

面向语义Web的旅游领域本体构建方法研究

古圣钰^{1,2}, 章牧²

¹惠州学院地理与旅游学院, 广东省 惠州市

²暨南大学深圳旅游学院, 广东省 深圳市

发布日期: 2025年12月1日

摘要

随着旅游行业智能化转型的不断推进, 如何实现多源旅游数据的语义集成与高效组织, 成为提升旅游信息服务质量的关键路径。语义Web技术为旅游本体构建提供了有力支撑, 其核心在于构建结构化、可推理的领域知识体系。本文以旅游领域本体为研究对象, 系统梳理了其定义、类型与构建流程, 深入分析了人工主导与自动化方法的技术路径与应用差异, 并对主流建模工具及典型案例进行了结构性比较。在此基础上, 文章归纳了当前旅游本体构建中面临的数据异构、本体一致性、跨文化适配等主要挑战, 并提出了包括本体学习、图神经网络、大语言模型 (LLM) 等在内的未来研究方向。研究结果可为相关科研人员与旅游智能服务系统开发者提供理论借鉴与实践参考。

关键词

语义Web; 旅游本体; 本体构建方法; 知识图谱; 本体学习; 大语言模型

Ontology Construction Methods for the Tourism Domain Oriented to the Semantic Web

Shengyu Gu^{1,2}, Mu Zhang²

¹School of Geography and Tourism, Huizhou University, Huizhou 516007, Guangdong, China

²Shenzhen Tourism College, Jinan University, Shenzhen 518053, Guangdong, China

Abstract

With the accelerating transformation of the tourism industry toward intelligent services, achieving semantic integration and efficient organization of heterogeneous tourism data has become a critical issue. Semantic Web technologies provide a solid foundation for ontology construction in the tourism domain, enabling the development of structured and machine-interpretable knowledge systems. This paper presents a comprehensive review of tourism ontology construc-

文章引用: 古圣钰, & 章牧. (2025). 面向语义 Web 的旅游领域本体构建方法研究. 计算机科学辑要, 1(2), 41-50.

ISSN Print: 3079-5362; ISSN Online: 3079-5370.

DOI: 10.63313/CS.8021

tion methods, covering key concepts, ontology types, and construction workflows. It further analyzes both manual and (semi-)automated approaches, compares mainstream ontology engineering tools, and summarizes representative case studies. The study highlights major challenges such as data heterogeneity, ontology consistency, and cross-cultural adaptability. Finally, the paper explores future directions including ontology learning, integration with knowledge graphs, large language model-assisted construction, and the application of graph neural networks. The findings aim to support researchers and practitioners in designing more adaptive, scalable, and intelligent semantic frameworks for tourism information systems.

Keywords

Semantic Web; Tourism Ontology; Ontology Construction; Knowledge Graph; Ontology Learning; Large Language Models

Copyright © 2025 by author(s) and Erytis Publishing Limited.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



1. 引言

1.1. 研究背景与动因

随着数字经济的蓬勃发展, 旅游产业逐步迈入信息化、智能化、平台化的新阶段。伴随在线旅游预订、用户点评系统、社交媒体打卡、移动导航服务等技术广泛应用, 旅游相关数据呈现出来源多样、结构复杂、语义异构、更新频繁等显著特点, 传统的人工处理和关键词检索模式难以满足对海量旅游信息的高效组织、精准理解与智能推荐的实际需求。在此背景下, 语义 Web 技术作为新一代知识表达与共享手段, 为旅游领域的数据整合与智能服务提供了有力支撑。语义 Web 强调通过统一的语义框架对异构数据进行语义标注、语义建模与本体推理, 使计算机能够“理解”信息的语义内涵, 从而实现基于本体的自动化推理与知识发现[1]。通过构建科学系统的旅游领域本体, 不仅可实现对景点、线路、活动、用户偏好等关键概念的形式化定义, 还可实现数据之间的语义关联与可计算性推理, 提升旅游信息服务系统的智能化水平与用户满意度。目前, 部分研究已在旅游本体构建方面开展了探索性工作, 如基于专家知识的手工本体建模、基于文本挖掘的半自动本体生成, 以及基于机器学习的知识抽取方法等。然而, 从总体来看, 现有研究仍存在理论建模不统一、方法体系不完备、工具适配性差、语义推理效率低、缺乏动态演化机制等问题, 亟须在更高层面系统梳理研究现状、归纳主流方法、剖析技术瓶颈, 并为后续研究指明方向。因此, 深入研究面向语义 Web 的旅游本体构建方法, 对于推动旅游大数据的语义集成、支撑智能推荐系统建设、构建旅游知识图谱以及提升旅游服务智能水平具有重要的理论价值与现实意义。

1.2. 本文综述的目的与价值

本研究旨在围绕“面向语义 Web 的旅游领域本体构建方法”主题, 系统回顾近年来相关研究成果, 梳理旅游本体建模所采用的典型方法、常用工具与平台, 分析不同方法的优劣及适用场景, 总结当前存在的挑战, 并对未来研究方向进行预测与展望。与已有综述文献相比, 本文强调旅游领域语义建模在多源异构数据环境下的适配能力, 关注语义 Web 技术与人工智能融合背景下本体构建方法的演进趋势,

力求为该领域的后续研究提供结构清晰、内容全面、问题导向明确的知识框架与参考资料。

1.3. 结构安排说明

全文共分为六个部分。第一部分为引言, 阐明研究背景、动因与价值; 第二部分介绍本体及语义 Web 的相关理论基础, 为理解构建方法提供理论支撑; 第三部分系统综述旅游领域本体构建的典型方法、建模工具及代表性案例; 第四部分探讨当前本体构建实践中存在的核心挑战; 第五部分提出未来研究的若干发展方向与关键议题; 第六部分为全文总结与归纳, 明确本研究的总体贡献。

2. 理论基础与关键技术

2.1. 旅游本体的定义与类型

本体 (Ontology) 是知识表示领域中的核心概念, 通常被定义为对现实世界某一领域内概念、实体及其相互关系的形式化、结构化和可计算的表达模型。在语义 Web 环境下, 本体不仅作为概念语义建模的工具, 更是支撑语义标注、语义推理与知识重用的基础框架[2]。根据抽象层次与应用范围的不同, 本体通常分为上位本体 (Upper Ontology)、领域本体 (Domain Ontology) 与应用本体 (Application Ontology) 三类, 三者概念粒度、适用范围和通用性方面具有显著差异, 具体如下: (1) 上位本体一般用于建模具有高度抽象性的通用概念, 如“实体”“事件”“时间”“空间”等, 是跨领域语义互通的桥梁。(2) 领域本体则面向特定行业或知识领域, 描述该领域内的基本概念及其关系, 如旅游中的“景点”“路线”“节庆”“服务”等。(3) 应用本体通常依附于具体的业务场景或系统平台, 在继承领域本体结构的基础上, 进一步定义特定系统中的功能需求与语义关系, 如“用户偏好”“预订流程”“评论标签”等。在旅游信息系统中, 上位本体提供通用语义支撑, 领域本体构建旅游知识结构主干, 而应用本体则与实际业务逻辑紧密结合, 三者共同构成多层次语义建模体系。表 1 对这三类本体在旅游语境中的应用功能进行了比较分析:

表 1. 旅游本体类型与功能比较

本体类型	定义特征	表示内容	在旅游领域中的作用
上位本体	抽象层次高, 跨领域通用	实体、事件、空间、时间等	提供通用概念体系, 支撑跨系统互联
领域本体	领域专属, 结构系统	景点、行程、交通、节庆等	构建旅游语义结构主干, 统一数据语义
应用本体	任务导向, 面向具体系统需求		

值得注意的是, 旅游本体的类型划分虽具有一定层级性, 但在实际系统中往往交叉使用, 需在建模阶段明确边界与耦合方式, 确保语义一致性与模型扩展性。针对旅游领域的语义复杂性, 构建结构合理、逻辑清晰的多层级本体体系已成为研究与工程实践的核心环节之一。

2.2. 语义 Web 支撑技术概览

语义 Web 作为万维网的演进形态, 旨在通过结构化语义增强, 使 Web 信息具备机器可读性与可推理性, 进而实现信息的自动处理与智能服务[1]。在旅游信息系统中, 语义 Web 技术提供了一套完整的知识建模与语义表达机制, 主要依托以下三项核心技术实现: 一是资源描述框架 (RDF)。RDF (Resource Description Framework) 是语义 Web 的基础数据模型, 其通过“主语—谓语—宾语”的三元组结构来描述资源间的语义关系。例如, “张三 喜欢 桂林”可建模为一条 RDF 语句, 明确指向个体与行为之间的逻辑连接。RDF 模型支持开放世界假设, 具备良好的可扩展性与互操作性, 适用于描述旅游领域中多元数据之间的关联结构。二是 Web 本体语言 (OWL)。OWL (Web Ontology Language) 是在 RDF 和 RDFS 基础上扩展的本体描述语言, 支持更复杂的逻辑约束和语义推理规则。通过 OWL 可精确表达旅游本体

中的类层级、属性约束、个体实例及其之间的推理关系，满足本体一致性校验、子类继承、类间推导等语义需求，是构建可推理旅游知识模型的重要支撑工具[3]。三是 SPARQL 查询语言。SPARQL 是用于查询 RDF 图数据的标准语言，支持复杂的图模式匹配、过滤和聚合操作。在旅游应用中，SPARQL 可用于查询“用户感兴趣的自然景点”“距离机场 3 公里以内的酒店”等复杂语义条件，极大提升信息检索的精确度与响应效率。上述三项技术共同构成了语义 Web 支撑架构，在旅游系统中分别承担语义建模、知识表达与语义查询推理等关键职能。表 2 展示了它们在旅游信息系统中的功能定位与相互作用机制：

表 2. 语义 Web 技术在旅游系统中的功能定位图

序号	功能模块	技术/标准	主要作用与功能描述
1	旅游原始数据采集	数据源（文本、API、数据库、图像）	提供基础数据来源，包括景点信息、用户行为、评论文本、地图数据等
2	语义建模	RDF（资源描述框架）	使用三元组方式对旅游实体、属性及其关系进行结构化建模
3	本体结构与规则建模	OWL（Web 本体语言）	定义旅游概念层级结构、类之间的继承关系、属性约束与语义逻辑规则
4	语义推理引擎	逻辑推理（基于 OWL）	基于定义的本体进行一致性检查、子类推导、实例归类等推理操作
5	语义查询接口	SPARQL	查询本体与 RDF 数据图，实现复杂语义检索（如用户兴趣、时空偏好匹配）
6	推荐系统与服务输出	Web 服务/前端组件	向用户展示个性化推荐结果，实现旅游产品展示、行程规划与交互服务

通过上述架构，可实现旅游数据的结构化处理、语义增强与智能推荐的无缝衔接，为本体构建与语义集成提供了坚实的技术基础。随着旅游大数据与人工智能的深度融合，语义 Web 将在旅游领域智能化发展过程中发挥更加核心的作用。

2.3. 本体构建的基本流程

在语义 Web 框架下，旅游领域本体的构建需遵循系统化、规范化和可复用的技术流程。一般而言，本体构建包括概念提取、模型设计、语义编码、一致性验证及动态演化等关键阶段，各阶段任务虽相对独立，但整体逻辑应形成闭环，确保知识模型的完整性、逻辑性与应用适配性[4]。第一，概念提取阶段，主要面向领域知识的语义识别与分类整理。该阶段需通过文献梳理、专家访谈、词频分析、语义抽取等方式，识别出旅游领域中的核心概念与基本术语，如“景区”“路线”“交通”“节庆活动”等。第二，模型设计阶段，即构建本体结构框架。此阶段通常以图形化建模工具为支撑（如 Protégé），在定义“类—子类”层级结构的基础上，进一步明确概念之间的属性关系、从属关系与限制条件。第三，语义编码与一致性验证阶段，需使用 RDF/OWL 等语义语言对建模结果进行形式化描述，并借助语义推理引擎（如 Pellet、Hermit）对本体的逻辑一致性进行系统校验，防止语义冲突、类层次矛盾或属性重复等问题。第四，动态演化阶段，面向本体随数据变化或业务需求扩展而进行的更新与优化。此阶段需引入版本控制机制，支持语义增量建模、结构调整与实例更新，并确保与原有模型的兼容性与稳定性。上述流程不仅适用于面向旅游场景的通用本体建构，同样也可为后续开展半自动本体学习、旅游知识图谱生成与语义推荐系统提供方法论基础。表 3 对本体构建流程中的关键阶段与常用技术方法进行了归纳总结。

表 3. 本体构建阶段与对应技术方法

构建阶段	核心任务	常用方法与工具	说明
概念提取	识别领域术语与基本概念	关键词抽取、实体识别、专家知识	可结合文本挖掘与 NLP 技术提取候选概念
模型设计	建立类层级结构与属性关系	Protégé、OntoEdit	支持图形化建模、层次结构可视化
语义编码与验证	使用形式语言编码并检验一致性	OWL/RDF、Pellet、Hermit	编码过程需符合 OWL DL 规范并通过推理校验
动态演化	模型更新、版本控制与兼容处理	Ontology Versioning 机制	支持模型可持续演化及与数据源同步更新

3. 旅游本体构建方法综述

3.1. 人工主导的本体构建方法

人工主导的本体构建方法，是在早期语义建模研究中广泛应用的一种传统策略，主要依赖领域专家经验、规范文档与人工知识抽取等手段，逐步完成概念体系的识别、分类与建模任务。该方法强调建模过程的严谨性、准确性与语义一致性，适用于知识稳定、领域明确、控制要求高的应用场景，曾广泛用于政府标准制定、行业规范发布与本体工程原型建设等方面。具体来看，该方法通常包括以下步骤：一是由领域专家根据专业知识与业务流程，提炼出旅游活动中的关键实体与概念；二是组织概念之间的层次结构，形成“类—子类”继承体系；三是定义概念间的属性关系与约束规则；四是采用本体建模工具（如 Protégé）进行结构化建模，并在小范围内进行初步语义验证。该方法的优势在于语义建模精准，逻辑结构清晰，便于实现高质量的本体框架输出；但也存在明显局限：一是知识获取依赖人工，过程耗时且人力成本高；二是缺乏自动化支持，难以应对旅游数据动态性强、语义复杂、应用快速迭代的现实需求；三是在多语、多源背景下可扩展性弱，难以适配跨平台语义集成与系统互操作。目前，人工建模方式主要用于旅游知识体系顶层设计、本体模板构建与高可信场景的本体微调等方面，并作为自动化方法的语义输入源或训练基础。

3.2. 半自动与自动化方法

随着旅游数据规模迅速扩张及语义复杂性不断提升，单纯依赖人工手段构建旅游本体已难以满足动态化、大数据环境下的实际需求。近年来，半自动与自动化本体构建方法逐渐成为研究热点，依托自然语言处理（NLP）、信息抽取、机器学习等先进技术，从多源数据中自动识别语义结构、生成本体草案、辅助建模过程，显著提升建模效率与适配性[5]。半自动建模方法主要指在人工知识框架下引入技术辅助，对本体构建的某一或多个阶段进行流程优化。例如，基于文本语料库的词频分析可自动抽取领域术语；基于句法分析的模板匹配方法可识别基本概念间的关系；实体链接与词义消歧算法可用于确定术语边界与语义指向。该类方法在降低人工工作量的同时，仍保留领域专家的知识审核与决策，确保语义准确性。自动化建模方法则尝试实现本体构建的全过程自动执行，适用于结构化或半结构化数据环境，尤其在语义开放的知识图谱构建场景中应用广泛。该类方法主要包括以下几类技术路径：一是基于信息抽取的本体生成。通过规则驱动或统计学习方法，从旅游文本、网页、评论等非结构化数据中提取实体、属性与关系，再映射为 OWL 格式的本体结构。二是基于机器学习的本体学习。利用监督学习、无监督聚类或嵌入模型（如 Word2Vec、BERT），对大规模旅游数据中的概念进行聚合与归纳，形成本体概念网络。近年来，图神经网络（GNN）与预训练语言模型也开始被引入到本体结构学习中，进一步提升建模智能化水平[6]。三是利用已有知识库进行本体对齐与迁移。通过与已有旅游相关本体（如 Schema.org、DBpedia、YAGO 等）对齐，快速建立本体框架，并据此进行语义继承、扩展与本地适配。虽然自动化方法在效率与适应性方面表现优越，但当前仍面临一些核心挑战：一是语义噪声与冗余高，自动识别的概念往往缺乏层次性与一致性；二是语境依赖性强，旅游术语在不同文化、地域语境下含义可能存在差异；三是缺乏统一评估标准，自动构建的本体难以在实际系统中直接部署。

综合来看，半自动与自动方法应作为传统人工建模的有力补充，在复杂旅游语义环境下实现“人工+技术”的协同建模机制。表 4 对当前主流的自动化本体构建技术流程作简要归纳，供后续系统构建与模型优化参考。

表 4. 自动化本体构建技术流程简表

流程阶段	主要技术方法	数据来源类型	工具/框架示例	技术优势
------	--------	--------	---------	------

概念抽取	词频统计、命名实体识别 (NER)	原始文本、用户评论、新闻语料	OpenNLP、Stanford NLP	自动识别核心概念词及实体边界，提高语义解析基础质量
关系抽取	句法依存分析、关系模板匹配	描述性文本、旅游文档等非结构化语料	spaCy、Text2Onto	准确提取概念间的语义关系与属性，有助于结构建模
概念聚合	聚类分析、主题建模、词向量嵌入算法	搜索日志、评论日志、大规模语料	LDA、Word2Vec、BERT	聚合同义或近义概念，提升本体一致性与泛化能力
本体编码	语义映射、OWL/RDF 转化	结构化语义数据或知识片段	Protégé 插件、RDFLib	快速构建机器可读的本体结构，支持语义推理
本体对齐	本体匹配、知识迁移、对齐验证	外部本体库、开放知识图谱（如 DBpedia、Wikidata）	OntoAlign、LogMap、AML	加速本体构建过程，实现异源本体间的高效集成

3.3. 常用本体建模工具

本体建模工具作为语义 Web 体系中支撑知识建构、可视编辑与语义编码的核心软件平台，其性能与适配性直接影响本体开发效率与部署效果。在旅游领域，因本体内容涉及概念层次复杂、属性关系多样、实例演化频繁，需借助功能完备、扩展性强的建模工具完成语义建模、验证与发布等全流程任务。目前，国际上较为主流的本体建模工具包括 Protégé、OntoEdit 与 WebProtégé 等，具备良好的可视化建模界面、OWL 语义支持能力以及插件扩展机制。其中，Protégé 由斯坦福大学开发，是应用最广泛的开源本体编辑平台，支持 RDF/OWL 标准，具备类层次图管理、属性定义、推理引擎集成等功能，适用于初学者与专家用户。WebProtégé 则为其轻量级在线版本，适合分布式协同建模、跨团队在线编辑与权限管理场景，逐渐成为云端语义建模的主流选择。OntoEdit 则更偏重于企业级建模需求，内嵌推理与一致性校验引擎，适用于本体与应用系统深度集成的需求环境。旅游本体建模实践中，工具选择应结合任务复杂度、用户水平、部署环境等因素综合考虑，确保功能匹配、数据互通与长期可维护性。表 5 汇总了三种主流工具的核心功能与适用场景，为本体构建实践提供参考依据。

表 5. 主流本体建模工具功能与适用性对比

工具名称	支持标准	核心功能	扩展能力	使用场景特点
Protégé	OWL、RDF(S)	类层次编辑、属性定义、推理集成	插件丰富	桌面端建模，适合单体开发与教学演示
WebProtégé	OWL	在线建模、协同编辑、权限管理	云端部署	分布式团队协作、远程访问
OntoEdit	RDF、RDFS、OWL	推理支持、企业流程集成	商业插件支持	企业系统对接、本体业务集成

3.4. 国内外旅游本体代表案例

在全球范围内，旅游领域的本体构建实践已逐步从理论探索走向工程应用，涌现出一批具有代表性的本体模型。这些模型既服务于旅游信息的语义组织与集成，也支撑旅游推荐、问答系统、知识图谱等智能服务功能，体现出良好的领域适配性与系统通用性。国外研究中，ONTOUR 是较早出现的旅游本体之一，由德意志研究中心（DFKI）提出，聚焦于旅游活动的概念建模，覆盖旅行、住宿、交通、天气等基本服务要素，并与本地行程规划服务深度融合。该本体采用 OWL 语言编写，支持多语言表达和语义查询，可嵌入旅游问答与路线推荐系统[7]。Hontology 是西班牙瓦伦西亚理工大学开发的本体项目，采用模块化建模思想，围绕旅游住宿、目的地特征、交通方式、旅游活动等维度进行本体分层管理。该本体在旅游数据集成和语义互操作性方面表现优异，已被多个欧洲智能旅游平台采用，具有良好的可重用性与标准化水平[8]。在国内，部分高校与政府项目也开展了旅游本体建模实践。如“智慧旅游知识工程”课题组结合我国文化旅游实际构建了涵盖自然景观、人文遗产、非遗活动、节庆时间轴等语义模块的旅游本体，注重本地文化表达能力与政务数据整合能力，已在若干政务信息平台中试点应用。

综合分析上述案例，旅游本体的建模策略呈现出以下趋势特征：一是普遍采用模块化、分层式建模方式，便于系统维护与扩展；二是强调与知识图谱、语义服务平台的耦合能力；三是逐步融入用户偏好、

评价反馈等动态语义元素，提升本体的实用性与交互性。表 6 对部分具有代表性的旅游本体进行了结构性对比，便于横向分析与实践借鉴。

表 6. 典型旅游本体对比一览表

本体名称	建立机构	涵盖内容范围	技术特征	应用场景与系统集成
ONTOUR	德国 DFKI	住宿、交通、活动、天气	OWL 建模、多语言、多源融合	路线规划、旅游问答系统
Hontology	西班牙瓦伦西亚理工大学	住宿、目的地、活动、交通	模块化结构、SPARQL 支持	欧洲旅游数据平台集成
智慧旅游本体	国内高校与政务平台	自然景观、人文遗产、节庆	强调本地性、非遗建模、与政务数据对接	

4. 面临的挑战

尽管旅游本体在语义 Web 背景下已取得一定的研究进展与应用成果，但在大数据、高动态、多语言环境中，旅游本体构建工作仍面临诸多现实挑战。尤其是在数据异构性、本体一致性、语义多样性等方面，制约因素明显，亟须理论与技术手段协同攻克。结合当前典型研究案例与工程实践经验，本文归纳出以下三项关键问题：

4.1. 数据异构性与建模复杂度

旅游领域的数据具有显著的多源异构性，涵盖文本、图片、表格、API、地理信息系统（GIS）等多种形式，来源包括旅游网站、社交媒体、政府门户、用户终端等，语义粒度与信息密度差异显著。这种数据异构性不仅体现在结构维度，更体现在语义层面——同一概念在不同数据源中存在表述差异、上下文歧义甚至概念重叠[9]。在构建旅游本体过程中，必须对这些异构信息进行统一的语义抽象与概念归一化处理，包括实体识别、概念消歧、上下位关系抽象、语义融合等环节。而现有工具与建模流程在多模态数据（如图文混合、地图叠加）上的支持能力仍较薄弱，缺乏面向异构数据的高效本体建模机制，极大增加了模型构建与维护的复杂度。同时，旅游数据的高时效性与情境性也对本体建模提出挑战。例如，“热门景区”随时间变动而变化，其定义需嵌入实时查询逻辑与上下文计算模型，这对传统静态本体结构提出新的适应要求。

4.2. 本体一致性与维护难题

本体一致性，是保障知识表达规范性、可推理性与可重用性的基础。然而在旅游场景中，因概念演变快、业务扩展强、系统异构严重，导致本体扩展过程易出现语义冲突、结构冗余与逻辑失效等问题[10]。一方面，多个团队或系统在并行建模时，可能因命名规范不一致或上下位划分差异，导致类继承结构紊乱；另一方面，随着业务需求变更，新增概念往往未经过统一审定即接入原有模型，造成语义重叠、属性冗余甚至逻辑矛盾。此外，实例数据的大量动态变更，也可能影响本体中已有规则的适用性与准确性。当前虽然存在若干本体版本控制工具（如 OWL Versioning、ChangeLog 等），但尚未形成面向旅游语义场景的统一变更管理机制与一致性约束策略。如何建立可演化、本体间可对齐、跨平台协同的维护机制，是未来工作的重点方向之一。

4.3. 跨文化适配与语言多样性

旅游活动天然具有跨地域、跨语言与跨文化的全球特性，本体模型若欲广泛部署应用，必须具备良好的语言适配能力与文化语义表达能力。目前多数旅游本体仍停留在以英文或本国语言构建的单语模型阶段，缺乏对多语言词汇对齐、文化习俗差异与表达方式差异的系统支持。例如，在中文语境中“景点”

常含“历史遗迹”“文化象征”意味，而在英语中“attraction”可能泛指一切“具有吸引力的旅游资源”，二者概念外延不一致。此外，一些文化特有概念（如节气、非遗活动、庙会等）在其他语种中缺乏对应表达，需引入文化中介概念或注解性建模策略。国际上如 W3C 的 Multilingual Web 标准、欧洲旅游语义框架已尝试推动多语言对齐与文化本体建模工作，但在旅游领域尚未形成通用技术路径。表 7 结合旅游活动中常见概念，展示不同语言语义适配的若干关键要素，以供参考。

表 7. 旅游本体适配多语多文化的示意表

核心概念	中文表达	英文对应术语	潜在语义差异	跨文化适配建议
景点	景点	Attraction	英语更宽泛	设置子类“文化类景点”以弥补
非遗节庆	非物质文化遗产	Intangible Festival	多数语种无直接对译	添加注释与链接文化背景
游客行为偏好	偏好	Travel Preferences	表达方式与动因差异明显	引入行为意图建模模块
民俗体验	民俗	Folk Experience	文化概念具象化程度不同	配合视频/图像辅助标注

5. 未来研究方向

随着旅游行业数字化、智能化的加速发展，旅游本体作为语义 Web 关键基础设施之一，正面临新一轮技术融合与应用拓展的历史机遇。为充分发挥本体在语义建模、知识组织与智能推理等方面的支撑作用，未来研究应紧扣智能计算、大模型赋能与数据演化等关键趋势，持续推进建模方法革新、技术工具升级与工程应用深化。结合当前研究现状与技术演进动向，本文提出以下三方面的发展重点方向：

5.1. 本体学习与知识图谱结合

传统本体构建依赖人工与规则主导，难以满足大规模、动态化语义建模需求。近年来，本体学习（Ontology Learning）技术不断成熟，可通过机器学习算法从文本、数据库、用户行为等多源数据中自动提取概念、属性及其关系，初步形成语义框架[11]。本体学习不仅提升建模效率，更为大规模知识获取与系统更新提供了方法支撑。与此同时，知识图谱（Knowledge Graph）作为现实世界语义网络的结构载体，日益成为智能搜索、推荐与问答系统的重要基础。旅游领域知识图谱已广泛应用于景点联想、行程规划、偏好匹配等场景，其语义结构与本体高度耦合。未来研究可在旅游本体与知识图谱之间建立双向连接机制，推动本体驱动的图谱结构规范化、推理逻辑显性化，同时利用图谱数据反哺本体动态演化与实例更新，实现语义模型与实际数据的同步发展。

5.2. 面向大规模数据的本体优化

随着旅游数据体量急剧增长，现有本体推理机制在处理大规模异构数据时面临计算效率低、响应延迟高、推理深度受限等问题。为支撑海量旅游信息的实时处理与智能服务，亟须探索面向大数据环境的本体优化策略。其一，可引入增量本体更新机制，在不重构整体本体结构的前提下实现概念扩展、关系调整与实例追加，提升建模响应能力。其二，应推动本体推理引擎的并行化与轻量化，如基于图数据库的分布式推理、基于缓存的推理优化策略等，提升处理性能与系统稳定性。其三，需构建支持“查询即建模”的按需本体加载机制，针对特定用户请求动态生成本体子图，实现语义查询资源最小化调度。

5.3. 与新兴技术融合趋势

新一代人工智能技术的飞速发展，正为本体构建提供新的突破口。特别是大语言模型（LLM）如 ChatGPT 的语义理解与生成能力，已在文本抽象、结构生成与知识归纳等方面展现出显著优势。未来旅游本体建模可引入大模型参与术语生成、属性推理、子类归纳等任务，通过自然语言提示构建本体结构草案，并结合人工审核机制提高自动建模质量[12]。此外，图神经网络（Graph Neural Network, GNN）在

知识表示与关系学习方面的表现，为复杂语义关系的建模提供新路径。通过将旅游本体结构映射为知识图谱节点图谱，可利用图卷积机制挖掘深层次概念关联，提升本体结构的表达能力与自动推理能力，实现“从图中学语义”的目标。未来旅游本体的发展趋势，正从结构设计走向数据驱动、从静态模型迈向动态智能。表 8 对上述关键研究热点进行了汇总，供后续工作参考。

表 8. 旅游本体未来研究热点与关键问题梳理

研究方向	技术路径	关键问题	发展潜力
本体学习与知识图谱结合	本体学习、实体抽取、图谱融合	自动抽取质量、知识融合难度	支撑语义建模自动化
大规模语义数据处理优化	并行推理、增量建模、按需加载	推理效率低、资源消耗大	提升实时语义服务能力
大模型辅助建模与生成技术	ChatGPT、Prompt Engineering	控制精度、语义结构合理性	降低门槛，提升扩展效率
图神经网络与本体表示融合	GNN、图嵌入、节点分类	网络结构复杂性、语义透明性	实现深层次语义建模

6. 结论与展望

本研究围绕“面向语义 Web 的旅游领域本体构建方法”主题，系统梳理了国内外相关研究成果与实践经验，全面分析了本体构建的理论基础、技术方法、工具支撑及典型案例，并重点讨论了当前面临的关键挑战，提出了未来发展方向与技术融合趋势，力求为旅游语义建模提供系统性、可操作的研究框架。本文的主要贡献体现在以下三个方面：一是对旅游本体的类型划分、构建流程与语义 Web 支撑技术进行了结构化梳理，厘清了从概念抽取到模型部署的关键环节；二是比较分析了人工、半自动与自动建模方法的优劣与适用场景，并通过多源案例展示实际建模路径，增强研究的实践可用性；三是结合大数据、人工智能与知识图谱的发展，提出本体学习、图神经网络、语言大模型等新兴技术在本体构建中的融合路径，具有较强的前瞻性与启发性。对于相关研究者而言，本文提供了一份旅游本体构建研究的系统综述，可作为开展后续专题研究、构建原型系统或开发工程应用的知识参考。对于旅游信息服务开发者与知识工程从业人员，本文所汇总的工具比较、流程模板与建模范式，亦具有重要的参考价值。

总体而言，旅游领域语义建模正处于由结构规范向智能建构转型的关键阶段。未来研究应进一步强化跨语言适配、多源知识融合与实时本体演化能力，推动旅游语义技术在推荐服务、智能问答、文化传播等场景中的深入应用，真正实现语义 Web 技术对旅游产业智能化升级的支撑作用。

致 谢

本研究工作得到了广东省教育厅计划国内访学项目的有力支持，特此致以诚挚感谢。访学期间，惠州学院（服务单位）和暨南大学（访学单位）为本人提供了良好的学术环境与周到的行政保障，切实保障了课题研究的顺利开展与阶段目标的如期实现。在此，谨向在本研究过程中给予悉心指导与无私帮助的各级领导、专家学者及相关工作人员表示衷心感谢。

基金项目

本研究获惠州学院横向科研项目【惠东非遗展示馆运营技术服务】（项目编号：18-35900259094），惠州学院高等教育教学改革项目【惠州学院旅游管理专业国际化人才培养路径研究】（项目编号：17-1702242054），惠州市哲学社会科学规划项目【从‘两创’到‘数创’：人工智能驱动下岭南书院文化赋能惠州‘大湾区文化高地’建设路径研究】（项目编号：HZSK2025GJ151/2025ZX114）共同资助。

参考文献

- [1] Lee, T. B., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific american*, 284(5), 34-43.
- [2] Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge acquisition*, 5(2), 199-220.
- [3] Antoniou, G., & Van Harmelen, F. (2004). *A semantic web primer*. MIT press.
- [4] Gómez-Pérez, A., Fernández-López, M., & Corcho, O. (2004). *Ontological Engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*. London: Springer London.
- [5] Maedche, A., & Staab, S. (2005). Ontology learning for the semantic web. *IEEE Intelligent systems*, 16(2), 72-79.
- [6] Al-Aswadi, F. N., Chan, H. Y., & Gan, K. H. (2020). Automatic ontology construction from text: a review from shallow to deep learning trend. *Artificial Intelligence Review*, 53(6), 3901-3928.
- [7] Wache, H., Voegelé, T., Visser, U., Stuckenschmidt, H., Schuster, G., Neumann, H., & Hübner, S. (2001, August). Ontology-based integration of information-a survey of existing approaches. In *Ois@ ijcai*.
- [8] Garcia, R., & Celma, O. (2005, November). Semantic Integration and Retrieval of Multimedia Metadata. In *SemAnnot@ ISWC*.
- [9] Fensel, D., Van Harmelen, F., Horrocks, I., McGuinness, D. L., & Patel-Schneider, P. F. (2005). OIL: An ontology infrastructure for the semantic web. *IEEE intelligent systems*, 16(2), 38-45.
- [10] Flouris, G., Manakanatas, D., Kondylakis, H., Plexousakis, D., & Antoniou, G. (2008). Ontology change: classification and survey. *The Knowledge Engineering Review*, 23(2), 117-152.
- [11] Buitelaar, P., Cimiano, P., & Magnini, B. (Eds.). (2005). *Ontology learning from text: methods, evaluation and applications* (Vol. 123). IOS press.
- [12] Funk, M., Hosemann, S., Jung, J. C., & Lutz, C. (2023). Towards ontology construction with language models. *arXiv preprint arXiv:2309.09898*.