描述制干品种枣制干率      | 12 | 适应性      | 描述该类枣适应性优点        |
| 5  | 可食率     | 描述枣的可食率         | 13 | 质地       | 描述枣果肉质地           |
| 6  | 含可溶性固体物 | 描述枣包含可溶性固体物含量   | 14 | 维生素 C 含量 | 描述该类枣维生素 C 含量普遍含量 |
| 7  | 味道      | 描述枣的口感味道        | 15 | 酸        | 描述该类枣酸含量          |
| 8  | 品种      | 描述枣所述品种         |    |          |                   |

表 3 病虫害类数据属性  
Table 3 Data properties of disease and pest classes

| 序号 | 数据属性名称 | 定义 (definition) | 序号 | 数据属性名称 | 定义 (definition) |
|----|--------|-----------------|----|--------|-----------------|
| 1  | 世代及发生期 | 描述虫害世代发生期       | 8  | 最早发生时间 | 描述每种病虫害最早发生时间   |
| 2  | 农业防治   | 描述病虫害农业防治办法     | 9  | 物理防治   | 描述病虫害物理防治办法     |
| 3  | 减产数量   | 描述病虫害对枣造成的减产数量  | 10 | 生物防治   | 描述病虫害生物防治办法     |
| 4  | 名称     | 描述病虫害名称         | 11 | 症状描述   | 描述病虫害作物症状表现     |
| 5  | 异名     | 描述病虫害其他名称       | 12 | 药剂防治   | 描述病虫害防治所需药剂     |
| 6  | 寿命     | 描述虫害发生周期        | 13 | 年发生代数  | 描述虫害在不同环境下年发生代数 |
| 7  | 形态特征   | 描述虫害当前形态特征      |    |        |                 |

由于知识组织的本质是构建现实世界与数字系统间的可计算语义映射，因此需通过形式化本体实现精确的机器可理解关联<sup>[48-49]</sup>。随着语义网技术的发展，目前鉴于基于 Web 的本体描述语言 OWL 具有丰富的语义表达能力，且可以同时支持多种语言版本兼容与交互等优点<sup>[50-51]</sup>。因此，本文采用 OWL 作为本体描述语言。其中具体如下：

红枣产业本体的类使用 owl:Class 进行表示，层次关系使用 subClassOf 表示（图 5）。

```
<owl:Class rdf:about="http://www.semanticweb.org/jujube#  
建园">  
<rdfs:subClassOf rdf:resource="http://www.semanticweb.or  
g/jujube#建园栽培"/>  
</owl:Class>
```

图 5 建园栽培类  
Fig. 5 Ontological class of orchard establishment and cultivation

对象属性（Object Property）通过 owl:ObjectProperty 显式声明，则定义域（domain）表示该属性适用的主体类，值域（range）表示属性值所属的类（图 6）。

数据属性（Datatype Property）的定义域（domain）表示属性适用的主体类，值域（range）表示属性值的数据类型（图 7）。

构建的红枣产业本体（图 8），其中 owl:Thing 为所有实例的顶级类；蓝色线表示类包含关系；其他颜色线定义对象属性。

3.1.3 知识处理

在知识图谱的构建过程中，本体概念层级下的知

```
<owl:ObjectProperty  
rdf:about="http://www.semanticweb.org/jujube#危害">  
  <rdf:type  
    rdf:resource="http://www.w3.org/2002/07/owl#TransitivePr  
operty"/>  
  <rdfs:domain  
    rdf:resource="http://www.semanticweb.org/jujube#病害"/>  
  <rdfs:domain  
    rdf:resource="http://www.semanticweb.org/jujube#虫害"/>  
  <rdfs:range  
    rdf:resource="http://www.semanticweb.org/jujube#枣类"/>  
</owl:ObjectProperty>
```

图 6 病害、虫害与枣类间对象属性（危害）  
Fig. 6 Object property between plant disease and pest class and jujube class: Harm

识实体质量是决定图谱质量与性能的关键因素，而完备、规范、关系明确、语义清晰且来源可靠的知识实体是构成高质量图谱的基础要素。由于现实世界的知识可分为：①以数据库、JSON 对象键值对等具有明显结构特征的结构化知识；②以 HTML 网页、XML 文档等为代表的半结构化知识；③以纯文本（如纸质书籍、文献）、多媒体等为主的非结构化知识。依之，针对结构化知识的数据抽取可通过直接映射或规则映射实现源数据与目标数据的高效、准确关联转换；半结构化、非结构化数据（知识），因其缺乏明确的层次结构且实体信息分布杂乱，通常需借助自然语言处理（NLP）技术实现命名实体识别（NER）、关系抽取及属性抽取等任务。

```

<owl:DatatypeProperty
rdf:about="http://www.semanticweb.org/jujube#价格">
  <rdfs:subPropertyOf
rdf:resource="http://www.semanticweb.org/jujube#枣数据属性"/>
  <rdfs:domain
rdf:resource="http://www.semanticweb.org/jujube#枣类"/>
  <rdfs:range
rdf:resource="http://www.w3.org/2001/XMLSchema#string"
"/>
</owl:DatatypeProperty>

```

图7 数据属性（价格）

Fig. 7 Data property: Price

### （1）结构化、半结构化知识抽取

农业领域信息知识涉及学科较多，知识颗粒度呈现深度细化的特征，使得其结构具有高度异质性和碎片化特点，在当前农业信息资源标准化建设尚不完善的情况下，红枣及红枣产业信息知识精准、高效获取面临诸多困难与挑战。为此，本文基于本体需求构建

信息单元，遵循数据库（如建库、建表、存储）范式，利用 ETL 技术对多源数据进行筛选、去重、标准化等处理<sup>[52-53]</sup>，存入 MySQL 数据库而形成结构化知识语料库（图9）。并且，利用 D2RQ(Database to RDF Query) 工具虚拟 RDF 映射：通过定制化修改映射规则<sup>[54]</sup>，将关系型数据按本体结构映射转换为 RDF 三元组，且建立起知识要素间的语义关联，实现从结构化数据到语义化知识的转化。

### （2）非结构化知识抽取

基于非结构化原始语料（如书籍文本）构建的红枣信息知识库，在预处理阶段通常以标题、副标题作为实例，关联文本内容作为属性值。这种粗粒度存储方式易导致深层信息知识（如隐含语义关系、多级属性）被隐藏或忽略，因此需要对话料内容通过细粒度信息抽取途径实现知识深化，而提升信息知识库的机器可读性和精准智能知识服务的能力等。

具体而言，书籍语料经文本化处理，呈现显著的半结构化特征。其中，一、二、三级标题与段落子标题构成显性的树状拓扑层级结构，且同级标题间存在并列、递进等隐性逻辑关联；从语义层面而言，这种层级显然形成“实体-属性-值”三元组（章节标题作为实体、子标题界定属性、段落内容提供属性值。

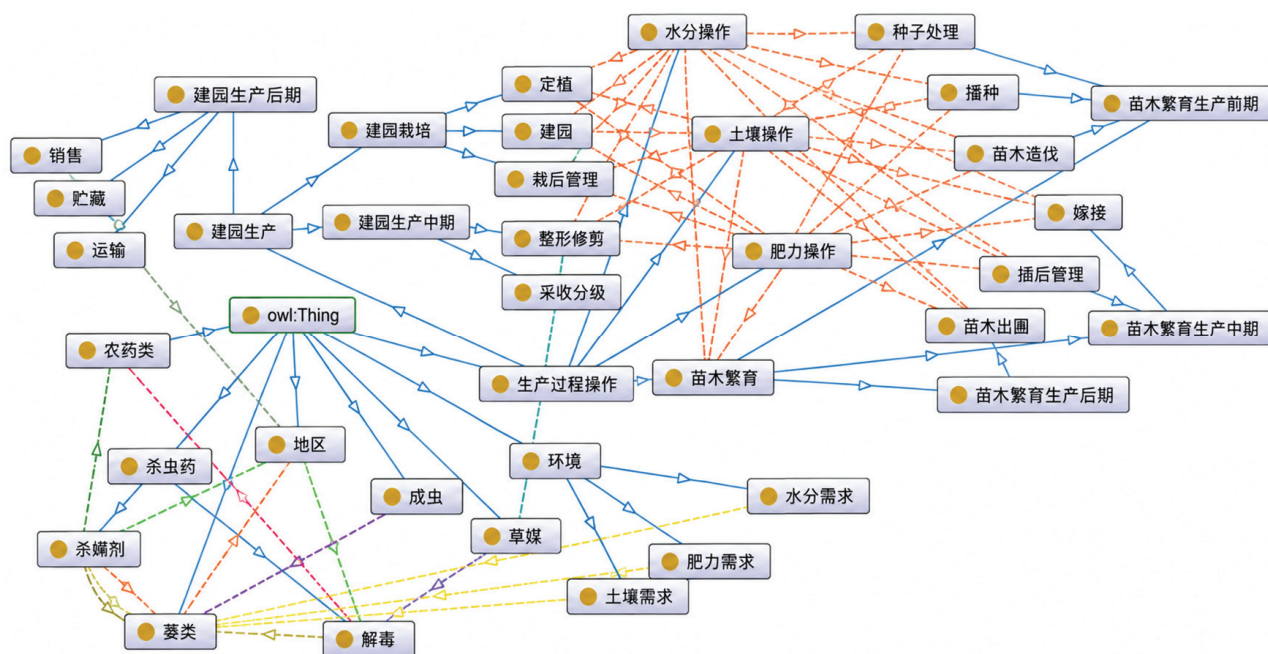


图8 红枣产业本体模型

Fig. 8 Ontology model of the jujube industry

例如,一级标题“新疆红枣”作为实体,二级标题“核心产区”作为属性,三级标题“阿克苏”和“和田”则作为具体属性值)。这种结构化知识图谱构建提供了重要先验信息内容,支持自动化实体关系抽取,以及多粒度知识检索和领域问答系统的优化升级等。

鉴之,本文采用来自转换器的双向编码器表征量(Bi-directional Encoder Representation from Transformers,

BERT)预训练语言模型<sup>[55]</sup>,结合双向长短期记忆网络(Bi-directional Long Short-Term Memory, BiLSTM)和条件随机场模型(Conditional Random Fields, CRF)<sup>[56]</sup>构成端到端深度学习模型,即 Bert-BiLSTM-CRF 模型组合方式,完成实体识别任务。它较其他神经网络模型,能同时捕捉字符级、词级和句级特征等优点,在 NER 等任务中表现优异,应用广泛。

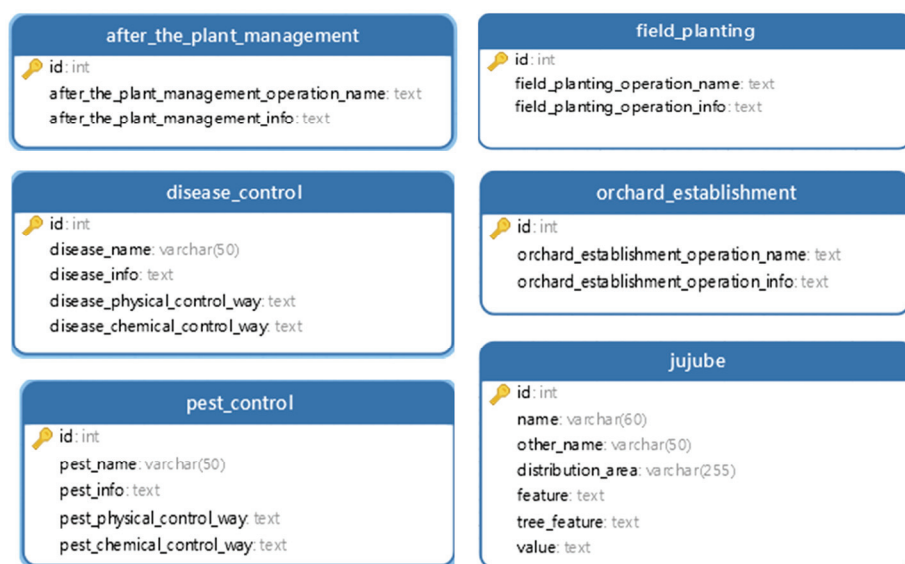


图9 红枣知识数据库(部分)

Fig. 9 Partial view of the jujube knowledge database

由于已构建的红枣产业本体体系对象属性存在两个典型的特征:①核心类之间的对象属性遵循稳定的三元组模式(Entity 1, Relation, Entity 2)(如“红枣品种-生产地-种植技术”,或“红枣品种-适宜种植-土壤类型”);②相关描述文本蕴含/隐含大量非显式语义关系(如“耐旱性”隐含红枣品种与气候的适应性关联),直接应用相关关系抽取方法(如深度学习)易引发关系误抽取和过拟合现象问题。因此,本文采用基于规则匹配的混合式关系抽取方案,即假设命名实体识别后存在实体对(Entity 1, Entity 2),通过预设关系集 $\{(Entity\ 1, Relation, Entity\ 2), \dots, (Entity\ n_1, Relation\ n, Entity\ n_2)\}$ 匹配得到对应关系;如果实体存在,但关系无法匹配,则存入候选关系集,后期进行相似度计算以补充预设关系集<sup>[57-58]</sup>。此外,当语料规模有限但结构规范、清晰,则采用专家人工标注补充实体属性字段而保障其良好的准确性(且显著优于自

动化采集方式)<sup>[59]</sup>。

通过知识抽取、数据填充等步骤,构建出红枣产业知识图谱且导入 Neo4j 图数据库。例如,以“金丝小枣”为核心节点进行检索查询,可以可视化呈现出该品种与产地、种植栽培、病虫害等相关的实体间的多维关联关系(图10),服务种植生产、产业分析等智能信息知识服务及其它相关应用。

## 3.2 知识智能服务与评估

### 3.2.1 知识智能问答服务

智能问答作为本体和知识图谱的重要应用领域,它力求高效、精准、便捷地为用户提供个性化知识服务,实现知识服务的智能化。其中,机器的语义理解、知识融合和推理,尤其由于数据的稀缺性和不完整性,如何利用有限的数据进行模型训练和知识抽取;并提供精准的有针对性的个性化知识回答和信息知识推送/推荐等,是目前智能问答知识服务的关键问题和难点<sup>[15,60-61]</sup>。

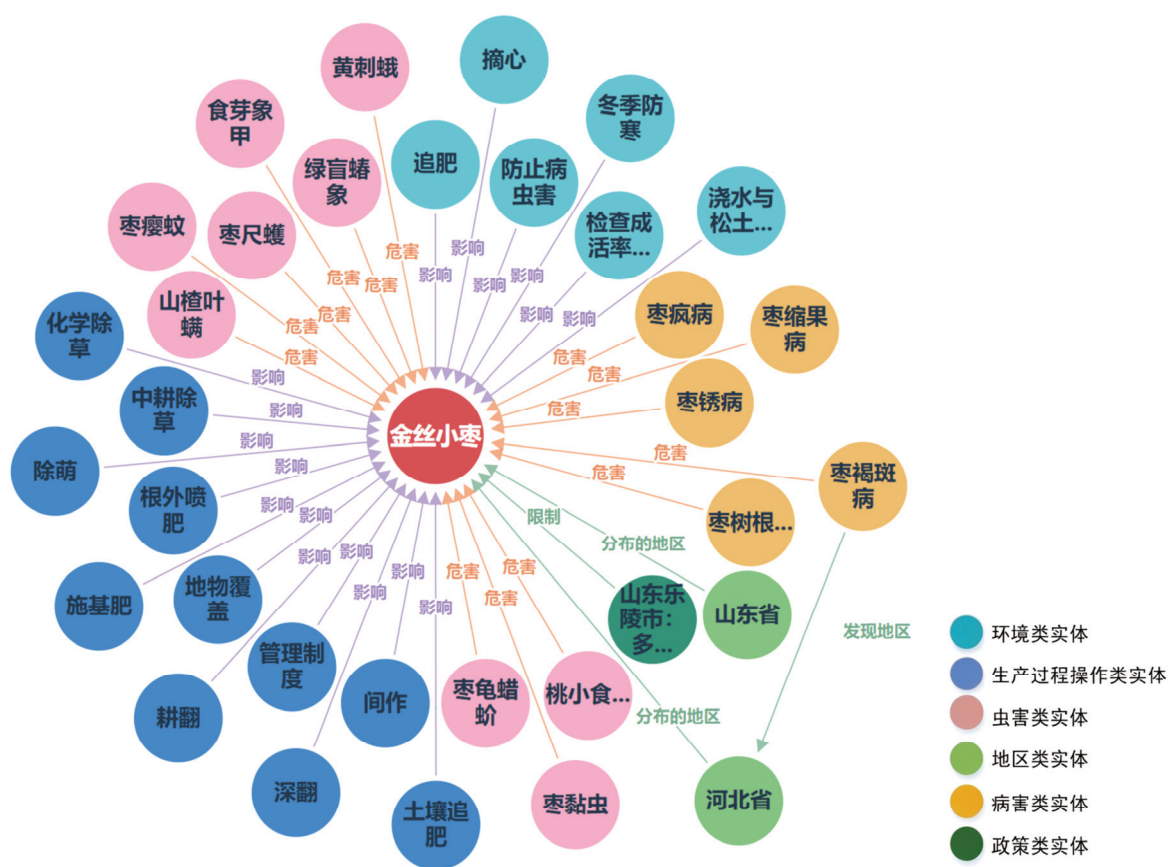


图 10 “金丝小枣”实例关系（部分）

Fig. 10 Entity relations of the Golden Silk Jujube (Partial)

为了提升新疆红枣产业活动智慧化和社会化服务智能化水平,面向南疆红枣领域知识智能服务的实际需要,研发红枣产业领域知识智能问答系统:基于领域知识服务串并行协同优化原理,以及用户使用便捷和系统问答具有可扩展性,系统采用 B/S 协同服务架构,捕获用户的自然语言问题输入和回答反馈生成,如图 11 所示。其中,利用大模型(LLM)功能解析用户问题,通过预设的动态规则集<sup>[62-63]</sup>进行问句意图匹配、判识及问题分类,而后在 Web 后端进行业务分发。(限定域)站内部问题,协同利用领域专用模型 BERT-BiLSTM-CRF 神经网络模型和大模型本地化迭代处理等,进行语句理解、语句分析(如精准地实体识别、意图分析、语义增强处理等);对结构化常规查询,通过预构建的 Cypher 模板库实现高效映射,生成 Cypher 语句,而对复杂开放性站内部问题,则利用大模型的涌现求解能力<sup>[64]</sup>动态生成 Cypher 语句,然后依之来检索、查询知识库等获取信息,最后以回

复填充模版或利用大模型检索/查询增强的方式生成回答语句,以完成语句的求解、生成。另外,对于分发至站内查询的问句如果存在未登录状态,则对其缓存和语句的理解、分析等而更新本地知识库。站外开放域问题,基于分发的信息向外部相关网站(如红枣相关网站、数据中心等)进行请求,并将外部大模型通过 API 接口接入,承担开放域语义理解与逻辑推理任务;知识库采用动态加载机制,如整合维基百科、行业数据库等外部结构化资源,且兼顾利用本地知识库等的支持,答案生成则采用双重校验机制(即首先基于大模型生成候选答案,再通过知识库事实性验证)而最后生成问题答案。

### 3.2.2 知识智能服务评估

知识智能服务评估是知识智能服务体系中的重要组成部分,它是通过对知识智能服务质量、功效、满意度等方面进行综合评价的过程,旨在满足用户多样化、个性化、精准化的需求目标<sup>[65]</sup>。对服务质量与功效的

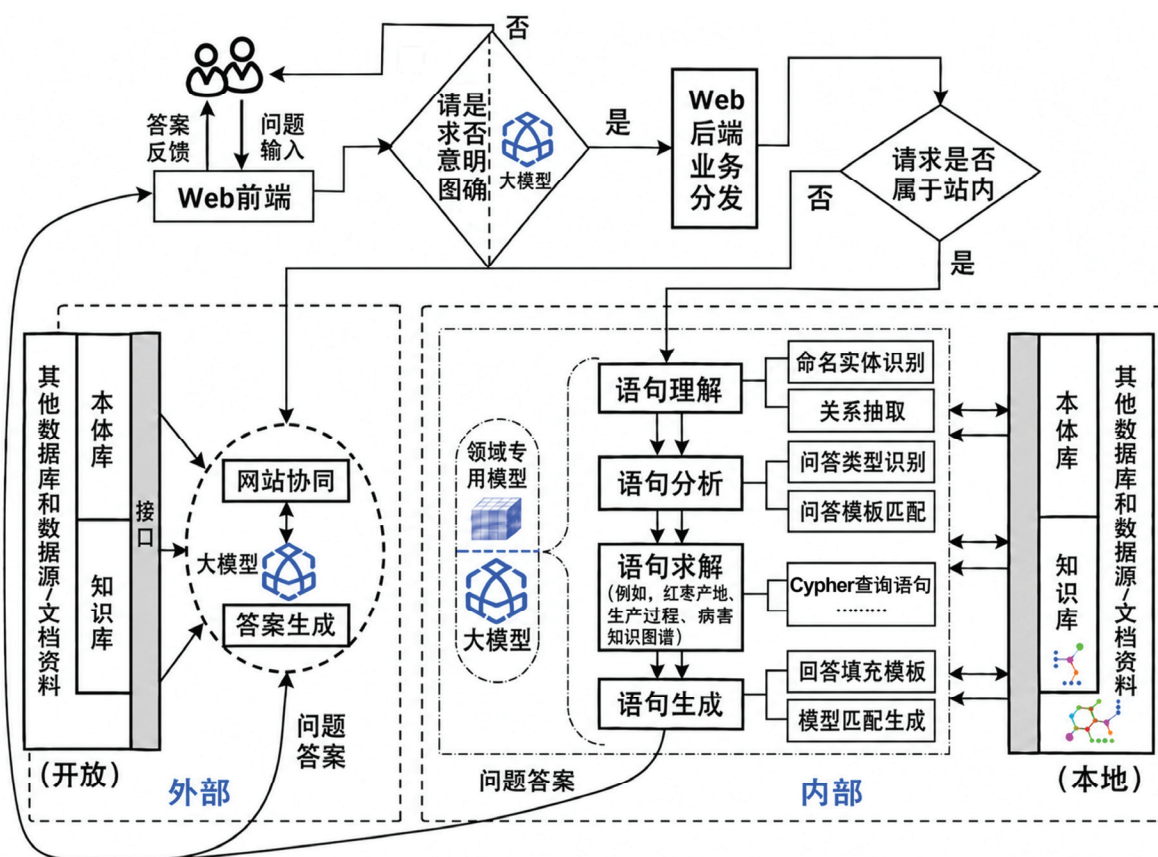


图 11 红枣领域知识服务流程

Fig. 11 Workflow of knowledge services in the jujube domain

衡量与评价是知识智能服务评估的核心内容，因此，针对知识智能服务的复杂性与动态性，目前常常从定性和定量两个途径来进行评估。其中，定性评估与分析，能够充分地利用专家知识和用户体验等，但缺乏系统性和完整性，存在对评估内容不够精准、效率不高等缺陷；定量化评估通过数据统计和分析，能够有效地对服务的各项指标进行精准量化评估，但如何构建全面、客观、实用的评估指标体系，并实现多指标体系的对接和一致性（如消除指标之间的歧义、标定指标的度量标准、实现指标的转换），解决指标之间的解耦与独立性、权衡分配与协同一致等问题，以及如何科学合理地各指标参量化，也是当前对知识智能服务评估需要从理论和实践应用上不断探索、解决的问题。

至于此，本体和知识图谱技术的应用，通过本体语义映射建模与图谱动态关联分析途径，解决多指标体系的连接与集成、指标语义对齐、定性指标量化与

标定等，能够显著地有效提升对知识智能服务多角度、多维度测度与综合评估的能力、提高服务评估的精准度与效率<sup>[66-67]</sup>。然而，针对这一评估的过程，构建优良的本体语义转换体系和高质量的知识图谱模型等尤为重要，以满足评估对知识智能服务的多场景实际的复杂应用需要；而之，目前针对本体和知识图谱（尤其诸多领域类型）的质量评估工作尚无建立统一评价标准<sup>[68-69]</sup>，研究多从多个角度与综合评价的方法提出评估参考框架或体系<sup>[70-71]</sup>，以及在应用实践中，常运用统计学方法人工参与本体和图谱内容的抽样评价等<sup>[72]</sup>。

在领域本体和知识图谱的质量评估方面，更需关注领域本体和知识图谱的一致性、完整性、简洁性、可拓展性等方面的评测，例如实体属性和关系与真实场景一致情况、本体类间关系无冲突且指向逻辑明确、完备的实体覆盖率与关系密度、简洁本体生成与动态可扩展以及知识图谱增量构建等方面评估评价<sup>[73-75]</sup>，

支持保障知识智能服务评估过程健壮、快捷、可靠且评估结果具有良好的准确性、可用性和实用价值。

总之,在实际的知识智能服务评估中,为了提高评估的可用性和实用性,通常基于定性和定量方法协同应用的途径,能够有效、快捷地实现对知识智能服务的评估和分析评价等,以确保知识智能服务在数据准确性、响应速度、用户体验等方面达到预期标准,从而为用户或系统提供更加优质、高效的服务。

## 4 总结与展望

### 4.1 总结

随着信息技术的飞速发展和数智时代的全面到来,传统农业更快向数字农业、智慧农业转型升级,智能知识服务作为核心支撑,发挥着愈来愈重要的基础性关键作用。本研究围绕着当前红枣领域知识组织表达与知识体系的构建及更新、领域知识智能服务机制与多场景实际应用的需要等,开展了基于本体的红枣产业领域知识智能服务研究,主要内容与总结如下:

(1) 分析了红枣产业信息知识服务过程体系,构建了红枣领域知识服务基本框架,利用本体技术建模与知识图谱构建,对红枣产业领域知识组织方式进行了探索,并基于领域知识本体与知识图谱开展了红枣领域知识智能服务机制与服务模式研究;

(2) 通过对红枣产业种植经营及相关领域结构化、半结构化、非结构化等多源异构语料进行系统整合和语义预处理,构建了包含品种特性、栽培技术、病虫害防治等多维度的规范化红枣产业领域知识库及本体库。并在文本有效解析的基础上,基于 Bert-BiLSTM-CRF 端到端混合模型,精准高效地实现对象目标实体的识别、实体关系的抽取等任务,构建了具有较好的完备性与消歧容错多源信息融合、有效的关联推理能力等优势的红枣产业领域知识图谱,以实现红枣及相关异构数据资源的知识统一组织及高效应用;

(3) 面向红枣领域知识智能问答服务的需要,基于本体和知识图谱及领域知识服务串并行协同优化原理,研发设计了协同利用领域专用模型和大模型(本地化)混合处理模式的红枣产业领域知识智能问答系统,以求高效、精准、便捷地为用户提供个性化知识服务。并且,讨论了知识智能服务评估的提升途径,而确保知识智能服务在数据准确性、响应速度、用户

体验等方面达到设定标准与预期目标,为用户或系统提供高质量、可信、满意的服务。

### 4.2 展望

伴随着数智经济的深入发展,红枣领域知识智能服务仍面临诸多的挑战。例如,海量多源异构数据(来源于种植、加工、流通等环节)存在数据分散、格式差异、标准不统一且信息知识获取碎片化等问题,仍很难在短时间内完全解决,多模态数据处理模型与算法仍需要加强;由于动态知识建模技术不足,现有本体构建方法难以适应领域知识的快速更新(如红枣新品种、新工艺等快速迭代),导致知识图谱准确率和覆盖率总体偏低,亟需构建有效的动态本体建模机制,不断完善红枣领域本体体系、更新优化知识图谱,以及更为精准、高效的红枣领域知识智能服务机制和多维度用户画像模型算法,包括与大模型的结合,而提高服务的精准度、时效性、灵活性,加强服务的可解释性,有效解决信息“孤岛”和冗余问题,提升红枣领域知识智能服务的实用性和广泛适用性等。

## 参考文献

- [1] 姜永常. 论知识服务与信息服务. 情报学报, 2001(5):572-578.  
Jiang Y C. On knowledge service and information services. Journal of the China Society for Scientific and Technical Information, 2001(5): 572-578.
- [2] 陈碧叶. 从信息服务到知识服务. 图书馆杂志, 2004(1):20-22.  
Chen B Y. From Information Service to Knowledge Service. Library Journal, 2004(1):20-22.
- [3] 陈洪澜. 论知识分类的十大方式. 科学学研究, 2007(1):26-31.  
Chen H L. Ten methods of knowledge classification. Studies in Science of Science, 2007(1):26-31.
- [4] 马海云,薛翔. 面向知识服务的领域知识结构研究. 情报学报, 2022,41(1):73-82.  
Ma H Y, Xue X. Research on domain knowledge structure oriented to knowledge services. Journal of the China Society for Scientific and Technical Information, 2022,41(1):73-82.
- [5] 严志雁,吴辉,丁建,等. 基于 WebGIS 的数字农业园区信息系统研究. 江西农业学报, 2021,33(12):77-83.  
Yan Z Y, Wu H, Ding J, et al. Research on digital agricultural park information system based on WebGIS. Acta Agriculturae Jiangxi, 2021,33(12):77-83.
- [6] 苏新宁. 面向知识服务的领域知识组织纵论. 情报学报, 2022,41(9):889-899.

- Su X N. Knowledge service-oriented domain knowledge organization. *Journal of the China Society for Scientific and Technical Information*, 2022,41(9):889-899.
- [7] 马冠南,丁春雨,谷金英,等. 基于物联网的智慧农业移动平台设计与实现. *现代农业科技*, 2022(2):224-225+228.
- Ma G N, Ding C Y, Gu J Y, et al. Design and implementation of an IoT-based smart agriculture mobile platform. *XIANDAI NONGYE KEJI*, 2022(2):224-225+228.
- [8] 刘焯宸,李华昱. 领域知识图谱研究综述. *计算机系统应用*, 2020,29(6):1-12.
- Liu Y C, Li H Y. Survey on domain knowledge graph research. *Computer Systems & Applications*, 2020,29(6):1-12.
- [9] 孙雨生,倪雨婷. 国内基于本体的知识服务研究进展:架构、技术与应用. *图书馆学研究*, 2019(22):2-8+101.
- Sun Y S, Ni Y T. Research development of knowledge service based on ontology in China: Architecture, technology and application. *Researches in Library Science*, 2019(22):2-8+101.
- [10] 许智双,张昆,范俊超,等. 基于本体的网络安全知识图谱构建方法. *信息网络安全*, 2025,25(3):451-466.
- Xu Z S, Zhang K, Fan J C, et al. Construction method of cybersecurity knowledge graph based on ontology. *Netinfo Security*, 2025,25(3):451-466.
- [11] 刘杨,王金云,方亮. 基于本体的葡萄酒酿造知识图谱构建与应用研究. *中外葡萄与葡萄酒*, 2025,(2):100-108.
- Liu Y, Wang J Y, Fang L. Study on the construction and application of knowledge graph of wine brewing based on ontology. *Sino-Overseas Grapevine & Wine*, 2025,(2):100-108.
- [12] 陈莉,吴洁,马静,等. 基于本体的领域知识搜索研究. *计算机工程*, 2008,34(24):49-51.
- Chen L, Wu J, Ma J, et al. Research of domain knowledge search based on ontology. *Computer Engineering*, 2008,34(24):49-51.
- [13] 陈明,朱钰樟,席晓桃. 基于知识图谱的花卉病虫害知识管理方法. *农业机械学报*, 2023,54(3):291-300.
- Chen M, Zhu Y Z, Xi X T. Knowledge management method of flower diseases and pests based on knowledge graph. *Transactions of the Chinese Society of Agricultural Machinery*, 2023,54(3):291-300.
- [14] 穆维松,刘天琪,苗子激,等. 知识图谱技术及其在农业领域应用研究进展. *农业工程学报*, 2023,39(16):1-12.
- Mu W S, Liu T Q, Miao Z W, et al. Research progress on knowledge graph technology and its application in agriculture. *Transactions of the Chinese Society of Agricultural Engineering*, 2023,39(16):1-12.
- [15] Dreyer S, Werth O, Olivotti D, et al. Knowledge management systems for smart services: A synthesis of design principles. *e-Service Journal*, 2021, 13(2):27-67. DOI:10.2979/eservicej.13.2.02.
- [16] 谢能付,王文生. 农业知识本体构建方法. *农业网络信息*, 2007(8):12-16.
- Xie N F, Wang W S. Ontology and acquiring of agriculture knowledge. *Agriculture Network Information*, 2007(8):12-16.
- [17] Bonephace E M, Gui X, Makawia P J. Assessment of awareness of farmers on agricultural information services in Mara Region, Tanzania. *Library Philosophy & Practice*, 2022:1-23.
- [18] Zhao R X, Yang C X, Zheng J H, et al. Agricultural intelligent knowledge service: overview and future perspectives. *Smart Agriculture*, 2022, 4(4):105-125.
- [19] 熊大红. 基于本体的农业知识服务机制研究[D]. 湖南农业大学, 2008.
- Xiong D H. Research on the Mechanism for the Agricultural Knowledge Service Based on Ontology[D]. Hunan Agricultural University, 2008.
- [20] 刘琼,李宝敏. 一种果品领域本体库的构建方法. *计算机技术与发展*, 2009,19(1):197-199+203.
- Liu Q, Li B M. A Construction method of ontology library in fruit areas. *Computer Technology and Development*, 2009,19(1): 197-199+203.
- [21] 刘超. 农业领域(茶学)本体云建模方法研究[D]. 安徽农业大学, 2020.
- Liu C. Research on the Cloud Modeling Method of Agricultural Ontology--Taking Tea Knowledge as an Example[D]. Anhui Agricultural University, 2020.
- [22] 朱越. 基于小麦病虫害知识图谱的推荐技术研究[D]. 河南农业大学, 2022.
- Zhu Y. Studies of Recommendation Technology based on Diseases and Pests Knowledge Graph of Wheat[D]. Henan Agricultural University, 2022.
- [23] 沈金梦. 基于知识图谱的水稻病虫害诊断方法研究[D]. 吉林农业大学, 2021.
- Shen J M. Research on the Diagnosis Method of Rice Pests and Diseases Based on Knowledge Graph[D]. Jilin Agricultural University, 2021.
- [24] 赵娉花,米春桥. 柑桔病虫害自适应本体构建方法. *吉首大学学报(自然科学版)*, 2018,39(5):86-89.
- Zhao C H, Mi C Q. Adaptive ontology construction method for citrus pest management. *Journal of Jishou University(Natural Sciences Edition)*, 2018,39(5):86-89.

- [25] 陈泉林,贾璐. 面向战略运筹分析的事件本体及数据集构建方法. 系统仿真学报, 2025,37(4): 943-952.
- Chen Q L, Jia J, An event ontology and dataset construction method for strategic operations. Journal of System Simulation, 2025,37(4): 943-952.
- [26] 董磊,吴福居,史健勇,等. 基于大语言模型的施工安全多模态知识图谱的构建与应用. 计算机工程与应用, 2025,61(9):325-333.
- Dong L, Wu F J, Shi J Y, et al. Construction and application of multimodal knowledge graph in construction safety field based on large language model. Computer Engineering and Applications, 2025,61(9):325-333.
- [27] 朱晓俊,黄慧欣,汪鹏,等. 基于大模型的机器人研制场景知识图谱构建方法. 控制工程, 2025,32(8):1405-1416.
- Zhu X J, Huang H X, Wang P, et al. A method for constructing knowledge graph of robot development scenarios based on large language model. Control Engineering, 2025,32(8):1405-1416.
- [28] 冯在文,许多,田芳,等. 设施农业智能决策大模型关键技术与发展方向. 农业工程学报, 2025,41(13):50-61.
- Feng Z W, Xu D, Tian F, et al. Key technologies and development trends of intelligent decision-making large models for facility agriculture. Transactions of the Chinese Society of Agricultural Engineering, 2025,41(13): 50-61.
- [29] 曹喆,张泽. 从内容生产到知识服务:大模型如何扩展教育出版边界. 出版广角, 2025(5):18-26.
- Cao J, Zhang Z. From content production to knowledge services: How large models expand the boundaries of educational publishing. View on Publishing, 2025(5):18-26.
- [30] 刘洋,李斌,董云成,等. 新疆红枣种植与收获现状调研. 新疆农机化, 2021(3):32-34+44.
- Liu Y, Li B, Dong Y C, et al. Research on current planting and harvest situation of red jujube in Southern Xinjiang. Xinjiang Agricultural Mechanization, 2021(3):32-34+44.
- [31] 谢学军,金东艳,戴俊生,等. 红枣产业发展情况调研报告. 中国农村科技, 2021(10):54-57.
- Xie X J, Jin D Y, Dai J S, et al. Research report on the development of the jujube industry. China Rural Science & Technology, 2021(10): 54-57.
- [32] 包艳丽,程红梅,张利召. 新疆红枣产业发展研究. 农村经济与科技, 2022,33(1):145-146+152.
- Bao Y L, Cheng H M, Zhang L Z. Research on the development of Xinjiang's jujube industry. Rural Economy and Science, 2022, 33(1):145-146+152.
- [33] 王建文. 知识图谱算法之于阅读与出版智能服务探析. 出版广角, 2022(20):88-92.
- Wang J W. Analysis of knowledge graph algorithms in intelligent reading and publishing services. View on Publishing, 2022(20):88-92.
- [34] 杜雪盈,刘名威,沈立炜,等. 面向链接预测的知识图谱表示学习方法综述. 软件学报, 2024,35(1):87-117.
- Du X Y, Liu M W, Shen L W, et al. Survey on representation learning methods of knowledge graph for link prediction. Journal of Software, 2024,35(1):87-117.
- [35] Schreiber G, Akkermans H, Anjewierden A, et al. Knowledge engineering and management. Artificial Intelligence in Medicine, 1999,24(1):97-102. DOI:10.1016/S0933-3657(01) 00090-2.
- [36] 郑军红,何利力,叶修梓. 基于本体的夹具设计知识表达. 计算机研究与发展, 2010,47(7):1276-1285.
- Zheng J H, He L Y, Ye X Z. Ontology-based knowledge representation for computer-aided fixture design. Journal of Computer Research and Development, 2010,47(7):1276-1285.
- [37] 刘萍,石湘,邢玉艳. 卓越学者隐性知识体系建模与表示方法. 情报学报, 2023,42(8):915-925.
- Liu P, Shi X, Xing Y Y. Modeling and representation of tacit knowledge infrastructure of outstanding scholars. Journal of the China Society for Scientific and Technical Information, 2023,42(8):915-925.
- [38] 黄毅,张雪英,盛业华,等. 大语言模型驱动的台风灾害知识服务:关键技术与应用. 地球信息科学学报, 2025,27(6):1249-1262.
- Huang Y, Zhang X Y, Sheng Y H, et al. Typhoon disaster knowledge service driven by large language models: Key technologies and applications. Journal of Geo-Information Science, 2025,27(6): 1249-1262.
- [39] 周新,陈旸. 面向知识服务的红色文化资源知识图谱构建与应用研究. 档案与建设, 2025(5):72-83.
- Zhou X, Chen Y. Knowledge mapping of red cultural resources for knowledge service and application research. Archives & Construction, 2025(5):72-83.
- [40] 张云中,郭冬,王亚鸽,等. 基于知识图谱的红色历史人物知识问答服务框架研究. 图书情报工作, 2021,65(16):108-117.
- Zhang Y Z, Guo D, Wang Y G, et al. Framework of knowledge Q & A service for red historical figures based on knowledge graph. Library And Information Service, 2021,65(16):108-117.
- [41] 王元胜,吴华瑞,赵春江. 农业知识驱动服务技术革新综述与前沿. 农业工程学报, 2024,40(7):1-16.
- Wang Y S, Wu H R, Zhao C J. Agricultural knowledge driven service technology innovation: Overview and frontiers. Transactions of the

- Chinese Society of Agricultural Engineering, 2024,40(7):1-16.
- [42] 俞德凤. 美国康奈尔大学图书馆隐私保护服务及启示. 农业图书情报学报, 2021,33(11):28-37.
- Yu D F. The practice and enlightenment on privacy protection service of cornell university library. Journal of Library and Information Science in Agriculture, 2021,33(11):28-37.
- [43] 李鸿亮,刘禹良,廖文辉,等. 大模型时代的光学文字识别:现状及展望. 中国图象图形学报, 2025,30(6):2023-2050.
- Li H L, Liu Y L, Liao W H, et al. OCR in the era of large models: Current status and prospects. Journal of Image and Graphics, 2025,30(6):2023-2050.
- [44] 刘峤,李杨,段宏,等. 知识图谱构建技术综述. 计算机研究与发展, 2016,53(3):582-600.
- Liu Q, Li Y, Duan H, et al. Knowledge graph construction techniques. Journal of Computer Research and Development, 2016,53(3):582-600.
- [45] 徐勇,甘国辉,牛方曲. 农业信息协同服务总体架构解析. 农业网络信息, 2009(9):10-12.
- Xu Y, Gan G H, Niu F Q. Analysis of cooperative service framework of agricultural information. Agriculture Network Information, 2009(9): 10-12.
- [46] 徐勇,高雅,刘艳华. 农业过程本体及其构建方法:以玉米为例. 农业网络信息, 2009(11):8-11.
- Xu Y, Gao Y, Liu Y H. Building method of agriculture process ontology: A case of corn. Agriculture Network Information, 2009(11): 8-11.
- [47] 王毅,刘云生,钟昊. XSD 优化存储与解析. 计算机工程, 2005(19): 46-48.
- Wang Y, Liu Y S, Zhong H. Optimized storage & parser for XSD. Computer Engineering, 2005(19):46-48.
- [48] 曹思源,马海云. 领域知识组织理论基础及方法分类简述. 情报资料工作, 2021,42(5):14-22.
- Cao S Y, Ma H Y. A brief introduction to the theoretical basis and method classification of domain knowledge organization. Information and Documentation Services, 2021,42(5):14-22.
- [49] 向安玲,高爽,彭影彤,等. 知识重组与场景再构:面向数字资源管理的元宇宙. 图书情报知识, 2022, 39(1):30-38.
- Xiang A L, Gao S, Peng Y T, et al. Knowledge reorganization and scene reconstruction: A metaverse for digital resources management. Document, Information & Knowledge, 2022, 39(1):30-38.
- [50] 岳静,张自力. 本体表示语言研究综述. 计算机科学, 2006(2): 158-162.
- Yue J, Zhang Z L. A survey on ontology specification languages. Computer Science, 2006(2):158-162.
- [51] 甘健侯,夏幼明,徐天任,等. 本体描述语言 OWL 知识表示的扩展. 云南师范大学学报(自然科学版), 2005(4):9-14.
- Gan J H, Xia Y M, Xu T R, et al. Extension of web ontology language OWL in knowledge representation. Journal of Yunnan Normal University(Natural Sciences Edition), 2005(4):9-14.
- [52] 王克龙,王玲,王平立,等. 数据仓库中 ETL 技术的探讨与实践. 计算机应用与软件, 2005 (11):32-33+80.
- Wang K L, Wang L, Wang P L, et al. Research on technology of ETL in data warehouse and its practice. Computer Applications and Software, 2005 (11):32-33+80.
- [53] Liang K K. Implementation of ETL scheme based on storm platform. Computer Science, 2019,46(S2): 208-211+240.
- [54] Wisnesky R, Filonik D. Relational to RDF data migration by query Co-evaluation[J/OL]. [2025-05-05]. <https://doi.org/10.48550/arXiv.2403.01630>.
- [55] Devlin J, Chang M W, Lee K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding[J/OL]. [2025-05-05]. <https://doi.org/10.48550/arXiv.1810.04805>.
- [56] Huang Z H, Xu W, Yu K. Bidirectional LSTM-CRF models for sequence tagging. Computer Science[J/OL]. [2025-05-05]. <https://doi.org/10.48550/arXiv.1508.01991>.
- [57] 陈芸,邓晖,吴道婷. 基于自然语言处理的教学设计学科知识图谱自动构建研究. 中国教育信息化, 2020(7): 15-19.
- Chen H, Deng H, Wu D T. Research on automatic construction of knowledge graphs for instructional design discipline based on natural language processing. China Education Info, 2020(7): 15-19.
- [58] 李阳,杜睿山,张豪鹏. 面向医药信息知识图谱构建. 计算机技术与发展, 2022,32(10):189-193.
- Li Y, Du R S, Zhang H P. Construction of knowledge graph for medical data. Computer Technology and Development, 2022,32(10): 189-193.
- [59] 李敏瑜,李永卉,卢章平. 古方志历史名人资源知识图谱构建与知识发现研究. 图书馆杂志, 2026,45(1):61-73.
- LI M Y, LI Y H, LU Z P. Research on knowledge graph construction and knowledge discovery of ancient local chronicles and historical celebrities resources. Library Journal, , 2026,45(1):61-73.
- [60] 高劲松,方晓印,刘思洋,等. 基于关联数据的馆藏文物资源知识关联与智能问答研究. 情报科学, 2021,39(5):12-20.
- Gao J S, Fang X Y, Liu S Y, et al. Research on knowledge association and intelligent question-and-answer of cultural relic resources based on linked-data. Information Science, 2021,39(5): 12-20.
- [61] 黄永文,孙坦,赵瑞雪,等. 大数据与人工智能背景下新型知识服务

- 研究与实践. 图书情报工作, 2022,66(19):36-46.
- Huang Y W, Sun T, Zhao R X, et al. Research and practice of new knowledge service in the context of big data and artificial intelligence. Library And Information Service, 2022,66(19):36-46.
- [62] 李博文,王永志,丁正江,等. 基于大模型与 GraphRAG 的胶东金矿智能搜索技术. 地学前缘, 2025, 32 (4):155-164.
- Li B W, Wang Y Z, Ding Z J, et al. Intelligent search technology for Jiaodong gold deposits based on large models and GraphRAG. Earth Science Frontiers, 2025, 32 (4):155-164.
- [63] 孙雨生,曾俊皓,陈思好,等. 知识图谱增强的政策大模型知识问答系统构建研究. 图书馆学研究, 2025(5):66-75.
- Sun Y S, Zeng J H, Chen S Y, et al. Construction of the large language model-based policy knowledge question and answering system enhanced by knowledge graph. Researches In Library Science, 2025(5):66-75.
- [64] 李卓伦,梁少军,摆卫兵,等. 基于大模型的武器装备故障知识图谱构建研究. 计算机测量与控制, 2025,33(9):176-181+190.
- Li Z L, Liang S J, Bai W B, et al. Large model-based fault knowledge graph construction method for weapon equipment. Computer Measurement & Control, 2025,33(9):176-181+190.
- [65] 范俊杰,马海群,刘兴丽. 数智时代下开源情报的军事知识图谱问答智能服务研究[J/OL]. 数据分析与知识发现. [2023-10-26]. <http://kns.cnki.net/kcms/detail/10.1478.G2.20231026.1305.002.html>.
- FAN J J, MA H Q, LIU X L. Research on Intelligent Question-Answering Services for Military Knowledge Graphs Based on Open Source Intelligence in the Era of Digital Wisdom[J/OL]. Data Analysis and Knowledge Discovery. [2023-10-26]. <http://kns.cnki.net/kcms/detail/10.1478.G2.20231026.1305.002.html>.
- [66] 李新. 基于本体的制造资源建模与智能搜索研究开发[D]. 西安建筑科技大学, 2016.
- Li X. Research and Development on Modeling and Intelligent Searching of Manufacturing Resources Based on Ontology[D]. Xi'an University of Architecture and Technology, 2016.
- [67] 陈德彦,赵宏,张霞. 专家视图与本体视图的语义映射方法. 软件学报, 2020,31(9):2855-2882.
- Chen D Y, Zhao H, Zhang X. Semantic mapping methods between expert view and ontology view. Journal of Software, 2020,31(9):2855-2882.
- [68] 熊鹏飞. 基于 DQV 的知识图谱数据质量完整性评估模型研究[D]. 东北石油大学, 2024.
- Xiong P F. Research on Date Quality Completeness Assessment Model of KnowledgeGraph based on DOV[D]. Northeast Petroleum University, 2024.
- [69] 袁满,刘国娇,袁靖舒,等. 标准驱动的本体质量评估模型研究. 吉林大学学报(信息科学版), 2025,43(3): 605-614.
- Yuan M, Liu G J, Yuan J S, et al. Research on assessment model of ontology quality based on standard-driven approaches. Journal of Jilin University (Science Edition), 2025,43(3): 605-614.
- [70] Gómez-Pérez A. Evaluation of ontologies. International Journal of Intelligent Systems, 2001,16(3):391-419.
- [71] 徐雷. 本体评估研究进展. 情报学报, 2016,35(7): 772-784.
- Xu L. Research advances in ontology evaluation. Journal of the China Society for Scientific and Technical Information, 2016, 35(7): 772-784.
- [72] 宋玲玲. 基于知识图谱和 LibQUAL 模型的高校图书馆学科服务质量评估指标监测机制研究. 图书馆学研究, 2019(10):75-83.
- Song L L. Research on the monitoring mechanism of subject service quality evaluation index of university library based on knowledge graph and LibQUAL. Researches in Library Science, 2019(10):75-83.
- [73] 许周. 基于叙词表的本体知识可信性评估工具的设计与实现[D]. 东南大学, 2021.
- Xu Z. Design and Implementation of Ontology Credibility Evaluation tool Based on Thesaurus[D]. Southeast University, 2021.
- [74] 张子辰,岳昆,祁志卫,等. 时序知识图谱的增量构建. 计算机科学与探索, 2022,16(3):598-607.
- Zhang Z C, Yue K, Qi Z W, et al. Incremental construction of time-series knowledge graph. Journal of Frontiers of Computer Science & Technology, 2022,16(3):598-607.
- [75] 王楚童,李明达,孙孟轩,等. 融合大规模医学事实的跨语言双层知识图谱. 软件学报, 2025,36(3):1240-1253.
- Wang C T, Li M D, Sun M X, et al. Cross-language bilayer knowledge graph with large-scale medical facts. Journal of Software, 2025,36(3): 1240-1253.
- 引用本文: 薛原,冯建中,杨浩,王剑,高祥伟,谢能付,曹丹,白林燕,陈彩铭. 基于本体的红枣领域知识智能服务研究[J]. 农业大数据学报, 2026,8(3): 321-337. DOI: 10.19788/j.issn.2096-6369.000133.
- CITE: XUE Yuan, FENG JianZhong, YANG Hao, WANG Jian, GAO XiangWei, Xie NengFu, CAO Dan, BAI LinYan, CHEN CaiMing. Research on Intelligent Services for Jujube Domain Knowledge, Based on Ontology[J]. Journal of Agricultural Big Data, 2026, 8(3): 321-337. DOI:10.19788/j.issn.2096-6369.000133.

## Research on Intelligent Services for Jujube Domain Knowledge, Based on Ontology

XUE Yuan<sup>1</sup>, FENG JianZhong<sup>1\*</sup>, YANG Hao<sup>1,2</sup>, WANG Jian<sup>1</sup>, GAO XiangWei<sup>2</sup>, XIE NengFu<sup>1</sup>, CAO Dan<sup>3\*</sup>, BAI LinYan<sup>4</sup>, CHEN CaiMing<sup>1</sup>

1. Agricultural Information Institute of Chinese Academy of Agricultural Sciences, Beijing 100081, China; 2. School of Marine Technology and Geomatics, Jiangsu Ocean University, Lianyungang, Jiangsu 222005, China; 3. China Fire and Rescue Institute, Beijing 102201, China; 4. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China

**Abstract:** To address the current issues in the red jujube domain, including fragmented knowledge systems, delayed updates, insufficient personalization of knowledge services, and low service quality and efficiency, and to meet the practical needs for efficient organization and intelligent services of red jujube domain knowledge, in this paper, we carried out a study about ontology-based intelligent knowledge service, which was associated with relevant aspects, for the red jujube industry. Firstly, a fundamental framework for jujube domain knowledge services was established through systematic analysis of the knowledge service processes in the jujube industry, and based on domain ontology and knowledge graphs, the intelligent service mechanism and patterns for the jujube domain were studied by exploring knowledge organization approaches about the jujube industry. Secondly, the structured, semi-structured and unstructured multi-source heterogeneous corpus information in terms of the cultivation and management of the jujube industry and its related domains was systematically integrated and processed (e.g., semantic preprocessing), and then, the multi-dimensional, standardized domain knowledge and ontology bases for the jujube industry were constructed, and further, the domain knowledge graph for the jujube industry, which showed good merits such as favorable completeness, well fault-tolerant and disambiguation capabilities, and strong relational reasoning competence, was built by using a BERT-BiLSTM-CRF hybrid model. Hence, to satisfy the requirements of intelligent question-and-answering (Q&A) services for the jujube domain, an intelligent Q&A system, based on the domain ontology and knowledge graph and keeping to the collaborative optimization principle of domain knowledge services, was designed while employing a hybrid processing framework with domain-specific models and (locally deployable) large language models (LLMs), and moreover, the improvement approaches for evaluation of knowledge intelligence services were discussed to ensure the provision of high-quality, reliable, and satisfactory services to users and/or systems. As for all the above, the paper would provide significant practical references for enhancing the practicality and applicability of intelligent services in the jujube domain, promoting the digital and intelligent transformation and upgrading of the jujube industry, and thus further exploring their feasible technical pathways.

**Keywords:** domain ontology; red jujube; knowledge graph; intelligent service; knowledge Q&A