Check for updates

CASE STUDY

REVISED An ontology-driven knowledge graph for tourism information management

[version 2; peer review: 1 approved, 2 approved with reservations]

SUBHASHIS DAS [1], Mayukh Bagchi[2]

[1]Department of Computer Science and Automation, Universidad de Salamanca, Salamanca, Castile and León, 37007, Spain
[2]Department of Information Engineering and Computer Science (DISI), Universita degli Studi di Trento, Trento, Trentino-Alto Adige/South Tyrol, 38123, Italy

v2 **First published:** 08 Jan 2025, **5**:1
https://doi.org/10.12688/openreseurope.17614.1
Latest published: 23 Jul 2025, **5**:1
https://doi.org/10.12688/openreseurope.17614.2

Abstract

Background

Lao Tzu, the ancient Chinese philosopher, had famously said, *"a journey of a thousand miles begins with a single step"*. The modern ecosystem of tourism is *extremely* multi-faceted and multi-context dependent, so much so that tourists, planning on an excursion, have an entire range of queries from as basic as *"Where to eat?"* to as complex as *"List all the eating establishments which are suitable for groups and also provide children menu"*. Such a list becomes even more complicated due to the intermixing of different perceptions, languages, cultures and economies in which tourists are enmeshed in. Accordingly, such increasing complexity in questions from tourists necessitate the need to provide specialized and fine-grained data and information services which are beyond the Information and communications technology (ICT) infrastructures in vogue in tourism today.

Methods

In this paper, we propose and advocate the use of, via a theoretically and implementationally *well-founded methodology*, the Artificial Intelligence (AI) backed technology of *ontology-driven Knowledge Graphs*, to provide fine-grained information services to tourists at any scale and at any level of complexity.

Results and Conclusions

The methodology is validated by developing a fully populated tourism-

Open Peer Review

Approval Status ? ✓ ?

	1	2	3
version 2 (revision) 23 Jul 2025		✓ view	
version 1 08 Jan 2025	? view	? view	? view

1. **Hongyin Zhu**, Qihoo 360, Beijing, China
2. **Rasha M. Ismail**, Ain Shams University Faculty of Computer and Information Science, Cairo, Egypt
3. **Stefan Neubig**, Outdooractive AG, Immenstadt, Germany
Technical University of Munich, ., Germany

Any reports and responses or comments on the article can be found at the end of the article.

specific ontology-driven knowledge graph having 4264 RDF statements and ready for *reuse*, *extension* and *exploitation* as a part of any of the current ICT infrastructure in tourism sector.

Plain Language Summary

Lao Tzu once famously said, “a journey of a thousand miles begins with a single step.” This quote captures the essence of embarking on any journey, including travel. Nowadays, the tourism industry is incredibly complex. When people plan trips, they have a wide range of questions, from simple ones like "Where should I eat?" to complex ones like "Which restaurants are good for groups and offer menus for children?" This complexity is further compounded by the diversity of perceptions, languages, cultures, and economies that tourists encounter.

As a result, there's a growing need for specialized and detailed information services to cater to tourists' varied queries. The current technology used in tourism, like Information and Communication Technology (ICT) infrastructures, may not fully meet these evolving needs.

In this paper, the authors propose using Artificial Intelligence (AI) technology, specifically ontology-driven Knowledge Graphs, to provide tailored information services to tourists regardless of the complexity of their questions. Ontology-driven Knowledge Graphs leverage AI to organize and connect information in a meaningful way, allowing for more nuanced and personalized responses to tourists' inquiries.

The authors validate their proposal by developing a tourism-specific Knowledge Graph populated with over 4000 statements. This Knowledge Graph is designed to be adaptable and expandable, making it a valuable resource for enhancing existing ICT infrastructure within the tourism sector.

In essence, the paper suggests that leveraging AI-driven technology like Knowledge Graphs can revolutionize how tourists access information, providing them with more precise and comprehensive assistance during their travels.

Keywords

Tourism, Ontology Modelling, Faceted Ontology, Knowledge Graph, Conceptual Disentanglement, Tourism and Artificial intelligence (AI)



This article is included in the Marie-Sklodowska-Curie Actions (MSCA) gateway.

Corresponding author: SUBHASHIS DAS (subhashis@usal.es)

Author roles: DAS S: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Bagchi M**: Conceptualization, Validation, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101034371. Funding was also received from the UNITN Matricola Fund 222879.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2025 DAS S and Bagchi M. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: DAS S and Bagchi M. **An ontology-driven knowledge graph for tourism information management [version 2; peer review: 1 approved, 2 approved with reservations]** Open Research Europe 2025, **5**:1
https://doi.org/10.12688/openreseurope.17614.2

First published: 08 Jan 2025, **5**:1 https://doi.org/10.12688/openreseurope.17614.1

REVISED **Amendments from Version 1**

In this revised version of the paper, we have introduced several key updates and refinements to enhance clarity, technical depth, and alignment with the real-world applicability of our approach.

First, we strengthened the motivational framing by explicitly connecting Lao Tzu's quote to the practical challenges tourists face when seeking information. This linkage now clearly illustrates how the conceptual foundation of our work is grounded in both philosophical insight and practical necessity, especially in dealing with heterogeneous and dynamic tourism data needs.

Second, we elaborated on the concept of conceptual disentanglement, emphasizing how different communities of practice—such as tourists, policymakers, and researchers—possess distinct perceptions of the tourism domain. This revision clarifies the necessity of modeling separate but interrelated concept sets within the Tourism Applications Ontology (TAO) and the Tourism Domain Knowledge Graph (TDKG), rather than aiming for a monolithic ontology.

Third, we expanded the discussion on the comparative strengths and limitations of knowledge graph-based AI services versus LLM-driven responses (e.g., ChatGPT). The paper now clearly outlines how the precision, structure, and explainability of KGs complement the flexibility and language understanding capabilities of LLMs, proposing a hybrid architecture for tourism AI.

We also refined the explanation of our four-step faceted ontology modeling methodology, improving the clarity of how competency questions (CQs) are mapped to class–attribute–property patterns. Additional emphasis is placed on the use of authoritative vocabularies (e.g., schema.org, WordNet) and OWL reasoning tools (HermiT, Fact++) to ensure logical and semantic consistency.

Lastly, we updated the completeness evaluation section with clearer references to SPARQL-based assessments (Figure 4–Figure 7), highlighting how each CQ is operationalized and validated using real-world data. These changes collectively enhance the paper's coherence, methodological transparency, and impact for both academic and applied audiences.

1 Introduction

Tourism has facilitated, since time immemorial, both intracultural and intercultural dialogue and exchange amongst otherwise diverse geographies and cultures. Even in its most modern form, tourism has been, and increasingly so, a *"key driver of the global economic recovery, and a vital contributor to job creation, poverty alleviation, environmental protection and multicultural peace and understanding across the globe"*[1]. An essential desideratum which underlies all of the aforementioned ramifications of tourism involves the availability and exploitation of multi-faceted and multi-contextual *information* [see, for instance, 1, 2]. In fact, it wouldn't be an overstatement to label *semantically* annotated and structurally *interrelated* information as the *gamechanger* for the success of a touristic excursion, all the more so in the face of *tourism big data* [see 3].

In continuation of the above context, let us consider the motivating example, for instance, of an *information services chatbot*[4] specialized in providing information about the touristic *Points of Interest* (PoIs) of a region. We are particularly interested in the *design of the knowledge base (KB)*[5] which underpin the multi-faceted informative replies provided by such a chatbot (or, any similar information service-oriented application). There are at least four *pivotal* considerations impacting the design of any such KB. Firstly, there should be a baseline set of *queries* which can be posed to the application (such as, the chatbot) and to which it is expected to factually reply. Secondly, grounded in the above queries, there should be a baseline set of linguistically named *concepts* which compose such a KB and which the KB exploits to generate factual replies. Thirdly, such concepts should be interrelated amongst themselves and organized into a baseline *graph-like conceptual hierarchy* in order to tackle, especially, the more complex queries posed, for example, to the chatbot. Fourthly, and last, each concept in the hierarchy should be described via a baseline set of *attributes* which are exploited to provide *fine-grained* informative answers to the queries vis-`a-vis *ground truth*. Notice that, as a key requirement, each of the above considerations should be *methodologically well-founded* in order to be modifiable and/or evolvable later on (for instance, due to a change in information *sources* or *services*[6]).

We posit that the fundamental bottleneck afflicting the aforementioned design considerations (individually and in unison) can be mapped to the more fundamental phenomenon of *conceptual entanglement*[7,8] ubiquitous in (semantic) data and information representation(s). Firstly, for any *domain of discourse*[9], there is always an unavoidable *manifoldness* (i.e., multiplicity, many-to-many correspondences) between the domain *as such* and its *perception* by an agent (thereby generating *diverse* queries in the process). Secondly, given perception, there

[1]United Nations World Tourism Organization (UNWTO) Report: https://www.e-unwto.org/doi/abs/10.18111/9789284416905

is always an unavoidable *manifoldness* between the perceived domain and the linguistically encoded *body of concepts* an agent utilizes to describe it. Thirdly, given the body of concepts, there is always an unavoidable *manifoldness* in how they can be *hierarchically organized and interrelated*. Finally, given the conceptual hierarchy, there is always an unavoidable *manifoldness* in the way in which each individual concept in the hierarchy can be *intensionally characterized* (i.e., via attributes).

Our proposed solution takes the form of an interdisciplinary, step-by-step *faceted ontology modelling methodology* towards *disentanglement* of the conceptual entanglement issues as described and exemplified above. A domain ontology, defined as *"a formal, explicit specification of a shared conceptualization"*[10], is essentially a computationally exploitable graph of concepts encoding knowledge about a target domain of discourse at different levels of abstraction. The methodology is founded in the interdisciplinary intersection of theories and technologies from faceted knowledge organization (KO)[11], Artificial Intelligence (AI)-based applied ontology[12] and semantic web standards such as RDF, OWL and SPARQL[13]. The key observation is that the methodology, via adapting the best practices individually from tried-and-tested arenas of information and knowledge (base) organization, generates a high-quality ontological knowledge base which is *interoperable* and *reusable* within the larger gamut of *linked open data*[14]. We also validate our methodology by developing a reusable and expandable *Tourism Applications Ontology (TAO)* which models and integrates real-life tourism data of tourist-specific eating establishments from the Italian province of Trentino into a *Tourism Domain Knowledge Graph (TDKG)* (Knowledge Graphs (KGs) being ontologies populated with data; see 15). The ontology-driven knowledge graph can then be exploited both by *humans*, for gaining structured information specific to different information queries, and by computer *applications* as a high-quality, expandable knowledge base for providing advanced information services. Notice that ontology-driven KGs have founded wide success in tackling socio-technical challenges at different levels of complexity in diverse domains as detailed in 15.

The novelty of the proposed solution can be noted from two different perspectives. Firstly, notice that our methodology being crucially based on the principles of *faceted classification*[11] - the principled analysis of a domain of discourse by hierarchically elucidating the *generic terms used to denote its conceptual components* at the requisite level of abstraction - is *innately* capable of disentangling the conceptual entanglement inherent in ontological representations produced by existing state-of-the-art methodologies. Hence, the TAO (at the conceptual level) and the TDKG (at the structured data level) produced following the methodology are taxonomically well-founded and ready to be exploited, expanded and *reused*. The second novelty is the fact that, from a socio-technical perspective, the methodology can provide the starting basis on which an all-encompassing FAIR information organization ecosystem[16,17] can be built which can subsequently be exploited by all players in tourism - tourists, businesses, governments, human and artificial agents, researchers, analysts, etc.

The rest of the paper is organized as follows: Section (2) details the research background of the work founded on a comprehensive review of relevant scientific literature in tourism and Information & Communication Technology (ICT). Section (3) elucidates the proposed faceted ontology modelling methodology in a step-by-step fashion and additionally illustrates each step with running examples grounded in real-world tourism data. Section (4) and Section (5), respectively, provide (via snapshots and illustrations) a glimpse of the technical implementation of the ontology and ontology-driven KG model and an illustrative evaluation of its performance as against (user) queries. Section (6) summarizes and concludes the paper by stressing the socio-technical ramifications of the proposed solution on the tourism domain (especially what concerns its information services).

2 Literature review

We organize the literature review around four dimensions. Firstly, we concentrate on works that deal broadly with the applications of ICT and Artificial Intelligence (AI) techniques in tourism information management. Secondly, we focus on literature which have employed some form of ontology and KG model to manage, exploit and analyze tourism data and information. Thirdly, we also review some of the prominent as well as newly proposed ontology and KG development methodologies (in general but also focused, if existent, on tourism). Finally, we also briefly explore the research in KO and its intersection with ontologies and KGs to contextualize and position the *interdisciplinarity* of the current work.

2.1 Tourism and ICT

The importance of Information and communications technology (ICT), consensually understood as the computing facilities employed for data(base), information and knowledge(base) management[18], in facilitating genuine information access to diverse players in tourism and in allied sectors was stressed in a detailed fashion in 19,20. The advantages that ICT applications and networked data analysis brings towards the fulfilment of Sustainable

Development Goals (SDGs) vis-`a-vis tourism was highlighted in 21. The work in 22 is very interesting from a quantitative perspective as it studied and concluded that ICT infrastructure had a *"positive, statistically significant relationship with tourism development"*, especially, in as large a continent as Africa. The descriptive study in 23 concluded that the tourism industry is heavily influenced by AI technologies, especially by information services *chat-bots* and internet of things (IoT) systems. The recent systematic study of the emerging applications of AI technologies in the tourism sector[24] also vindicated the fact that, even though human touch is still the *"determinant of experiential tourism"*, AI technologies are steadily emerging as high-quality complementary support for the future of tourism. The recent study in 25 also explicated the importance of the availability of structured data and information to develop and strengthen AI technologies in the tourism sector.

2.2 Tourism and ontology-driven KGs

We now briefly review research works which proposed some form of ontology model to facilitate tourism information management. The Harmonise ontology by 26 was mainly proposed for tourism data exchange. The authors in 27 proposed the QALL-ME technique which used controlled vocabulary and top level ontology in addition to WordNet and SUMO to develop the knowledge base for their information management framework (which can be exploited for any domain including tourism). The use of ontology to find semantic similarity between concepts for use in e-Government tourism service recommendation system has been discussed by 28. 29 listed the existing ontologies on the tourism domain. The Mondeca ontology used concepts from the World Tourism Organization thesaurus. The different concepts related to tourism information dissemination are present in the OnTour ontology, which was developed by the eTourism Semantic Web Portal. An ontology-driven tourism recommender system was designed by 30 where algorithms have been used to deal with context of users. An ontology for intelligent transportation systems was developed by 31 using semantic clustering algorithms. Ontology has been also used for medical and leisure tourism recommender systems as is evident from the works in 32 and 33 respectively. 34 tried to model the context in tourism. More recently, the work in 35 proposed an Event-centric Tourism Knowledge Graph (ETKG) which has the capability to model both the temporal and spatial dynamics of tourist trips.

2.3 Ontology and KG development methodologies

Finally, let us concentrate briefly on ontology and KG development methodologies. The survey in 36 provides an overview of some prominent early-generation ontology construction methodologies. The architecture in METHONTOLOGY[37] proposed a *"life cycle to build ontologies based in evolving prototypes"*. Ontology Development 101[38], instead, proposed a flexible methodology which allowed to choose and combine top-down, bottom-up or middle-out approaches in modelling ontologies. The NeOn methodology[39] offers a set of general use patterns for reuse, remixing and merging of ontological resources. More recently, a modular modular ontology modelling approach was proposed in 40 encouraging the reusability of ontologies. Distinct from the above, the use of facet analysis as the key methodological foundation in mainstream ontology engineering was proposed in 41–44 and expanded to ontology-driven KG modelling in 45–47 respectively. Finally, a recent paper[48], though in a slightly different yet related context, elucidated the key critical advantages and differences amongst (some of) the different methodologies referenced above. The present work also exploits the longstanding body of research on foundational ontologies (see, 49–51) and exploits, in particular, the proposal advanced by 49. Two observations. First, notice that while none of the above methodologies or ontological theories were instantiated to tourism data/information, all of them can be, potentially, with requisite tweaks, applied to model and analyze tourism data. Second, none of the above research considers the full spectrum of conceptual entanglement from perception to intensionality (possibly, due to their different goals or focus) and possible solution strategies to disentangle existing entanglement.

2.4 KO and ontology-driven KG development

KO, defined as *"description, indexing and classification performed in libraries, bibliographical databases"*[52], has long been a crucial component of any information management system (whether manual or (semi)automatic). In addition to the concrete classification and indexing systems KO has developed (see, e.g., 53–55) there has been a lot of foundational research on domains, facet analysis and how they can be exploited as a crucial part of any well-founded methodological pipeline to manage data and information. For instance, the very notion of domain analysis[56] and its extended application to modern ontology-based knowledge organization formalisms[57], is functionally linked to *perception* as discussed in conceptual entanglement and can be potentially adapted (in future works) to provide a foundation to disentangle perceptual manifoldness. Within the various theories of knowledge organization (see 58 for an overview), the facet-analytical paradigm of organizing and representing knowledge originally advanced by Ranganathan[59,60] is key to how, in our methodology, the ontological hierarchy

is modelled via facets and specialization of (sub)facets, thereby, facilitating disentanglement of manifoldness in concepts and ontological hierarchies. In fact, recent cumulative work in knowledge organization on how perception, language and ontological hierarchies are functionally linked (see 61–65) has also been key to understand our proposed facet-based ontology modelling methodology. Finally, our work is also orthogonal to recent research within KO as to how the foundational theories and methodologies of KO can be leveraged to model semantically well-founded and ontologically explainable KGs (see, for instance, 66–68). Notice, however, the fact that from none of the above research has been directly instantiated to tourism-based data and information. To that end, the present work is a first preliminary effort.

Chessa *et al.* (2023)[69] propose a robust data-driven methodology for automatically generating tourism-specific knowledge graphs. Their approach leverages structured and unstructured data sources and integrates schema design, data extraction, and KG population into a single pipeline. This work exemplifies how existing methodologies can be pragmatically adapted and implemented in tourism scenarios without necessarily reengineering the conceptual modeling process. Simsek *et al.* (2022)[70] offer a broad knowledge graph engineering perspective, including insights into domain-specific applications such as tourism. Their work emphasizes practical design patterns, toolchains, and knowledge reuse strategies, showing that existing frameworks can be effectively instantiated even in complex domains like tourism without extensive foundational modifications.

2.5 Critical analysis of research gaps

We noticed three research gaps within the aforementioned literature that we reviewed. Firstly, from the broad perspective of ICT and AI applications in tourism data and information management, there is a major *under-utilization* of semantic technologies and semantic knowledge management methodologies. This has consequential (mostly, negative) effects on annotating, modelling, using, analyzing, and reusing tourism data and information by almost all players party to the tourism sector. Secondly, even when some form of ontology (aka KG schema) formalization (or, semantic technologies in general) is employed to model tourism data/information, there is no *methodological framework* to harmonize different perspectives of the same target reality of tourism, resulting in a *conceptually entangled* representation suitable for some and unsuitable for the rest. Thirdly, none of the methodologies surveyed above provides an explicit, step-by-step approach to disentangle the conceptually entangled ontology (aka KG schema) models that are produced (this, being a major factor behind the lack of reusability and expandability of such models).

2.6 Novelty

Notice that, as per the aforementioned overviews of the different dimensions in research literature, the applications of ontologies, conceptual models and/or KGs for tourism data and information modelling are *limited*, and, when some form of ontological model is employed, the focus is on the concrete knowledge artifact (i.e., a specific ontological model or KG) *and not* on the underlying general methodological approach to support its implementation and potential reuse. In fact, none of the above research papers encompass within a single methodological framework the crucial dimensions of *conceptual disentanglement* as follows:

1. *Perception*, a key level towards eliciting the precise purpose behind the ontology-driven knowledge base modelling exercise, thereby, reducing perceptual ambiguity;

2. *Concepts*, a key (knowledge) representational level to elicit how the perception and analysis of a domain (e.g., tourism) translates into semantically unambiguous concepts (classes and properties), thereby, reducing terminological ambiguity;

3. *Hierarchy* which, assuming a shared perception and body of concepts, is important to determine both the *ontological commitment* of concepts[49] (i.e., whether a concept is an artifact, event, mental object, etc.) as well as the precise hierarchical organization and subsumption of the concepts, thereby, eliminating ontological and hierarchical ambiguity, and,

4. *Intensionality* which, assuming all the above levels, is crucial to model the object properties and data properties which interrelated and describe each concept in the hierarchy, thereby, eliminating intensional ambiguity.

The current work is a first attempt to advance a conceptual approach and a well-founded and exemplified methodology to tackle the complexity of the aforementioned levels of conceptual (dis)entanglement within a single approach within the spectrum of tourism data and information. To that end, the first step of the methodology, namely, domain and competency question identification explicates the otherwise implicit and entangled perception of the domain (here, tourism) specific to the purpose of the design of the ontological model composing

a knowledge base. The second step focuses on precisely explicating the otherwise implicit and entangled (linguistically labelled) body concepts which would constitute the ontology-based tourism KG. The third step of the methodology concentrates on explicating the otherwise implicit and entangled ontological and taxonomical assumptions which complicate and potentially entangled the development of an ontology-driven tourism KG. The subsequent steps of the methodology focuses on using standard KR and ontology modelling technologies to model an instantiation of the concrete tourism ontology and ontology-driven KG. Let us now concentrate on the concrete methodology.

3 Methods

As a unified, *well-founded solution* to the aforementioned research gaps, we now concentrate on the steps of our proposed faceted ontology modelling methodology. Before elucidating and illustrating each step, we ground the applicability and exploitability of the methodology in a concrete real-life research background, namely, the different genres of eating establishments available for tourists to choose from while visiting the Italian province of Trentino.

3.1 Research background

The proposed methodology and the base version of the ontology-driven KG model developed following the methodology are mainly inspired from datasets[2] on tourist-specific eating establishments situated in the Italian province of Trentino Alto-Adige[3] (such as, from the data catalog OpenData Trentino). The conceptual analysis of such datasets resulted in a first set of concepts related to tourist-specific *eating establishments* as well as bootstrapping a first version of the hierarchical taxonomy to be adopted as the backbone of the ontology/KG model. However, a central lacking of the datasets was that they contained only very basic information (such as, for instance, the address and the telephone number) about the tourist-specific eating establishments. Therefore, we decided to scrape other sources such as Tripadvisor and Yelp in order to extend the informational attributes provided by OpenData Trentino (and consequently, to design a more generalized ontology i.e. KG schema model). The names of the eating establishments and their geographical placement, including city and address, were freely available on OpenStreetMap and therefore, no copyright infringement occurred. However, some of the labels of the additionally incorporated attributes and fields were altered to avoid copyright infringement. Moreover, synthetic data was created when no open information were found. In addition to the above, further inspiration came from the schema.org concepts of *Food Establishment*[4] and from the *BBC ontology on Food*[5]. These, along with the previously cited sources of inspiration, were reused and appropriately incorporated as reference information modelling standards. The stress on reusing such standards, in particular, was in order to make data integration, i.e. Integrating data from multiple sources to provide users with a unified view[71], as efficient as possible. Given the aforementioned research background, we now sequentially elucidate and illustrate each step of our proposed methodology.

3.2 Step-1: Domain and competency question identification

This step initiates the proposed methodology by conducting a detailed analysis of the domain of interest[72] of which the data is to be structured and modelled as a KB. Such an analysis, in addition to a preliminary conceptual clarification regarding the types of concepts, relations and attributes to be included, also yields concrete requirements modelled as *Competency Questions* (CQs)[73]. A CQ is thus encoded as a natural language sentence that defines a pattern for the types of questions users anticipate an ontology-driven knowledge graph-based knowledge base to address.

For example, tourism as a domain is very vast and therefore it is vital to identify *CQs* as concrete requirements in order to define the boundaries of tourism that one wants to conceptually model as an ontology-driven KG. Consequently, the queries that a user might want to perform on such an ontology-driven KG (such as, particular to tourist-specific eating establishments) are also many and can be of different nature. We mention some illustrative CQs which were found to be interesting as well as simple to elucidate the methodology.

---

[2]see, for e.g., (1) https://dati.trentino.it/dataset/poi-altopiano-di-pine-e-valle-di-cembra; (2) https://dati.trentino.it/dataset/osterie-tipiche-trentine

[3]https://www.regione.taa.it/

[4]https://schema.org/FoodEstablishment

[5]https://iptc.org/thirdparty/bbc-ontologies/fo.html

A general CQ, may be of more interest to foreign tourists, could be *to have a list of the types of eating facilities present in Trento*. This is especially interesting as different countries might have different types of eating establishments (some of which might not exist elsewhere). A prominent example is represented by *trattorias* and *osterias* which are characteristic of Italy. Another CQ that one might want to model in the ontology-driven KG is to *find all places where one can eat a certain meal (e.g. breakfast, lunch or dinner) in Trento or in one of the Trento's suburbs*. A specific example of this type of query is *"Where is it possible to eat lunch in Cognola?"*. It is also possible to ask even more specific queries, such as "Where is it possible to eat pizza for lunch in Cognola?". Notice that, as already alluded to in the introduction, this step is crucial to disentangle the entanglement evident in the perception of a domain (such as tourism).

In fact, queries can be performed to elicit information from the ontology-driven *TDKG* about ratings, prices and amenities of the eating establishments at different levels of abstraction. Some illustrative examples of such queries are listed as follows:

- *CQ1:* "Whether a transportation stop has wheelchair access?"
- *CQ2:* "Whether bicycle is allowed in the trip?"
- *CQ3:* "List all the eating establishments which are suitable for groups and also provide children menu."
- *CQ4:* "What are the nearest eating establishments surrounding a trans-portation stop?"
- *CQ5:* "List all the places where one can eat which have ratings higher than 4 stars."
- *CQ6:* "Give me eating establishments with vegetarian menu and price range medium-low."
- *CQ7:* "Give me the contacts of all the eating establishments that accept reservations and have a parking lot."

As seen from above, these queries not only provide very useful information about eating establishments, but as in the latter examples, they may even provide the direct means to contact the eating establishment. Further, in order to fulfill user expectations, queries might provide options for specifying filters such as menu type or suitability for a specific occasion, such as a romantic date or a business meeting. There can be further queries that can involve the events that are held in eating establishments. For example, one can ask *"Which events will be scheduled in this eating establishment in the future?"*. Or, similarly, given the knowledge of an event, one might ask *"Where and when will this event be held?"*. Tourists might also be interested in knowing whether some famous or extraordinary events were held in the place they are eating right at the moment.

The list of CQs presented above, although already complex to a certain degree, might get even more fine-grained if the domain is expanded in order to include other, for instance, complementary domains. Examples of possible expansion are the domain of transportation, which might be exploited in order to allow the user to know how to reach a certain eating establishment, or the domain of food[74]. Notice also the fact that these CQs perfectly map to our introductory motivating example of developing an ontological KB for a tourism information services chatbot that can provide informative answers to user queries. Finally, observe how this step of domain analysis and competency question identification is functionally linked to explicit disentanglement of the otherwise implicit entanglement and manifoldness observed in instantiations of perception (e.g., in the form queries) in mainstream ontological models and KGs on tourism.

3.3 Step-2: Concept vocabulary determination

Given the conceptual analysis of the domain (e.g., tourism) and the elicitation of requirements as CQs, the second step of the methodology - *concept vocabulary determination* - is focused on choosing *standardized linguistic labels* with unambiguous meaning from relevant lexical resources[75] and/or controlled vocabularies[76] for encoding the concepts expressed in natural language CQs. This facilitates disentanglement of the entanglement evident in the linguistic labelling of the perception of a domain (such as, tourism).

Given the elicited CQs, the main source of the concept vocabulary used for the definition of the terminology used in order to build the *TAO* and later *TDKG* was WordNet[77], a large lexical database of English where nouns[6] are organized into sets of cognitive synonyms (synsets), each representing a unique concept. An exception was made for *osteria* and *trattoria*, two forms of eating establishments typical of Italy. These words, belonging to the Italian

[6]Also verbs, adjectives and adverbs (not relevant here).

language, are also adopted for use in English, as a proper translation is not possible due to the well-known linguistic phenomenon of *lexical gaps*[78]. Since WordNet and other sources, such as Oxford Dictionary, provided only a generic definition of *osteria* and *trattoria* as an *"an Italian restaurant"*, it was decided to use Wikipedia's definition for these two eating establishments as it was the only source providing a clear meaning on what these two concepts are.

The standardized disentanglement of the semantics of linguistic labels encoding perceived concepts is key to mitigate the *semantic heterogeneity* problem, i.e., the difficulty of establishing a certain level of connectivity between people, agents or software systems[79] for the purpose of enabling each of the parties to the appropriately understand exchanged information[80], e.g., as in tourism domain.

3.4 Step-3: Faceted hierarchy creation

Given the elucidation of CQs coupled with the aforementioned concept vocabulary creation, the *faceted hierarchy creation* step models the interrelated hierarchy of classes, relations, and attributes from amongst the unambiguous concept labels obtained in Step (2). The methodology followed to model the classes, relations, and attributes is an adaptation of *faceted classification*[11], a *tried-and-tested* dynamic knowledge organization methodological framework used to classify and interrelate information resources in libraries and information institutions. The illustration of the conceptual hierarchy obtained as a result is illustrated in Figure 1. Notice that this step of facet creation is extremely crucial to disentangle the entanglement implicit in the hierarchical model and intensional definition of any ontology-driven KG-based KB. We now provide an explanation for each of the three sub-components of facet creation, namely, class, attributes, and relations.

3.4.1 Class. Firstly, we model *classes* which represent *"concepts in the domain"*[38]. Some of the classes we considered for *TAO* and which were exploited to model *TDKG* are illustrated as follows:

- *Eating Establishment:* This class describes a general place where people go to eat. It is a specialization of *commercial building* which in turn is a specialization of a *building*. They can be further specialized into several (sub-)classes depending on the differentiating characteristics of the eatery itself, which, in some cases can be very subtle. A list of such sub-classes are presented below:

  – *Bar:* A bar is defined as either a counter where food or drinks can be obtained or a space or establishment where alcoholic beverages are served over a counter.

  – *Bistro:* Bistros are small and informal restaurants that typically serve wine.

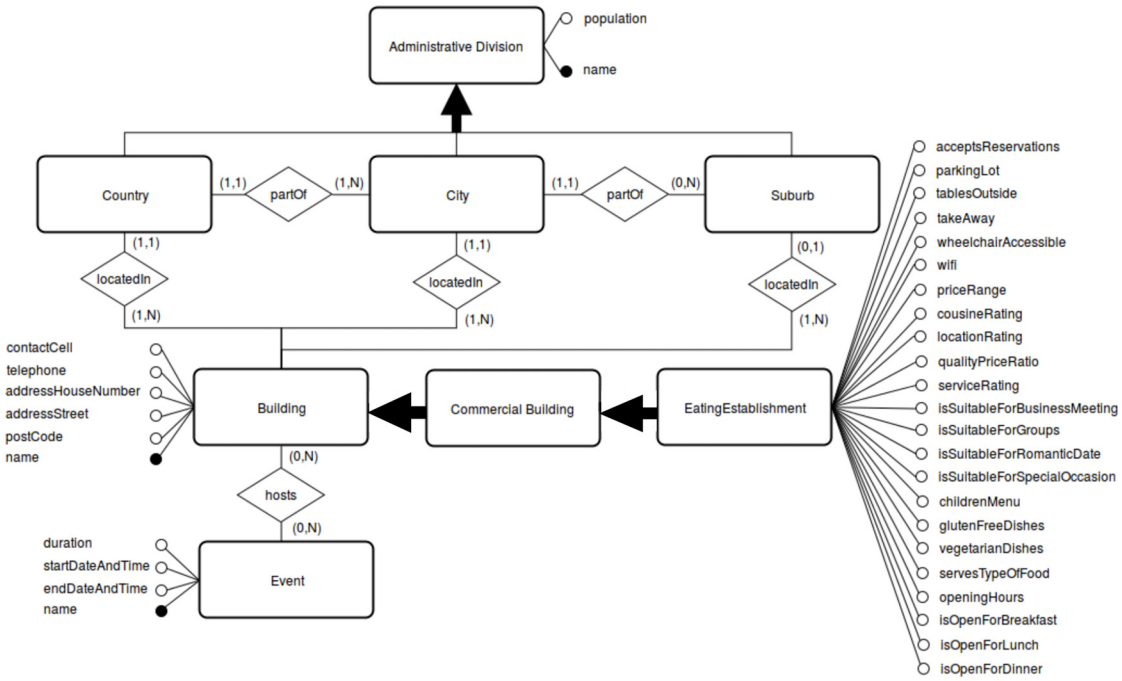

Figure 1. A partial illustration of the Extended ER (EER) Model.

– *Cafe:* A Cafe is similar to a bar but additionally house a small restaurant where drinks and snacks are sold.

– *Fast Food Joint:* Fast food joints are eating establishments where inexpensive food, such as hamburgers or chicken or milkshakes, are prepared and served quickly.

– *Alpine Hut:* Alpine huts are characterized by their geographical position as they are usually located high up in the mountains and are usually accessible only by foot. They are intended to provide food and shelter to tourists who intend to engage in mountaineering, climbing and/or hiking.

Notice also that this concept is more *situational* given that

– *Osteria:* Osterias, as noted beforehand, are typical of Italy. It is sometimes roughly translated to be an inn or tavern. Osterias were originally establishments offering wine and simple meals. Recently, the emphasis has shifted to food, with menus typically remaining concise and highlighting local specialties like pasta, grilled meat, or fish, often served at communal tables. *Pizzeria:* Pizzerias are also originally from Italy but are now commonplace in the rest of the world. They are shops where pizzas are made and sold.

– *Pub:* This type of eating establishment is comparable to a tavern which comprises a building with a bar and public rooms. In pubs, light meals are often provided.

– *Steakhouse:* Steakhouses or chophouses generally serve exclusively meat dishes, typically steaks or chops.

– *Trattoria:* Trattoria is defined as an Italian-style eating establishment, less formal than a restaurant, but more formal than an osteria. Generally, there are no printed menus and wine is sold by the decanter rather than the bottle. Typically, the service is casual and prices are low with the food served being plentiful, mostly following regional and local recipes. In some instances, trattorias even serve food in family-style, i.e., at common tables.

– *Restaurant:* This is a quite general type of eating establishment and is simply defined as a building where people go to eat. Although the definition is *similar* as the one of an eating establishment, it was decided to keep it separate from the main class as in the specific context of Italian culture, it can represent a unified class which can be a (multiple) combination of any of the above classes.

- *Event:* An event is something that happens at a given place and time. We have reused the notion of an *event* as described in the top-level ontology *DOLCE*[81]. An event can be further divided into several sub-classes depending either on the theme or the people it is devoted to. For instance, a children's event is an event devoted to children, whereas, a cultural event relates to the arts and aesthetics that a community practices. A seminar is a special type of cultural event that has the goal of exchanging ideas on a specific theme. A social event, instead, is an event attended by sociable people and has the purpose of sociability. Examples of themes that might characterize an event are music or food.

- *Administrative Division:* Administrative or territorial division is an entity that contains all sets of districts defined for administrative purposes. Examples of these districts used in *TAO*-driven *TDKG* are Country (e.g., Italy), City (e.g., Trento) and Suburb (e.g., Povo, Villazzano and Cognola).

3.4.2 Attributes (Data properties). Attributes (aka data properties) are associated with classes in order to assert their distinctive characteristics. In the proposed *TAO*-driven *TDKG*, the classes for which attributes were defined are *Administrative Division, Building, Commercial Building, Eating Establishment* and *Event*. These attributes can be seen in Figure 1 and are listed as follows:

- We modelled several amenity-specific attributes which include any feature that provides comfort, convenience, or pleasure to tourists. Among these are whether an establishment has a parking lot, tables outside, a take-away service and/or free WiFi. Further, we modelled attributes which also encoded information as to whether an eating establishment is wheelchair accessible and whether it accepts reservation.

- We also modelled attributes that allows users to retrieve information on how to contact the selected eating establishment via a cell phone or telephone number.

- The *is suitable for* attributes was used in order to group all the possible occasions for which an eating establishment might or might not be suitable. Such occasions include business meetings, romantic dates and/or special occasions (each of which are crucial for different genres of tourism[82]). The latter is typically used for tourist-specific celebratory gatherings and can also be used to encode whether an eating establishment is suitable for numerous groups.

- Several attributes were modelled to encode menu options, including, for instance, the different types of menus that might be available in an eating establishment, e.g. children's menu, vegetarian dishes, gluten-free dishes. It can also be exploited to encode the type of cuisine that is served, e.g. Italian, American, Thai etc.
- Rating, as the name suggests, contains numerical evaluations for elements related to the eating establishments such as the cuisine, the service, the location and quality-price ratio.
- Spatial attributes modelled included data such as the address street, the house number and the post code.
- Temporal attributes were used for encoding information related to dates and/or time. It includes attributes for both eating establishments and events. For the former, the opening hours attribute was used in order to list when an eating establishment is open to the public. Moreover, three attributes were used to formally specify whether a place is open for breakfast, lunch and dinner, respectively. For an event which tourists might be interested in, duration (i.e., the number of days), start date-time and end date-time were specified.

Additionally, several other attributes, that were modelled in the *TAO*-driven *TDKG* were *population* (used to capture how many people reside in each administrative division) and *price range* (specific to an eating establishment). The latter is crucial, for instance, for a tourist-specific information services chatbot to encode how expensive an eating establishment is, for instance, in terms of 'Low', 'Medium', 'High', 'Medium-Low' and 'Medium-High'.

3.4.3 Relations (Object properties). Relations (aka object properties) are employed to inter-relate classes among each other as inferred from the requirements modelled as CQs. For example, in our proposed *TAO*-driven *TDKG*, the *part-of* relation has been modelled to encode the inter-relation between a Suburb and a City and between a City and a Country. This is used to express, for instance, that Povo is part of Trento, and Trento is part of Italy. Similarly, the relations *addressCountry*, *addressCity*, and *addressSuburb* are employed to specify geographical locations (the country, city and suburb) of a building. As for the events, we assumed that the events we are interested in are all hosted in a building. Therefore, the hosts and *isHostedBy* relations, which are symmetrical, were defined. There were several other relations modelled in the *TAO*-driven *TDKG* which we don't describe in detail here, such as, *operatedBy*, *stopeAt*, *origin* and *destination*.

Finally, observe how this step of faceted hierarchy creation is functionally linked to explicit disentanglement of the otherwise implicit entanglement and manifoldness observed in two aspects of tourism ontologies and KGs: their ontological commitment and the explainability behind their chosen subsumption hierarchy.

3.5 Step-4: Informal modelling

The informal modeling step is a flexible and repeatable phase where the ontology-driven KG-based KB modeller can interact with different representative stakeholders having differing view points of the domain of discourse (e.g., different players in the tourism sector) to *tune*, consolidate and produce a unified disentangled conceptual model. This tuning is performed using the *Extended Entity-Relationship (EER)* model, an extension of the ER model originally proposed by 83. An Extended Entity-Relationship model is a conceptual data model designed to capture the data requirements for a new knowledge base-driven information system using clear and intuitive graphical notation. The result of the tuning and consolidation performed for tourist-sepcific eating establishments can be partially seen in Figure 1. Notice that, oftentimes, this step yields a harmonized conceptual model encoding different mutually complementary conceptual views (via classes, attributes and relations) achieved in achieved in Step (3) and is, thus, crucial for validating the stratified conceptual disentanglement performed in the previous step. A crucial observation. There have been recent proposals (see, e.g., 40,84) clarifying and advocating the foundations and processes involved in modelling ontologies by reusing existing ontological models or fragments of ontologies. For example, the proposal in 40 advances a step-by-step general workflow to model ontologies by reusing ontology design patterns[85] via a combination of project teamwork and specialized ontology engineering implemented by ontology engineers. The current work has similarities and differences with proposals like the one above. First, the methodology proposed by the current work, though instantiated for tourism domain, is generic and scope, and, therefore, has partial commonalities with any other ontology modelling methodology/process/workflow (including the one above). The *key* difference is, however, the diagnosis and acknowledgement of the stratified problem of conceptual entanglement across the functionally interlinked levels of representation (perception, terminology and concepts, hierarchy, intensionality) and its mapping as to how the proposed facet-analysis based methodology can aid in disentangling the entanglement in each level. In fact, based on the stratification of conceptual entanglement, the reusability of *conceptually disentangled* ontologies and ontology-driven KGs would also be equally stratified: resuability at the level of perception and/or terminological concepts and/or facets and faceted hierarchy and/or intensionality and/or any combinations of the above.

Finally, note that while the methodology is presented in a linear step-by-step fashion in the description (for the sake of simplicity), it is flexible enough to be repeated again and again in a cyclic fashion. To that end, the *TAO* and *TDKG* (described in the next section) versions modelled in a first cycle of the methodology can be, for instance, enriched and expanded with more concepts, properties and corresponding entities if there is a need for the KG to target extended coverage of specific tourism data. Similarly, the reverse can be achieved in subsequent cycles of the methodology if there is a requirement for the KG-backed tourism information service to cover a very specialized tourism theme, e.g., within an exclusive tourism zone.

4 Implementation of *TAO* and *TDKG*

We now validate our proposed methodology by computationally implementing the informal conceptual model arrived in Step (4) of the proposed methodology described and partially illustrated in the previous section.

The computational implementation of the informal conceptual modelling was effectuated via Prot´eg´e, an open-source ontology editor freely available, developed by the Stanford Center for Biomedical Informatics Research at Stanford University School of Medicine. Prot´eg´e employs the OWL ontology language as its underlying Knowledge Representation (KR) formalism. An OWL ontology is composed of three principle elements: individuals, properties (which are divided into object properties and datatype properties) and classes, thereby, mapping one-to-one with our proposed methodology. Individuals represent the objects within a domain, properties define binary relations between them, and classes are interpreted as sets encompassing these individuals. The naming convention that we adopted for modelling the classes, attributes and relations in the OWL version of *TAO* via the Prot´eg´e GUI was grounded in the notation *camelCase*.

All the entities described in Section 3.4.1 were modeled as OWL classes in the OWL version of *TAO*. For some of the classes mentioned in Section 3.4.1, we modelled their Prot´eg´e GUI labels after the preferred names of their semantically equivalent synsets as defined in WordNet. This was done in order to allow, for instance, unhindered data integration in the case of the expansion of the datasets which compose the *TAO*-driven *Tourism Domain Knowledge Graph*. In fact, different sources often use different labels for the same concept. For some classes such as artifact, event and geographical object, *logical disjunction* was defined in order to formalize the fact that these classes represent completely disjoint sets (and thereby, completely different concepts in the target reality). Similarly, disjunction was also defined between *country, city* and *suburb*. On the contrary, for sub-classes of events and of eating establishments, it made no sense to define disjunction due to the fact that a single instance might belong to more than one of these sub-classes. For instance, it is common to have a restaurant that is also a pizzeria, and therefore, the instance needs to be modelled in both classes. Moreover, an alignment with DOLCE top-level ontology[49] was formalized, as can be seen in Figure 2. This is *key* because of the fact that top-level ontologies describe abstract concepts which not only provides the knowledge model with the desired *formal ontological semantics* but also facilitates, non-trivially, data and information interoperability[86].

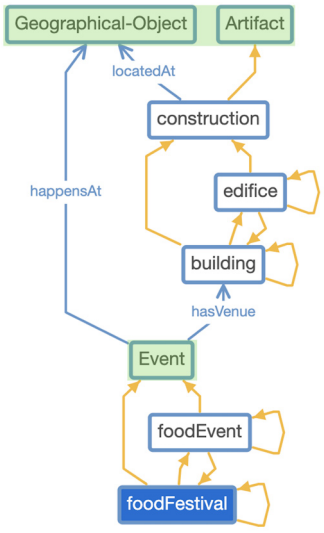

Figure 2. *TAO* alignment with DOLCE Top-level Ontology (DOLCE classes are in snap green).

The relations that were defined in Section 3.4.3 were modelled in the OWL version of *TAO* as object properties. Since it is not desirable to have two relations with the same linguistic label, and since *partOf* is supposed to be used both between city and country and between suburb and city, one of these was modified into *Part of*. While building *TAO*, the cardinalities defined in the EER model in Figure 1 were respected using property restriction constraints as is standard in OWL ontology modelling[87]. The attributes, that can be seen both in Subsection (3.3) and in Figure 1 with a direct connection to the classes they belong to, were modelled as datatype properties in Prot´eg´e. In specific, all the attributes for the class *building* were defined as strings. For events, instead, the properties *start date-time* and *end date-time* were modeled with a with the *dateTime* data type, while for the duration, which represents the number of days a tourist event can last, integer format was used. The integer data type was also used for encoding the population of the administrative divisions. For most of the attributes of the eating establishments, however, the data type employed was *boolean*. Exceptions are represented by the attributes *servesTypeOfFood* and *openingHours* which were modelled strings. The attribute *priceRange* was also modelled as a string, but, additionally, a restriction was implemented in order to allow the selection of one amongst the following five strings: *Low, Medium-Low, Medium, Medium-High, High*. Similarly, for the attribute *rating* which was modelled as a *float*, restrictions were used to allow only these values: *1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5 and 5*. The OWL version of the *TAO* ontology can be accessed here. TAO ontology has Axiom 1172, Logical axiom count 570, Declaration axioms count 351, Class count 186, Object property count 39 and Data property count 76. The full populated Knowledge Graph has 4264 RDF statements.

Given the mapping and population of the *Tourism Applications Ontology (TAO)* with relevant datasets[7], we obtained the *Tourism Domain Knowledge Graph (TDKG)*. The data mapping and population was done semi-automatically using the free and openly-available data mapping tool *Karma*[88]. A fragment of the *TDKG* specific to eating establishments is illustrated in Figure 3. The GraphDB SPARQL construct has been used to create the visualization of *TDKG*. It is to show that how different elements of the *TDKG* are interrelated at the data level. From the graph, the tourist, for instance, can get information that *bus number 5* stops at *Gardolo Materna Paludi* and this bus service is operated by *Trentino Transporti*. It further shows that *Trentino Transporti* also operates regional train service (e.g. Regional5401) which stops at stations *Trento Station FTM* and *Bassano Del Grappa FS*. Other information that we can from the visualized fragment of the *TDKG* is the name of origin and destination of a train trip (e.g. TrentoRoma Termini). The illustration of the entire *TAO* and *TDKG* was omitted as it results in an extremely complex visualization (also, not the focus of this paper).

Finally, having computationally implemented *TAO* and *TDKG*, we are able to produce SPARQL queries[89][8] which are computational query renderings of the natural language CQs elicited in Step (1) of the methodology. The performed queries and the results obtained can be seen in the next section.

5 Evaluation

We performed an evaluation of the ontology and the ontology-driven KG both as computational artifacts but also as a *proxy* for the validity of the proposed methodology itself.

[7]We provided links of some representative datasets in the subsection on *Research Background*. The details of such datasets are not the focus of this paper.

[8]SPARQL being the mainstream query language in ontology-driven KG paradigm to evaluate whether a model has satisfied the requirements it set out to encode.

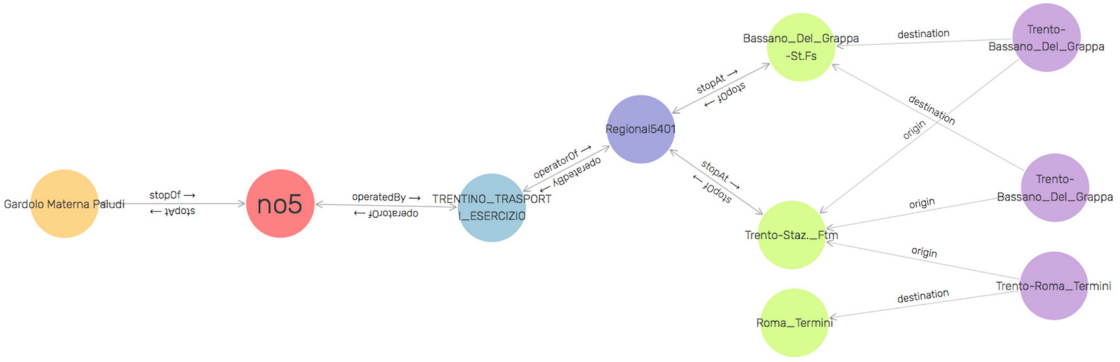

Figure 3. A fragment of the Tourism Domain Knowledge Graph (TDKG) focused on transportation.

For the first aspect, we followed the evaluatory guidelines (both qualitative and quantitative) proposed by Gómez-Pérez[90]. According to Gómez-Pérez[90], the overall goal of evaluating an ontology development process is to check whether the developed ontology definition is, cassificatorily as well as semantically, correct or incomplete or largely incorrect. In concrete terms, the two steps to be followed are: *verification and validation*. The purpose of verification is to check the computational *classificatory correctness* of the developed ontology-driven KG. The purpose of validation, instead, is to check its consistency, completeness, and conciseness. We ascertained the verification dimension of our model via the facility provided by the ontology editor Protégé, by using available reasoners such as HermiT and Fact++ which are available as plugins. The model is concise if it does not accommodate redundancies. In fact, our ontology-driven KG model was found to be largely complete and concise due to careful incorporation of the necessary entity types and properties from the competency questions, thus, capturing what it was purported to represent of the target fragment of tourism. On the other hand, the semantic quality of the ontological commitment (and thereby, of the ontology-driven KG) in terms of consistent modelling of meta-properties like *identity* and *(non)rigidity* was checked and ensured following the *OntoClean* methodology suggested by Guarino and Welty[91].

For the second aspect, we evaluated our ontology-driven KG by modelling the CQs elicited from Step (1) in Section 3.2 as SPARQL queries[89] and adjudicating the answers returned. In fact, to answer *CQ1*, we attached the data property *wheelchairAccess* with the class *transportationPoint* and set the datatype of the data property as Boolean. A snapshot of the SPARQL query structure and the corresponding answer is shown in Figure 4.

```
PREFIX rdf : <http://www.w3.org/rdf-syntax-ns#>
PREFIX owl:< http://www.w3.org/2002/07/owl#>
PREFIX rdfs : <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd:< http://www.w3.org/2001/XMLSchema#>
PREFIX gtfs : < http://purl.org/net/tour-gtfs#>
SELECT? StopCode? StopName? wheelchair Access
WHERE {
 ? bus Stop gtfs : stopCode? StopCode.
 ? bus Stop gfs : name? StopName.
 ? bus Stop gtfs : wheelchair Access? wheelchair Access.
}ORDERBY?name LIMIT 5
```

Similarly, for *CQ2*, we created *bikesAllowed* as a datatype property and attached it with the class *trip*. Figure 5 shows a snapshot of the result from the query.

```
PREFIX rdf : <http://www.w3.org/rdf >
PREFIX owl : <http://www.w3.org/2002/07/owl#>
PREFIX rdfs : <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd : <http://www.w3.org/2001/XMLSchema#>
PREFIX gtfs : <http://purl.org/net/tour-gtfs#>
SELECT? TripName? Trip Id? bikes Allowed? bar
WHERE { ? TripName gtfs:trip Id? Trip Id.
       ? TripName gtfs:bikes Allowed? bikes Allowed.
? TripName  gtfs:bar? bar }
```

StopName	wheelchairAccess	StopCode
"S.Ilario Sav"^^<ht	"true"^^<http://	"31171x"
"Lamar Ftm"^^<ht	"false"^^<http:/,	"22205z"
"Cognola Grezoni"^	"true"^^<http://	"24030z"
"Cognola Toresela"^	"true"^^<http://	"24015x"
"Cognola S.Vito"^^	"true"^^<http://	"24035x"

Figure 4. A snapshot of the SPARQL query result for wheelchair access.

The results of the SPARQL query for *CQ3* is partially illustrated in Figure A1[92]. To answer *CQ4*, we used Apache Jena Fuseki[9] plugin to allow running this typical type of geospatial query. A snapshot of the answer to this query is shown in Figure 6.

```
PREFIX rdf : <http://www.w3.org/rdf>
PREFIX owl : <http://www.w3.org/2002/07/owl#>
PREFIX rdfs : <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd : <http://www.w3.org/2001/XMLSchema#>
PREFIX gtfs : <http://purl.org/net/tour-gtfs#>
PREFIX spatial : <http://jena.apache.org/spatial#>
SELECT * {
   ? busStop spatial : nearby (46.0811. 061’ miles’);
       gtfs : name?name
}
```

TripName	TripId	bikesAllowed	bar
Trento-Bassano_Del_Grappa	"5401"^^"true"^^<ht	"false"^^	
Trento-Bassano_Del_Grappa	"5405"^^"true"^^<ht	"false"^^	
Trento-Roma_Termini	"9477"^^"false"^^<ht	"true"^^	

Figure 5. A snapshot of the SPARQL query result for bikes allowed.

```
SELECT  distinct (?e as ?facility) ?type ?street ?number ?schedule
WHERE {
        ?e rdf:type ?type .
        ?type rdfs:subClassOf* rest:eatingEstabilishment .
        ?e rest:isSuitableForGroups ?g .
        ?e rest:childrenMenu ?c .
        ?e rest:addressStreet ?street .
        ?e rest:addressHouseNumber ?number .
        ?e rest:openingHours ?schedule
        FILTER(?c = true) .
}
```

facility	type	street	number	schedule
OrsoGrigio	trattoria	"Via Degli Orti"@	"19"@	"lunedì 12:00 - 00:00 martedì 12:00 - 00:00 mercoledì 12:00 - 00:00 giovedì 12:00 - 00:00 venerdì 12:00 - 00:00 sabato 12:00 - 0
RoadhouseGrill	steakhouse	"Via del Brennero"@	"146"@	"domenica 12:00 - 15:0018:30 - 23:30 lunedì 12:00 - 15:0018:30 - 23:30 martedì 12:00 - 15:0018:30 - 23:30 mercoledì 12:00 - 15
McDonald's	fastFood	"Via di Madonna Bianca"@	"5"@	"domenica 11:00 - 01:00 lunedì 11:00 - 01:00 martedì 11:00 - 01:00 mercoledì 11:00 - 01:00 giovedì 11:00 - 01:00 venerdì 11:00
RosaD'Oro	restaurant	"Piazza Santa Maria Maggiore"(	"21"@	"lunedì 12:00 - 14:30 19:00 - 00:00 martedì 12:00 - 14:30 19:00 - 00:00 mercoledì 12:00 - 14:30 19:00 - 00:00 giovedì 12:00 - 14
Tognola	mountainHut	"Via Passo Rolle| Alpe Tognola"	"21"@	"domenica 08:30 - 17:00 lunedì 08:30 - 17:00 martedì 08:30 - 17:00 mercoledì 08:30 - 17:00 giovedì 08:30 - 17:00 venerdì 08:30
Saporoso	restaurant	"Via Dei Mille"@	"1"@	"domenica 12:00 - 14:30 18:30 - 22:30 martedì 12:00 - 14:30 18:30 - 22:30 mercoledì 12:00 - 14:30 18:30 - 22:30 giovedì 12:00 -
HappyHour	restaurant	"Via Fratelli Perini"@	"57"@	"domenica 11:30 - 14:30 martedì 11:00 - 14:3018:30 - 22:30 mercoledì 11:00 - 14:3018:30 - 22:30 giovedì 11:00 - 14:3018:30 - 2
Piedicastello	trattoria	"Piazza di Piedicastello"@	"2"@	"domenica 12:00 - 14:00 lunedì 12:00 - 14:0019:30 - 22:00 martedì 12:00 - 14:0019:30 - 22:00 mercoledì 12:00 - 14:0019:30 - 22
AnticaBirreriaPedavena	pub	"Piazza di Fiera"@	"13"@	"domenica 09:00 - 00:30 lunedì 09:00 - 00:30 martedì 09:00 - 00:30 mercoledì 09:00 - 00:30 giovedì 09:00 - 00:30 venerdì 09:00
AnticaTrattoriaDueMori	trattoria	"Via San Marco"@	"11"@	"domenica 19:00 - 22:0012:00 - 14:30 martedì 12:00 - 14:3019:00 - 22:30 mercoledì 12:00 - 14:3019:00 - 22:30 giovedì 12:00 - 1
Anfiteatro	steakhouse	"Piazzetta Dell'Anfiteatro | Vicol	"10"@	"lunedì 12:00 - 14:3018:00 - 00:00 martedì 12:00 - 14:3018:00 - 00:00 mercoledì 12:00 - 14:3018:00 - 00:00 giovedì 12:00 - 14:3
SaporiMediterranei	restaurant	"Via R. da Sanseverino | Vicino	"39/4"@	"domenica 12:00 - 14:3019:00 - 23:30 lunedì 12:00 - 14:3019:00 - 23:30 martedì 12:00 - 14:3019:00 - 23:30 mercoledì 12:00 - 14
SaporiMediterranei	pizzeria	"Via R. da Sanseverino | Vicino	"39/4"@	"domenica 12:00 - 14:3019:00 - 23:30 lunedì 12:00 - 14:3019:00 - 23:30 martedì 12:00 - 14:3019:00 - 23:30 mercoledì 12:00 - 14
OldWildWest	fastFood	"Piazza Delle Ddonne Lavoratri	"13"@	"domenica 11:30 - 00:30 lunedì 12:00 - 15:3018:30 - 00:30 martedì 12:00 - 15:3018:30 - 00:30 mercoledì 12:00 - 15:3018:30 - 00
Niky's	restaurant	"Via San Pio X"@	"29"@	"lunedì 12:00 - 14:3018:30 - 23:00 martedì 12:00 - 14:3018:30 - 23:00 mercoledì 12:00 - 14:3018:30 - 23:00 giovedì 12:00 - 14:3

Figure A1. snapshot of the SPARQL query result for eateries suitable for groups and that have children menu.

StopId	busStop	hotel
"21" ^^xsd:string	"Cognola Centro Civicc	"Hotel e Ristorante Villa Madruzzo"
"21" ^^ xxd: strin	"Cognola Centro Civicc	"Albermonaco"
"21" ^^ xxd: strin	"Cognola Centro Civicc	"Hotel Everest-Trento"

Figure 6. A snapshot of the SPARQL query result for near by services.

[9]https://jena.apache.org/documentation/fuseki2/

Finally, we also show, partially the results for the SPARQL query pertaining to the *CQ5 - "List all the places where one can eat which have ratings higher than 4 stars"* in Figure 7. We omit the illustration of the rest of the CQs elucidated in Step (1) of the methodology to avoid repetition.

```
PREFIX rdfs : <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf : <http://www.w3.org/rdf-syntax-ns#>
PREFIX rest : <http://purl.org/net/tour#>
SELECT  DISTINCT? eating Establishment? Service Rating? Cousine Rating
WHERE
        { ? eating Establishment rdf : type? type.
? type rdfs : sub Class Of * rest:eating Estabilishment.
? eating Establishment rest : service Rating? Service Rating.
? eating Establishment rest : cousine Rating? Cousine Rating
FILTER (? Cousine Rating > 4).
}
Order  By? eating Establishment
```

In addition to the above twin evaluation, we also validate qualitatively as to how the proposed faceted ontology modelling methodology tackles the disentanglement of the otherwise ubiquitous problem of conceptual entanglement as stated in the introduction. To that end, the first step (Step-(1)) of the methodology concentrates on a precise documented characterization of domain analysis and competency questions identification, thereby, disentangling the complexity and multiplicity of assumptions in tourist perception (and queries). Notice that, due to the very (non-capturable) nature of perception[93], this step attempts reduction (and not complete elimination) of perceptual entanglement and perceptual complexity. Notice that, as briefed before, there can be successive cycles of domain analysis and competency questions enrichment/extension/abstraction in sync with successive implementational cycles of the methodology to satisfy the target coverage and informational scalability of the KG-backed tourist information service. The second step (Step-(2)) focuses on concept vocabulary determination and to that end, disentangles, the terminological entanglement otherwise prevalent in tourism ontological models. This facilitates elimination of the complexity and multiplicity of assumed meanings of terms which are used to linguistically denote concepts and properties composing the schema, e.g., *TAO* of the ontology-driven KG, e.g., *TDKG*. This step, as before, can also be repeated in successive cycles to enrich/extend/abstract the body of terms in response to increase or decreased requirements of scalability in tourist information services. The next step (Step-(2)) concentrates on the faceted hierarchy creation and sequentially tackles the rest of the levels evident in conceptual entanglement. It eliminates the complexity and multiplicity of assumptions in hierarchical entanglement by precisely modelling the subsumption hierarchy of facets and sub-concepts denoting the concepts composing the ontology-driven KG. It also eliminates the otherwise complexity and multiplicity of assumptions in intensional entanglement by precisely decorating each concept with a precise set of object properties (relating it with other concepts) and a precise set of data properties

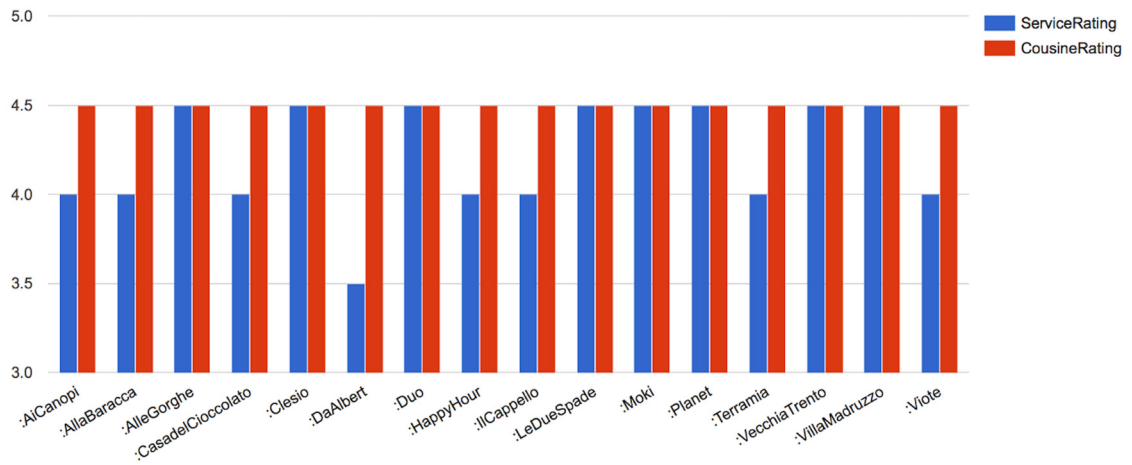

Figure 7. A snapshot of the SPARQL query result for quality of restaurants in Trento.

(describing it via a purpose-specific set of attribute metadata). As with the other steps, successive cycles of faceted hierarchy modelling is *key* to manage the desired level of data and informational scalability to be provided to tourists via the KG-backed information service.

Notice that, as briefed before, there can be successive cycles of domain analysis and competency questions enrichment/extension/abstraction in sync with successive implementational cycles of the methodology to satisfy the target coverage and informational scalability (e.g. with new data as in 94, 95) of the KG-backed tourist information service.

6 Conclusive discussion

In this paper, we proposed a faceted ontology modelling methodology for modelling, managing and analyzing diverse genres of tourism-specific domain information. The proposed methodology is theoretically and implementationally well-founded and can, to a promising extent, disentangle the conceptual entanglement inherent in state-of-the-art tourism-specific ontology-driven KGs. We also validated the proposed methodology by modelling a base version of a tourism ontology-driven KG, the results of which are promising and *extensible for reuse*. The *key* conclusive observation, from a socio-technical perspective of tourism, is that the paper highlighted a *duality* which is *ubiquitous* for tourism-specific information services, i.e., the need for acknowledging the different perceptions of different stakeholders from the same piece of target reality *as well as* the need to accommodate and harmonize such different views into a methodologically well-founded information services solution. Notice that the proposed methodology (and the baseline *TAO* and *TDKG* produced) are capable, to a large extent, to disentangle the entanglement inherent in answers provided, for instance, by the tourist information services chatbot in the motivating example (and similar information and knowledge-intensive applications).

Declarations

- Ethics approval: Not applicable
- Consent to participate: Not applicable
- Consent for publication: Yes
- Authors' contributions: Both authors contributed equally to this work.

Data availability

No data are associated with this article.

Software availability

Repository Name	Dataset description
Zenodo	Web Ontology Language (OWL) file of the DAS, S. (2024). Restaurant OWL file [Data set]. Zenodo. https://zenodo.org/doi/10.5281/zenodo.12731788 Data are available under the terms of the GNU General Public License v3.0 or later.
Zenodo	List of the SPARQL queries used to generated the answer available in text format. DAS, S. (2024). SPARQL query for eating establishment [Data set]. Zenodo. https://doi.org/10.5281/zenodo.12799173 Data are available under the terms of the GNU General Public License v3.0 or later.
Zenodo	Snapshot of the SPARQL query result for eateries suitable for groups and that have children's menus. The table of the Image is generated using open source ontology editor Protégé (https://protege.stanford.edu/) DAS, S. (2024). snapshot of the SPARQL query result for eateries suitable for groups and that have children menu. Zenodo. https://doi.org/10.5281/zenodo.13120538 Data are available under the terms of the GNU General Public License v3.0 or later.

References

1. Manfredo MJ: **An investigation of the basis for external information search in recreation and tourism.** *Leis Sci.* 1989; **11**(1): 29–45.
Publisher Full Text

2. Kambele Z, Li G, Zhou Z: **Travelers' information-seeking behaviors.** *Journal of Travel & Tourism Marketing.* 2015; **32**(1–2): 141–152.
Publisher Full Text

3. Li J, Xu L, Tang L, *et al.*: **Big data in tourism research: a literature review.** *Tour Manag.* 2018; **68**: 301–323.
Publisher Full Text

4. Bagchi M: **Conceptualising a library chatbot using open source conversational Artificial Intelligence.** *DESIDOC Journal of Library & Information Technology.* 2020; **40**(6): 329–333.
Reference Source

5. Guarino N, Giaretta P: **Ontologies and knowledge bases. Towards very large knowledge bases.** 1995; 1–2.
Reference Source

6. Singh G: **Information sources, services and systems.** PHI Learning Pvt. Ltd., New Delhi, 2013.
Reference Source

7. Bagchi M, Das S: **Conceptually disentangled classificatory ontologies.** In: *15th Seminar on Ontology Research in Brazil (ONTOBRAS), Virtual.* 2022.
Reference Source

8. Bagchi M, Das S: **Disentangling domain ontologies.** In: *19th IRCDL (The Conferenceon Informationand Research Science Connecting to Digital and Library Science)*. Bari, Italy, 2023.
Reference Source

9. Guarino N, Oberle D, Staab S: **What is an ontology?** In: *Handbook on Ontologies.* Springer, Berlin/Heidelberg, 2009; 1–17.
Publisher Full Text

10. Studer R, Benjamins VR, Fensel D: **Knowledge engineering: principles and methods.** *Data Knowl Eng.* 1998; **25**(1–2): 161–197.
Publisher Full Text

11. Ranganathan SR: **Prolegomena to Library classification.** Asia Publishing House, New York, 1967.
Reference Source

12. Guarino N, Musen M: **Applied ontology: the next decade begins.** *Appl Ontol.* 2015; **10**(1): 1–4.
Publisher Full Text

13. Hitzler P: **A review of the semantic web field.** *Commun ACM.* 2021; **64**(2): 76–83.
Publisher Full Text

14. McCrae JP, Abele A, Buitelaar P, *et al.*: **The linked open data cloud.** Lod- cloud. net 2019.
Reference Source

15. Hogan A, Blomqvist E, Cochez M, *et al.*: **Knowledge graphs.** *ACM Computing Surveys (CSUR).* 2021; **54**(4): 1–37.
Publisher Full Text

16. Wilkinson MD, Dumontier M, Aalbersberg IJJ, *et al.*: **The FAIR guiding principles for scientific data management and stewardship.** *Sci Data.* 2016; **3**(1): 160018.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

17. Bagchi M: **Towards Knowledge Organization Ecosystem (KOE).** *Cat Classif Q.* 2021; **59**(8): 740–756.
Publisher Full Text

18. Bagchi M: **Smart cities, smart libraries and smart knowledge managers: Ushering in the neo-knowledge society.** 50 Years of LIS Education in North East India, 2019.
Publisher Full Text

19. Shanker D: **ICT and tourism: challenges and opportunities.** 2008.
Reference Source

20. Bethapudi A: **The role of ict in tourism industry.** *Journal of Applied Economics and Business.* 2013; **1**(4): 67–79.
Reference Source

21. Gössling S: **Technology, ict and tourism: from big data to the big picture.** *J Sustain Tour.* 2021; **29**(5): 849–858.
Publisher Full Text

22. Adeola O, Evans O: **ICT, infrastructure, and tourism development in Africa.** *Tourism Economics.* 2020; **26**(1): 97–114.
Publisher Full Text

23. Zsarnoczky M, *et al.*: **How does Artificial Intelligence affect the tourism industry?** *VADYBA.* 2017; **31**(2): 85–90.
Reference Source

24. Samala N, Katkam BS, Bellamkonda RS, *et al.*: **Impact of AI and robotics in the tourism sector: a critical insight.** *Journal of Tourism Futures.* 2022; **8**(1): 73–87.
Publisher Full Text

25. Bulchand-Gidumal J: **Impact of Artificial Intelligence in travel, tourism, and hospitality.** In: *Handbook of e-Tourism.* Springer, 2022; 1943–1962.
Publisher Full Text

26. Fodor O, Werthner H: **Harmonise: a step toward an interoperable e-tourism marketplace.** *Int J Electron Comm.* 2005; **9**(2): 11–39.
Publisher Full Text

27. Ferrández Ó, Spurk C, Kouylekov M, *et al.*: **The QALL-ME framework: a specifiable-domain multilingual Question Answering architecture.** *J Web Semant.* 2011; **9**(2): 137–145.
Publisher Full Text

28. Al-Hassan M, Lu H, Lu J: **A semantic enhanced hybrid recommendation approach: a case study of e-Government tourism service recommendation system.** *Decis Support Syst.* 2015; **72**: 97–109.
Publisher Full Text

29. Mathur A, Akshatha K, Shastry A, *et al.*: **A survey on existing tourism ontologies.** *IJRET: International Journal of Research in Engineering and Technology.* 2015; 20–23.
Reference Source

30. Chu Y, Wang H, Zheng L, *et al.*: **TRSO: a tourism recommender system based on ontology.** In: *International Conference on Knowledge Science, Engineering and Management.* Springer, 2016; 567–579.
Publisher Full Text

31. Gregor D, Toral S, Ariza T, *et al.*: **A methodology for structured ontology construction applied to intelligent transportation systems.** *Comp Stand Inter.* 2016; **47**: 108–119.
Publisher Full Text

32. Frikha M, Mhiri M, Zarai M, *et al.*: **Using TMT ontology in trust based medical tourism recommender system.** In: *2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA).* IEEE, 2016; 1–8.
Publisher Full Text

33. Lee CI, Hsia TC, Hsu HC, *et al.*: **Ontology-based tourism recommendation system.** In: *2017 4th International Conference on Industrial Engineering and Applications (ICIEA).* IEEE, 2017; 376–379.
Publisher Full Text

34. Kashevnik AM, Ponomarev AV, Smirnov AV: **A multimodel context-aware tourism recommendation service: approach and architecture.** *J Comput Syst Sci Int.* 2017; **56**(2): 245–258.
Publisher Full Text

35. Wu J, Zhu X, Zhang C, *et al.*: **Event-centric tourism knowledge graph—a case study of Hainan.** In: *International Conference on Knowledge Science, Engineering and Management.* Springer, 2020; 3–15.
Publisher Full Text

36. Fernández-López M: **Overview of methodologies for building ontologies.** In: *IJCAI99 Ontology Workshop.* Citeseer, 1999; 430.
Reference Source

37. Fernández-López M, Gómez-Pérez A, Juristo N: **Methontology: from ontological art towards ontological engineering**. 1997.
Reference Source

38. Noy NF, McGuinness DL: **Ontology development 101: a guide to creating your first ontology**. Stanford knowledge systems laboratory technical report KSL-01-05 and, 2001. **Reference Source**

39. Suárez-Figueroa MC, Gómez-Pérez A, Fernández-López M: **The neon methodology for ontology engineering.** In: *Ontology Engineering in a Networked World.* Springer, Berlin/Heidelberg, 2012; 9–34.
Publisher Full Text

40. Shimizu C, Hammar K, Hitzler P: **Modular ontology modeling.** *Semant Web.* 2023; **14**(3): 459–489.
Publisher Full Text

41. Das S, Roy S: **Faceted ontological model for brain tumour study.** *Knowl Org.* 2016; **43**(1): 3–12.
Publisher Full Text

42. Das S: **Domain modeling theory and practice**. PhD thesis, University of Trento, 2018.
Reference Source

43. Ghosh SS, Das S, Chatterjee SK: **Human-centric faceted approach for ontology construction.** *Knowl Org.* 2020; **47**(1): 31–44.
Publisher Full Text

44. Das S, Naskar D, Roy S: **Reorganizing educational institutional domain using faceted ontological principles.** *Knowl Org.* 2022; **49**(1): 6–21.
Publisher Full Text

45. Bagchi M: **A knowledge architecture using knowledge graphs**.

Master's thesis, Indian Statistical Institute, Bangalore, 2019.

46. Bagchi M: **A large scale, knowledge intensive domain development methodology.** *Knowl Org.* 2021; **48**(1): 8–23.
Publisher Full Text

47. Bagchi M: **A diversity-aware domain development methodology.** *PhD Symposium, 41st International Conference on Conceptual Modelling (ER 2022).* 2022.
Publisher Full Text

48. Das S, Bagchi M, Hussey P: **How to teach domain ontology-based Knowledge Graph construction? An irish experiment.** *Knowl Orgn.* 2023; **50**(3): 182–201.
Publisher Full Text

49. Borgo S, Ferrario R, Gangemi A, *et al.*: **DOLCE: a Descriptive Ontology for Linguistic and Cognitive Engineering.** *Appl Ontol.* 2022; **17**(1): 45–69.
Publisher Full Text

50. Smith B, Kumar A, Bittner T: **Basic formal ontology for bioinformatics**. 2005.
Reference Source

51. Guizzardi G, Botti Benevides A, Fonseca CM, *et al.*: **UFO: Unified Foundational Ontology.** *Appl Ontol.* 2022; **17**(1): 167–210.
Publisher Full Text

52. Hjørland B: **What is Knowledge Organization (KO)?** *Knowl Orgn.* 2008; **35**(2–3): 86–101.
Publisher Full Text

53. Satija MP: **The theory and practice of the Dewey Decimal Classification system**. Elsevier, Amsterdam, 2013.
Reference Source

54. McIlwaine IC: **The Universal Decimal Classification: some factors concerning its origins, development, and influence.** *J Am Soc Inf Sci.* 1997; **48**(4): 331–339.
Publisher Full Text

55. Ranganathan S: **A descriptive account of the colon classification**. Sarada Ranganathan Endowment for Library Science, 1967.
Reference Source

56. Hjørland B, Albrechtsen H: **Toward a new horizon in information science: domain-analysis.** *J Am Soc Inf Sci.* 1995; **46**(6): 400–425.
Publisher Full Text

57. Smiraglia R: **Domain analysis for Knowledge Organization: tools for ontology extraction**. Chandos Publishing, Witney, England, 2015.
Reference Source

58. Hjørland B: **Theories of Knowledge Organization—theories of knowledge.** *Knowl Orgn.* 2014; **40**(3): 169–181.
Publisher Full Text

59. Ranganathan S: **Depth classification.** In: *Annals of Library Science.* NISCAIR-CSIR, India, New Delhi, 1954.
Reference Source

60. Ranganathan SR: **Prolegomena to library classification**. Asia Publishing House, Bombay and New York, 1967.
Reference Source

61. Giunchiglia F, Bagchi M: **Millikan + Ranganathan - from perception to classification.** In: *5th Cognition and OntologieS (CAOS) Workshop, Colocated with the 12th International Conference on Formal Ontology in Information Systems (FOIS)*. 2021.
Reference Source

62. Giunchiglia F, Bagchi M: **Object recognition as classification via visual properties.** In: *17th International ISKO Conference and Advances in Knowledge Organization.* Aalborg, Denmark, 2022.
Publisher Full Text

63. Giunchiglia F, Bagchi M, Diao X: **A semantics-driven methodology for high-quality image annotation.** In: *26th European Conference on Artificial Intelligence (ECAI 2023).* arXiv preprint arXiv: 2307.14119, Kraków, Poland, 2023.
Publisher Full Text

64. Giunchiglia F, Bagchi M, Diao X: **Aligning visual and lexical semantics.** In: *International Conference on Information (iConference) 2023.* Springer, 2023; 294–302.
Publisher Full Text

65. Giunchiglia F, Bagchi M: **From knowledge representation to Knowledge Organization and back.** In: *International Conference on Information (iConference).* arXiv Preprint arXiv: 2312.07302, Springer, 2024.
Publisher Full Text

66. Xiao G: **Explainable AI: linking human and machine.** *Knowl Orgn.* 2019; **46**(5): 398–399.
Publisher Full Text

67. Qin J: **Knowledge Organization and representation under the AI lens.** *J Data Inf Sci.* 2020; **5**(1): 3–17.
Publisher Full Text

68. Xiao G: **The brain is a knowledge graph.** *Knowl Org.* 2019; **46**(1): 71–71.
Publisher Full Text

69. Chessa A, Fenu G, Motta E, *et al.*: **Data-driven methodology for knowledge graph generation within the tourism domain.** *IEEE Access.* 2023; **11**: 67567–67599.
Publisher Full Text

70. Simsek U, Kärle E, Angele K, *et al.*: **A knowledge graph perspective on knowledge engineering.** *SN Comput Sci.* 2022; **4**(1): 16.
Publisher Full Text

71. Lenzerini M: **Data integration: a theoretical perspective.** In: *Proceedings of the Twenty-first ACM Sigmod-sigact-sigart Symposium on Principles of Database Systems.* 2002; 233–246.
Publisher Full Text

72. Hjørland B: **Domain analysis.** *Knowl Org.* 2017; **44**(6): 436–464.
Publisher Full Text

73. Grüninger M, Fox MS: **The role of competency questions in enterprise engineering.** In: *Benchmarking—Theory and Practice.* Springer, Berlin/Heidelberg, 1995; 22–31.
Publisher Full Text

74. Infomuse: **Instructions for developing faceted classification.** 2013; accessed 29 November 2022.
Reference Source

75. Batsuren K: **Understanding and exploiting language diversity**. PhD thesis, University of Trento, 2018.
Reference Source

76. Fidel R: **Who needs controlled vocabulary**. Special libraries, 1992; **83**(1): 1–9.
Reference Source

77. Miller GA: **Wordnet: a lexical database for english.** *Commun ACM.* 1995; **38**(11): 39–41.
Publisher Full Text

78. Bentivogli L, Pianta E: **Looking for lexical gaps.** In: *Proceedings of the Ninth EURALEX International Congress*. Stuttgart: Universität Stuttgart, 2000; 8–12.
Reference Source

79. Uschold M, Gruninger M: **Ontologies and semantics for seamless connectivity.** *ACM SIGMod Record.* 2004; **33**(4): 58–64.
Publisher Full Text

80. Pollock J: **Integration's dirty little secret: it'sa matter of semantics**. Whitepaper, Modulant: The Interoperability Company, 2002.

81. Gangemi A, Guarino N, Masolo C, *et al.*: **Sweetening ontologies with Dolce**. In: *International Conference on Knowledge Engineering and Knowledge Management*. Springer, 2002; 166–181.
Publisher Full Text

82. Tureac CE, Turtureanu A: **Types and forms of tourism. Acta Universitatis Danubius.** *Œconomica.* 2010; **4**(1): 92–103.
Reference Source

83. Chen PPS: **The entity-relationship model—toward a unified view of data.** *ACM Transactions On Database Systems (TODS).* 1976; **1**(1): 9–36.
Publisher Full Text

84. Katsumi M, Grüninger M: **What is ontology reuse?** In: *FOIS*. 2016; **283**: 9–22.
Publisher Full Text

85. Gangemi A, Presutti V: **Ontology design patterns**. In: *Handbook on Ontologies*. Springer, 2009; 221–243.
Publisher Full Text

86. Gómez-Pérez A, Fernández-López M, Corcho O: **Ontological Engineering: with examples from the areas of knowledge management, e-Commerce and the semantic web**. Springer, Berlin/Heidelberg, 2006.
Publisher Full Text

87. Horridge M, Jupp S, Moulton G, *et al.*: **A practical guide to building owl ontologies using Protégé 4 and co-ode tools edition1. 2**. The university of Manchester, 2009; 107.
Reference Source

88. Knoblock CA, Szekely P, Ambite JL, *et al.*: **Semi-automatically mapping structured sources into the semantic web**. In: *The Semantic Web: Research and Applications: 9th Extended Semantic Web Conference, ESWC 2012*. Heraklion, Crete, Greece, May 27–31, 2012. Proceedings Springer, 2012; **9**: 375–390.
Publisher Full Text

89. DuCharme B: **Learning SPARQL: querying and updating with SPARQL 1.1**. O'Reilly Media, Inc. California, 2013.
Reference Source

90. Gómez-Pérez A: **Evaluation of ontologies.** *Int J Intell Syst.* 2001; **16**(3): 391–409.
Publisher Full Text

91. Guarino N, Welty C: **Evaluating ontological decisions with ontoclean.** *Commun ACM.* 2002; **45**(2): 61–65.
Publisher Full Text

92. Das S: **Figure A1- snapshot of the SPARQL query result for eateries suitable for groups and that have children menu**.

Zenodo. 2024.
http://www.doi.org/10.5281/zenodo.13120538

93. Merleau-Ponty M: **The world of perception**. Routledge, Oxfordshire, 2004.
Publisher Full Text

94. Xie J, Wen Q, Zhu H, *et al.*: **Construction of multimodal chinese tourism knowledge graph.** Springer Singapore, 2021.
Publisher Full Text

95. Zhu H, Peng H, Lyu Z, *et al.*: **Pre-training language model incorporating domain-specific heterogeneous knowledge into a unified representation.** *Expert Syst Appl.* 2023; **215**: 119369.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ? ✓ ?

Version 2

Reviewer Report 07 August 2025

https://doi.org/10.21956/openreseurope.22224.r56994

© **2025 Ismail R.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✓ **Rasha M. Ismail**
Information systems, Ain Shams University Faculty of Computer and Information Science, Cairo, Cairo Governorate, Egypt

Following the review of the revised article and consideration of the author's response, the paper has been accepted for indexing.

References

1. Haridy S, Ismail R, Badr N, Hashem M: An Ontology Development Methodology Based on Ontology-Driven Conceptual Modeling and Natural Language Processing: Tourism Case Study. *Big Data and Cognitive Computing*. 2023; **7** (2). Publisher Full Text

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Data Science, Artificial Intelligence, Ontology Engineering, Information Retrieval, Software Engineering, Bioinformatics, Data Mining and Knowledge Discovery

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 06 March 2025

https://doi.org/10.21956/openreseurope.19033.r49900

© **2025 Neubig S.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Stefan Neubig
[1] Outdooractive AG, Immenstadt, Germany
[2] Technical University of Munich, ., ., Germany
[3] Outdooractive AG, Immenstadt, Germany
[4] Technical University of Munich, ., ., Germany

First of all, I would like to thank the authors for the opportunity to review their work. The paper proposes a methodology for ontology and knowledge graph development in tourism, specifically addressing and systematically resolving conceptual entanglements based on a structured, faceted classification approach.

While I appreciate the contribution, I do have some reservations regarding the current version of the paper, which I outline in detail below. My intention is to support the authors in further improving their work.

General Comments
All in all, the paper is well written, but the sentence structure is sometimes too complex and should be simplified for the sake of readability.

Literature Review

- The literature review, while broad, lacks transparency (how did the authors identify the described literature?) and important details. For example, in the introduction, the authors state that the "phenomenon of conceptual entanglement" is ubiquitous in the semantic data domain; however, the authors do not mention earlier approaches to tackle this topic in the specific field of ontologies. Since this topic is the paper's primary focus, I expect a more in-depth analysis of existing approaches.
- Instead, the authors cover well-known general-purpose ontology engineering methodologies (OEMs) like METHONTOLOGY and NeOn and conclude that these methods were not instantiated to the field of tourism and/or may require tweaks. However, this ignores the fact that these methods have indeed been applied successfully and without adaption in tourism (including in my work). Moreover, the authors miss out on existing domain-specific approaches (e.g., [1, 2]).
- In their analysis of research gaps, the authors note a "major under-utilization of semantic technologies" in tourism, which I need to disagree with. A short search at Google Scholar for "tourism semantic data" uncovers a large body of literature, let alone the numerous works by Dieter Fensel (well-known in this area). Also, searching Scopus for 'TITLE-ABS-KEY (tourism AND ("semantic web" OR "knowledge graph" OR "ontolog*"))' reveals more than 1200 peer-reviewed articles. What I tend to agree with, however, is the lack of a "methodological framework to harmonize different perspectives of the same target reality".
- When discussing the novelty of their approach, the authors state that the application of ontologies and KGs for tourism data is limited. But why? There are enough counterexamples, such as the works mentioned above.

- **Recommendation:** What makes the paper special is its focus on conceptual entanglements and a structured, faceted classification approach - a valuable contribution. Therefore, I believe that the authors should focus on this aspect and emphasize it even more. The literature review should be adapted accordingly; in my opinion, analyzing the use or non-

use of semantic web technologies in tourism is unnecessary.

Methodology

- The methodology is clearly and comprehensibly demonstrated and underpinned with sufficient graphics. Only Section 3.3 (Step-2) would benefit from a little more clarity.
- What I am missing, however, is a detailed discussion of the methodology developed in the light of other existing methodologies and related literature (perhaps in a discussion section?). Competency questions and the subsequent derivation of concept hierarchies, including taxonomic and non-taxonomic relations, are generally not new. The differences should, therefore, be highlighted in much more detail.

Evaluation

- The ontology and the KG are evaluated based on established methods. However, the actual artifact developed in this work should instead be the proposed methodology. The evaluation carried out, although helpful, therefore misses the actual objective.
- Therefore, to effectively prove the claimed advantage of the explicit conceptual disentanglement of the developed method compared to established OEMs, an initial comparison with another (albeit inadequate) methodology would be necessary.

References

1. Chessa A, Fenu G, Motta E, Osborne F, et al.: Data-Driven Methodology for Knowledge Graph Generation Within the Tourism Domain. *IEEE Access*. 2023; **11**: 67567-67599 Publisher Full Text
2. Simsek U, Kärle E, Angele K, Huaman E, et al.: A Knowledge Graph Perspective on Knowledge Engineering. *SN Computer Science*. 2022; **4** (1). Publisher Full Text

Is the background of the case's history and progression described in sufficient detail?

Partly

Is the work clearly and accurately presented and does it cite the current literature?

No

If applicable, is the statistical analysis and its interpretation appropriate?

Not applicable

Are all the source data underlying the results available to ensure full access and reuse by other researchers?

Yes

Are the conclusions drawn adequately supported by the data and analysis?

No

Is the case presented with sufficient detail to be useful for teaching or other practitioners?

No

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Artificial Intelligence, Ontologies, Knowledge Graphs, Tourism

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 23 May 2025

SUBHASHIS DAS

Thank you very much for your constructive comment. Please find our explanation below.

- The literature review, while broad, lacks transparency (how did the authors identify the described literature?) and important details. For example, in the introduction, the authors state that the "phenomenon of conceptual entanglement" is ubiquitous in the semantic data domain; however, the authors do not mention earlier approaches to tackle this topic in the specific field of ontologies. Since this topic is the paper's primary focus, I expect a more in-depth analysis of existing approaches.

ANS: We agree that transparency in literature selection is important and appreciate the opportunity to clarify. The literature cited in Section 2 was expertly curated through a structured review of works in four intersecting domains: tourism information systems, ontology/KG modeling in tourism, general ontology engineering methodologies and knowledge organization (KO) theory. We intentionally selected works that represent foundational methodologies and recent trends, including those often cited in ontology development practice, such as METHONTOLOGY, NeOn, and Ontology Development 101. Regarding the point on "conceptual entanglement" and prior attempts to address it: we respectfully note that, as discussed in Sections 2.3–2.6, no existing methodology explicitly addresses the entire stratified structure of conceptual entanglement—from perception, to concept vocabulary, to hierarchy, to intensional properties—in a unified and methodological way. While related issues are indeed touched upon across disparate research threads (e.g., terminological heterogeneity, class hierarchy ambiguity), they are not framed or treated as an integrated conceptual modeling problem. To the best of our knowledge, this paper is the first to diagnose and formally structure the notion of conceptual entanglement in ontology-driven knowledge graph design tuned to tourism (and, therefore, the general focus of the methodology in and of itself is beyond the scope of the paper), and to propose a faceted methodology to methodically disentangle it across all levels of representation. We appreciate the reviewer’s suggestion regarding the novelty of our integrated treatment of conceptual entanglement.

- Instead, the authors cover well-known general-purpose ontology engineering methodologies (OEMs) like METHONTOLOGY and NeOn and conclude that these methods were not instantiated to the field of tourism and/or may require tweaks. However, this ignores the fact that these methods have indeed been applied successfully and without adaption in tourism (including in my work). Moreover, the authors miss out on existing domain-specific approaches (e.g., [1, 2]).

ANS: We thank the reviewer for pointing out the omission of key domain-specific literature, particularly the works by Chessa et al. (2023) and Simsek et al. (2022), which indeed present relevant approaches to knowledge graph construction in the tourism domain. We acknowledge this oversight and appreciate the opportunity to improve our literature coverage. At the same time, we respectfully maintain our position that existing methodologies - whether general-purpose like METHONTOLOGY or domain-specific—have

not addressed the issue of conceptual entanglement in a structured and stratified way across all levels of representation (perception, vocabulary, hierarchy, intensionality). Our aim is not to dismiss existing OEMs, but to emphasize that our contribution lies in methodologically diagnosing and disentangling conceptual complexity, which we believe is an under-addressed yet foundational problem in ontology-based KG modeling. We will make this distinction clearer in the literature review by adding the suggested papers to avoid misinterpretation. We have added new text and citations to add (for the two referenced papers) Chessa et al. (2023) propose a robust data-driven methodology for automatically generating tourism-specific knowledge graphs. Their approach leverages structured and unstructured data sources and integrates schema design, data extraction, and KG population into a single pipeline. This work exemplifies how existing methodologies can be pragmatically adapted and implemented in tourism scenarios without necessarily reengineering the conceptual modeling process. Simsek et al. (2022) offer a broad knowledge graph engineering perspective, including insights into domain-specific applications such as tourism. Their work emphasizes practical design patterns, toolchains, and knowledge reuse strategies, showing that existing frameworks can be effectively instantiated even in complex domains like tourism without extensive foundational modifications.

- In their analysis of research gaps, the authors note a "major under-utilization of semantic technologies" in tourism, which I need to disagree with. A short search at Google Scholar for "tourism semantic data" uncovers a large body of literature, let alone the numerous works by Dieter Fensel (well-known in this area). Also, searching Scopus for 'TITLE-ABS-KEY (tourism AND ("semantic web" OR "knowledge graph" OR "ontolog*"))' reveals more than 1200 peer-reviewed articles. What I tend to agree with, however, is the lack of a "methodological framework to harmonize different perspectives of the same target reality".

ANS: The authors fully acknowledge that numerous works — including many by Dieter Fensel and others — have explored semantic technologies in tourism. To clarify: by “under-utilization,” we intended to convey not a lack of research output, but rather a limited adoption of semantically rigorous, reusable, and perceptually disentangled frameworks for modeling tourism data. As the reviewer rightly notes, there is still a gap in methodologically grounded approaches that explicitly address the problem of harmonizing divergent stakeholder perspectives within a unified ontology framework. We appreciate this point and we are glad that the reviewer aligns with our core thesis on the need for a principled framework to manage representational plurality in tourism knowledge modeling.

- When discussing the novelty of their approach, the authors state that the application of ontologies and KGs for tourism data is limited. But why? There are enough counterexamples, such as the works mentioned above.

ANS: The authors agree that numerous studies have applied ontologies and knowledge graphs (KGs) in the tourism domain, including notable efforts such as Harmonise, OnTour, QALL-ME, and the work by Dieter Fensel’s group. By "limited," we, however, do not refer to the number of applications, but rather to their scope and methodological grounding based multiperspective modelling of tourism data. As discussed in Section 2.6, most existing works either focus on domain-specific instantiations (e.g., recommendation systems or itinerary planners) or adopt ontological models without addressing the underlying issue of conceptual entanglement inherent in multiperspectivism, i.e., the ambiguity introduced by divergent stakeholder perceptions, inconsistent terminologies, and informal conceptual

hierarchies. Our contribution lies in proposing a structured, reusable methodology and artefacts to explicitly disentangle such entanglements — across perception, vocabulary, hierarchy, and intensional properties — which is not the focus of most prior work. In this sense, our claim of novelty pertains to how ontology-driven modeling is approached rather than whether it has been applied in tourism at all.

- Recommendation: What makes the paper special is its focus on conceptual entanglements and a structured, faceted classification approach - a valuable contribution. Therefore, I believe that the authors should focus on this aspect and emphasize it even more. The literature review should be adapted accordingly; in my opinion, analyzing the use or non-use of semantic web technologies in tourism is unnecessary.

ANS: We are grateful for the reviewer's appreciation of the paper's novelty in addressing conceptual entanglement through a structured, faceted classification approach. We agree that this is a key strength of the work. However, we respectfully maintain that the discussion of the use (or lack thereof) of semantic web technologies in tourism (Section 2.1–2.2) is necessary for two reasons:

- Contextual Motivation: It highlights that most current systems lack foundational semantic structure, thereby making the case for why conceptual entanglement remains unresolved. This sets the stage for our faceted methodology as a conceptual advancement.
- Contrastive Justification: The literature on semantic web in tourism (e.g., Harmonise, QALL-ME, OnTour) often stops at technical integration or schema creation. We show (in Section 2.5–2.6) that even these efforts do not tackle conceptual entanglement — thus reinforcing our paper's distinct contribution.

That said, we appreciate the suggestion and will consider emphasizing the faceted classification angle more strongly in any follow-up work.

Methodology

- The methodology is clearly and comprehensively demonstrated and underpinned with sufficient graphics. Only Section 3.3 (Step-2) would benefit from a little more clarity.
- What I am missing, however, is a detailed discussion of the methodology developed in the light of other existing methodologies and related literature (perhaps in a discussion section?). Competency questions and the subsequent derivation of concept hierarchies, including taxonomic and non-taxonomic relations, are generally not new. The differences should, therefore, be highlighted in much more detail.

ANS: We thank the reviewer for their positive remarks on the clarity and completeness of the methodology and graphics. Regarding the comment on methodological novelty, we respectfully point out that our manuscript does include a dedicated section (Section 2.3 and 2.4) that reviews existing ontology and knowledge graph development methodologies, including prominent approaches such as METHONTOLOGY, NeOn, and Ontology Development 101. Furthermore, in Section 2.6 ("Novelty"), we explicitly articulate how our methodology differ - particularly in addressing the stratified nature of conceptual entanglement (from perception to intensionality), which, to our knowledge, has not been comprehensively treated in existing methodologies (even when tuned for tourism). While competency questions and hierarchical modeling are indeed well-established components in ontology design, our key contribution lies in systematically aligning each step with the disentanglement of conceptual ambiguity across multiple levels, grounded in principles

from faceted knowledge organization and applied ontology. We acknowledge that the distinction may not stand out clearly to readers unfamiliar with KO theory, and we refer the following publications regarding the details of the methodology in and of itself:

- Bagchi M, Das S: Conceptually disentangled classificatory ontologies. In: 15th Seminar on Ontology Research in Brazil (ONTOBRAS), Brazil, 2022.
- Bagchi M, Das S: Disentangling domain ontologies. In: 19th IRCDL (The Conference On Information And Research Science Connecting to Digital and Library Science), Bari, Italy, 2023.

Evaluation

- The ontology and the KG are evaluated based on established methods. However, the actual artifact developed in this work should instead be the proposed methodology. The evaluation carried out, although helpful, therefore misses the actual objective.
- Therefore, to effectively prove the claimed advantage of the explicit conceptual disentanglement of the developed method compared to established OEMs, an initial comparison with another (albeit inadequate) methodology would be necessary.

ANS: The combined contribution of this work is the methodology for faceted ontology modeling and conceptual disentanglement and the resulting semantic artifacts (TAO and TDKG). Our choice to evaluate the resulting ontology and knowledge graph using established qualitative and quantitative methods (Section 5) was intended as a proxy for validating the methodology itself, by showing that it yields structured, reusable, and query-competent outputs from complex, perceptually diverse domains. We fully agree that a direct methodological comparison with other ontology engineering methodologies (OEMs) would further strengthen the case for our disentanglement approach. However, such a comparative study would require extending the current work into a broader empirical or meta-ontological benchmarking framework, which we view as an important next step rather than part of the current scope of this paper. That said, we do highlight (in Sections 2.3 and 2.6) the conceptual limits of existing OEMs in addressing entanglement at multiple representational levels. The authors hope that it positions the proposed methodology as a theoretically motivated and practically viable contribution toward bridging that gap.

Competing Interests: No competing interests were disclosed.

Reviewer Report 07 February 2025

https://doi.org/10.21956/openreseurope.19033.r50249

© 2025 Ismail R. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Rasha M. Ismail
[1] Information systems, Ain Shams University Faculty of Computer and Information Science, Cairo, Cairo Governorate, Egypt
[2] Information systems, Ain Shams University Faculty of Computer and Information Science, Cairo,

Cairo Governorate, Egypt

The paper "An ontology-driven knowledge graph for tourism information management" offers a solution to address the complexity of modern tourism-related queries through an AI-driven ontology-based knowledge graph.

- The paper addresses a highly relevant problem in the tourism industry by suggesting an innovative use of AI and ontology-driven knowledge graphs. The methodology is well-founded, and the paper provides a good example of its implementation with a populated knowledge graph.
- While the paper demonstrates the use of the knowledge graph with 4264 RDF statements, it would be useful to explore how these graphs can be continuously updated and managed in real-time, especially in dynamic environments like tourism.
- The author need to give more detail on how the proposed system performs in practice, such as user experience, accuracy of the AI's responses, and comparisons with existing systems.
- The paper claims to validate the proposed methodology by developing a knowledge graph, but it does not provide a rigorous empirical evaluation or real-world testing with actual users. Without user feedback or performance benchmarks, it is unclear how effective the proposed solution is in practical tourism applications.
- The paper does not discuss strategies for maintaining data quality, handling inconsistencies, or updating the graph in response to new tourism information.
- The paper does not compare its proposed approach to existing tourism ICT solutions, such as recommendation systems, chatbots, or other AI-driven knowledge representation techniques.
- Some sentences are long and repetitive, so please simplify these sentences and Some minor grammatical errors are needed to be reviewed.

References

1. Haridy S, Ismail R, Badr N, Hashem M: An Ontology Development Methodology Based on Ontology-Driven Conceptual Modeling and Natural Language Processing: Tourism Case Study. *Big Data and Cognitive Computing*. 2023; **7** (2). Publisher Full Text

Is the background of the case's history and progression described in sufficient detail?

Yes

Is the work clearly and accurately presented and does it cite the current literature?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full access and reuse by other researchers?

Partly

Are the conclusions drawn adequately supported by the data and analysis?

Yes

Is the case presented with sufficient detail to be useful for teaching or other practitioners?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Data Science, Artificial Intelligence, Ontology Engineering, Information Retrieval, Software Engineering, Bioinformatics, Data Mining and Knowledge Discovery

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 23 May 2025

SUBHASHIS DAS

Thank you very much for your constructive comment. Please find our explanation below.

- The paper addresses a highly relevant problem in the tourism industry by suggesting an innovative use of AI and ontology-driven knowledge graphs. The methodology is well-founded, and the paper provides a good example of its implementation with a populated knowledge graph.

ANS: Thank you. We appreciate the reviewer's thoughtful observation on our work.

- While the paper demonstrates the use of the knowledge graph with 4264 RDF statements, it would be useful to explore how these graphs can be continuously updated and managed in real-time, especially in dynamic environments like tourism.
- The paper does not discuss strategies for maintaining data quality, handling inconsistencies, or updating the graph in response to new tourism information.

ANS: for both points together: The esteemed reviewer's concern regarding the continuous update, real-time management, and data quality maintenance of the ontology-driven knowledge graph (KG) is highly pertinent, especially in dynamic domains such as tourism. While our current work focuses on the foundational methodology and an initial instantiation of the Tourism Applications Ontology (TAO) and Tourism Domain Knowledge Graph (TDKG), we fully acknowledge the importance of addressing these lifecycle management concerns. We would like to respectfully highlight that the modular and faceted design of our ontology, as described in Sections 3.4 and 3.5 was specifically chosen to support future extensibility and evolution of the KG. This is also briefly touched upon in the concluding discussion (Section 6), where we note that our methodology can be "repeated again and again in a cyclic fashion," precisely to accommodate ongoing updates and maintenance. While real-time updates, inconsistency resolution, and dynamic synchronization mechanisms are indeed essential next steps, they fall beyond the immediate scope of the current study. We consider this a valuable direction for future research building on the reusable infrastructure proposed here.

- The author need to give more details on how the proposed system performs in practice, such as user experience, accuracy of the AI's responses, and comparisons with existing systems.

ANS: The observation regarding practical evaluation aspects such as user experience, response accuracy, and benchmarking against existing systems is insightful. We agree that

such dimensions are critical for a full-scale deployment. However, we would like to clarify that the current work is primarily methodological and conceptual in nature. The focus of the paper is on advancing a well-founded faceted ontology modeling methodology, exemplified through the construction and population of a tourism-specific Knowledge Graph (TDKG). As noted in Sections 4 and 5, we validate the model by executing a series of representative SPARQL queries derived from real-world competency questions (CQs), which serve as a proxy for assessing the responsiveness and expressivity of the system. We acknowledge that detailed user experience studies and empirical comparisons with existing tourism information systems are important, but they are intentionally out of the scope of this first methodology-driven ontology and KG paper. We believe this work lays the groundwork for such evaluations in subsequent studies and systems-level implementations.

- The paper claims to validate the proposed methodology by developing a knowledge graph, but it does not provide a rigorous empirical evaluation or real-world testing with actual users. Without user feedback or performance benchmarks, it is unclear how effective the proposed solution is in practical tourism applications.

ANS: The methodology's novelty lies in its stratified approach to disentanglement (perception, terminology, hierarchy, intensionality), which is not directly comparable to monolithic ontology-building techniques. To that end, the paper emphasizes methodological validation and evaluation (e.g., adherence to OntoClean principles, SPARQL query results) over competitive benchmarking, as the latter requires a shared dataset and task framework and is beyond this study's scope. The evaluation further aligns with Gomez-Perez's guidelines for ontology assessment, prioritizing correctness and fitness-for-purpose over comparative metrics. All in all, an in-depth control variable analysis is beyond the scope of this work.

- The paper does not compare its proposed approach to existing tourism ICT solutions, such as recommendation systems, chatbots, or other AI-driven knowledge representation techniques.

ANS: We appreciate the reviewer's suggestion to include comparative analysis with existing tourism ICT solutions, such as recommendation systems and chatbots. Our intent in this paper, however, is not to propose an end-user system like a chatbot or a recommender but rather to present a generalizable and modular ontology-driven knowledge graph modeling methodology which could serve as the semantic backbone for such systems. As described in Section 2.2 and Section 2.3 of the literature review, we have provided an extensive overview of existing ontology-based tourism systems and methodologies (often tuned to tourism). We also differentiate our contribution by emphasizing the issue of conceptual entanglement and by proposing a faceted, step-by-step methodology for its systematic disentanglement - an aspect largely overlooked in current systems as noted in Section 2.6. Therefore, rather than a direct performance comparison with specific ICT tools, our contribution lies in addressing foundational limitations in schema and knowledge base design, which are often assumed as given in other tourism applications. Given the above, we fully agree that integrating our model into existing ICT systems for benchmarking purposes is an important avenue for future work, building on the ontological infrastructure developed here.

- Some sentences are long and repetitive, so please simplify these sentences and Some minor grammatical errors are needed to be reviewed.

ANS: Thank you. We've adjusted as much as possible following the limitations of the paper argumentation.

Competing Interests: No competing interests were disclosed.

Reviewer Report 05 February 2025

https://doi.org/10.21956/openreseurope.19033.r50241

© 2025 Zhu H. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Hongyin Zhu
[1] 360 search department, Qihoo 360, Beijing, China
[2] 360 search department, Qihoo 360, Beijing, China

Pros:
This paper proposes a novel perspective, considering the semantic knowledge construction of multi-party interest scenarios. Since the personalized needs of user groups are taken into account when querying, the implementation is more complex and challenging. I think the author has thought deeply about the tourism knowledge graph component, which is very valuable for solving the ambiguity of large language model technology.

Cons:
1. "Lao Tzu once famously said, “a journey of a thousand miles begins
with a single step.” The summary in the Abstract is very interesting. How do you think this wise quote relates to the tourism KG?
2. "be exploited by all players in
tourism - tourists, businesses, governments, human and artificial agents, researchers, analysts, etc." How to determine the concept set of all parties?
3. What are the advantages and disadvantages of tourism AI services based on knowledge graphs and direct replies based on chatgpt? How do you think fine-grained knowledge graph modeling and coarse-grained knowledge graph modeling are balanced in tourism AI?
4. Aspect-oriented ontology modeling is a novel method in tourism knowledge graphs. The tourism knowledge graph constructed in this paper makes full use of structured data modeling. Are the unstructured and semi-structured data in the tourism field mentioned in [1][2] helpful to this paper?
[1] Xie, Jiawang, et al., 2021 [Ref-1].
[2] Zhu, Hongyin, et al.,2023 [Ref-2]
5. How do you define the scope of concepts in the ontology you built, and on what principles?
6. How are the ontology construction and instance-level modeling aligned? How complete are the instance attributes? What is the number of entities and relations?
7. The concept entanglement problem in the tourism field is a challenging problem. Why are existing methods difficult to solve? Integrating the needs of multiple stakeholders is a challenging problem. What is the difficulty? How to solve it?
8. The author mentioned many novel concepts and understandings, such as medical tourism,

Osteria, Bar, Bistro, Alpine Hut. Is it necessary to make the ontology of food and beverage [3] more refined?
[3] Zhu, H.,2022 [Ref-3].
9. The subsection Critical analysis of research gaps is very well written and has a novel perspective. It is recommended to put it in the main text instead of the literature review.
10. This paper analyzes user queries in more detail. Does it support the conversion of natural language queries to SPARQL query language?
11. Does the tourism AI system have an accessible website? What are your insights on spatiotemporal modeling and event modeling in tourism knowledge graphs?
12. The Evaluation section lacks a clear subsection hierarchy and quantitative experimental comparisons, including comparisons with other works and in-depth control variable analysis.

References

1. Jiawang Xie, Qinghua Wen, Hongyin Zhu, Hailong Jin, Lei Hou & Juanzi Li: Construction of Multimodal Chinese Tourism Knowledge Graph. *Springer Singapore*. 2021. Reference Source
2. Zhu H, Peng H, Lyu Z, Hou L, et al.: Pre-training language model incorporating domain-specific heterogeneous knowledge into a unified representation. *Expert Systems with Applications*. 2023; **215**. Publisher Full Text
3. Hongyin Zhu: MetaAID: A Flexible Framework for Developing Metaverse Applications via AI Technology and Human Editing. *arXiv*. 2022. Reference Source

Is the background of the case's history and progression described in sufficient detail?
Yes

Is the work clearly and accurately presented and does it cite the current literature?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full access and reuse by other researchers?
Partly

Are the conclusions drawn adequately supported by the data and analysis?
Yes

Is the case presented with sufficient detail to be useful for teaching or other practitioners?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Knowlege graph, natural langauge processing

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 23 May 2025

SUBHASHIS DAS

Thank you very much for your constructive comment.
Please find our explanation below.

1. "Lao Tzu once famously said, “a journey of a thousand miles begins with a single step.” The summary in the Abstract is very interesting. How do you think this wise quote relates to the tourism KG? ANS: The quote by Lao Tzu, while bearing rich socio-literary wisdom, is also central as a general message and motivation to the technological approach proposed in the paper. The modern ecosystem of tourism is extremely multi-faceted and multi-context dependent, so much so that tourists, planning on an excursion (*“a journey of a thousand miles”*), have to make the first step (*“a single step”*) of gathering information relative to an entire range of queries from as basic as “Where to eat?” to as complex as “List all the eating establishments which are suitable for groups and also provide children menu”. Such a list becomes even more complicated due to the intermixing of different perceptions, languages, cultures and economies in which tourists are enmeshed in. This is the motivational essence that the authors have tried to capture bidirectionally between the quote and the motivation of the work behind the tourism KG.

2. "be exploited by all players in tourism - tourists, businesses, governments, human and artificial agents, researchers, analysts, etc." How to determine the concept set of all parties? ANS: This is an important question. In fact, no single ontology-driven KG can model and represent the concept set of all parties. That is the very basis of the methodological framework of conceptual disentanglement underlying the modelling of the TAO and TDKG, i.e., the fact the perception of a domain of discourse (here, tourism) is specific to the community of practice (“parties”) who are interested in modelling and instantiating the domain and the concept set for that community of practice emerges from that specific perception. Therefore, while encoding some overlapping concepts, the different parties, e.g., tourists, researchers, governments, etc., have an overall different concept set specific to their perception.

3. What are the advantages and disadvantages of tourism AI services based on knowledge graphs and direct replies based on chatgpt? How do you think fine-grained knowledge graph modeling and coarse-grained knowledge graph modeling are balanced in tourism AI? ANS: The reviewer raised an important perspective regarding the comparative positioning of knowledge graph-based tourism AI services versus large language model-based systems, such as ChatGPT. As clarified in our manuscript, the core contribution of our work lies in proposing a theoretically grounded methodology for ontology-driven knowledge graph (KG) construction - particularly focused on handling the multi-layered conceptual entanglement inherent in tourism information modeling. While we acknowledge the growing prominence of open-domain conversational agents, our work deliberately centers on building high-quality, semantically explainable, and reusable domain-specific KGs. Such structured KGs, as evidenced through the competency question-driven evaluation in Section 5, enable fine-grained, trustworthy, and context-aware responses - especially suited to tourism use cases involving complex filters (e.g., amenities, accessibility, suitability, events). Unlike purely generative systems, our model ensures transparency and ontological clarity, which are

often critical for decision-making in tourism contexts. On the modeling granularity, the faceted ontology approach adopted in our methodology allows natural balancing between fine and coarse grained modeling. By structuring information hierarchically (as shown in Section 3.4), we support both detailed queries (e.g., children menus in specific suburbs) and abstract reasoning (e.g., types of eating establishments). This enables scalable reuse without overfitting to one level of granularity, in line with our aim to build an expandable and interoperable tourism information infrastructure.

4. Aspect-oriented ontology modeling is a novel method in tourism knowledge graphs. The tourism knowledge graph constructed in this paper makes full use of structured data modeling. Are the unstructured and semi-structured data in the tourism field mentioned in [1][2] helpful to this paper? ANS: Thank you very much for bringing our attention to the (un/semi)structured data as referenced to in the cited literature. They would be essential for the next phase of developing TAO and TDKG and we'll cite them to that effect in the paper. (Added new text and citation in the papers in section -5 as below :-)) Notice that, as briefed before, there can be successive cycles of domain analysis and competency questions enrichment/extension/abstraction in sync with successive implementational cycles of the methodology to satisfy the target coverage and informational scalability (e.g., with new data as in [1.2]) of the KG-backed tourist information service.

5. How do you define the scope of concepts in the ontology you built, and on what principles?

6. How are the ontology construction and instance-level modeling aligned? How complete are the instance attributes? What is the number of entities and relations? ANS: combined for 5 and 6: The scope of concepts in our ontology is defined through a systematic methodology grounded in domain and competency question (CQ) analysis (Section 3.2). This ensures alignment between the ontology and real-world information needs in tourism. The conceptual model is shaped using faceted classification principles (Section 3.4), which allow for organizing concepts hierarchically and systematically, reducing ambiguity and enabling reusability. We deliberately adopted WordNet, schema.org, and other authoritative sources (Section 3.3) to standardize the vocabulary and ensure semantic clarity. The ontology construction and instance-level modeling are tightly coupled through our four-step faceted ontology modeling methodology (Section 3), culminating in a structured OWL implementation (Section 4). The Tourism Applications Ontology (TAO) provides the schema, and the Tourism Domain Knowledge Graph (TDKG) instantiates it using real-world data. Each competency question in Step 1 maps to concrete class-attribute-property combinations, ensuring that instance-level data are both semantically justified and structurally compliant with the ontology. This alignment was verified using OWL reasoners (e.g., HermiT and Fact++). The attributes of instances in the TDKG were carefully curated to reflect the full spectrum of information required to answer the identified competency questions (Section 3.2 and 5). The attributes span amenities, contact details, spatial and temporal metadata, menu options, and ratings (Section 3.4.2), with data types rigorously defined (e.g., boolean, string, float, dateTime) in the OWL implementation (Section 4). Completeness was assessed through SPARQL-based evaluations (Figures 4 - 7), demonstrating that all CQs could be operationalized using available data.

7. The concept entanglement problem in the tourism field is a challenging problem. Why are existing methods difficult to solve? Integrating the needs of multiple stakeholders is a challenging problem. What is the difficulty? How to solve it? ANS: Thank you for raising very interesting points. Existing methods often struggle with concept entanglement because they assume a fixed perspective, whereas tourism involves multiple overlapping perspectives via; perceptions, terminologies, and hierarchies. This leads to ambiguity and poor reusability (Section 2.5). Integrating stakeholder needs is difficult due to diverse, sometimes conflicting viewpoints across tourists, providers, and agencies. Without a structured way to separate and reconcile these, models become inconsistent. Our methodology addresses this via faceted ontology modeling (Section 3), which disentangles complexity across four levels — perception, terminology, hierarchy, and intensionality — and allows modular incorporation of stakeholder views. This ensures clarity, scalability, and interoperability (Sections 3.5 and 6).

8. The author mentioned many novel concepts and understandings, such as medical tourism, Osteria, Bar, Bistro, Alpine Hut. Is it necessary to make the ontology of food and beverage [3] more refined? [3] Zhu, H.,2022 [Ref-3]. ANS: The focus of our ontology, as described in Section 3.1 and elaborated in Steps 1 - 3 (Sections 3.2–3.4), is tourism-specific information services and not food and beverage categorization *per se*. Concepts such as Osteria, Bistro, and Alpine Hut are included only insofar as they contribute to modeling eating establishments as part of the tourist experience in Trentino. We acknowledge that specialized ontologies like [3] may offer deeper and valuable refinements in food and beverage classification for future adaptations. However, for this specific paper, our goal is not to replicate or extend such vertical domain ontologies, but to ensure interoperability and reusability by incorporating selected elements (e.g., from schema.org/FoodEstablishment, BBC ontology) relevant to tourist decision-making (see Section 3.1 and 3.3). Thus, while further refinement in food/beverage subtypes could be pursued in future domain extensions, we believe the current level of granularity is both sufficient and appropriate for the intended scope of tourism knowledge representation, without compromising conceptual clarity or usability. Q. The subsection Critical analysis of research gaps is very well written and has a novel perspective. It is recommended to put it in the main text instead of the literature review. ANS: Thank you. We've adjusted as much as possible following the limitations of space and length. Q. This paper analyzes user queries in more detail. Does it support the conversion of natural language queries to SPARQL query language? ANS: The conversion of natural language queries to SPARQL query language is extrinsic to the scope and theme of the paper. However, the authors encourage the readers, following the comment of the esteemed reviewer, to try out any of the free and open source NL to SPARQL query rewriters for purposes of the ontology/KG requirements satisfaction and evaluation. Q. . Does the tourism AI system have an accessible website? What are your insights on spatiotemporal modeling and event modeling in tourism knowledge graphs? ANS: The paper proposed an ontology and an ontology-driven KG for the tourism domain and not a full-fledged AI system with a website (which, however, can have the KG as one of its underlying knowledge base). While the authors discuss the potential for the ontology-driven knowledge graph to underpin applications like a tourism information services chatbot, and validate their methodology by developing a populated knowledge graph (TDKG), there is no indication that this system is available as a user-facing website, a proposal of the full fledged system with website is beyond the scope of the

paper. The following points are to be noted regarding insights on spatiotemporal modelling and event modelling in tourism knowledge graphs: The paper details how some of these aspects are modelled within the Tourism Applications Ontology (TAO) and its populated knowledge graph, the Tourism Domain Knowledge Graph (TDKG). Event Modelling:

- The ontology includes an `Event` class, drawing inspiration from the DOLCE top-level ontology.
- Events can be specialised into subclasses based on their theme or the people they are for, such as children's events, cultural events, seminars, or social events.
- Examples of themes mentioned are music or food.
- Events are associated with temporal attributes, including duration (in days), start date-time, and end date-time.
- Events are linked to the buildings where they are hosted via the symmetrical relations hosts and isHostedBy.
- The knowledge graph supports queries related to events, such as asking which events will be scheduled at a specific eating establishment in the future, or where and when a known event will be held.

Spatiotemporal Modelling:

- The ontology models spatial information through the Administrative Division class, which includes Country, City, and Suburb.
- These administrative divisions are interrelated using the part-of relation, for example, showing that a Suburb is part of a City, and a City is part of a Country.
- Buildings, including Eating Establishments, are linked to their geographical locations (country, city, suburb) using relations like `addressCountry`, `addressCity`, and `addressSuburb`.
- Specific spatial attributes like `address street`, `house number`, and `postcode` are defined for buildings.
- The developed knowledge graph supports geospatial queries such as finding the nearest eating establishments surrounding a transportation stop. This type of query requires modelling and leveraging spatial relationships between different entities in the graph.
- The paper also notes an Event-centric Tourism Knowledge Graph (ETKG) capable of modelling both temporal and spatial dynamics of tourist trips. This reinforces the relevance of these modelling aspects in the tourism domain.

The methodology's step-by-step approach, including Competency Question identification and Faceted Hierarchy Creation, is designed to capture the necessary concepts, attributes, and relations to support complex queries involving both spatial and temporal information related to tourism entities like eating establishments and events. The resulting TAO and TDKG are structured to provide fine-grained information services by interrelating concepts spatially and temporally where relevant. Q. The Evaluation section lacks a clear subsection hierarchy and quantitative experimental comparisons, including comparisons with other works and in-depth control variable analysis. ANS: The methodology’s novelty lies in its stratified approach to disentanglement (perception, terminology, hierarchy, intensionality), which is not directly comparable to monolithic ontology-building techniques. To that end, the paper emphasizes methodological validation (e.g., adherence to OntoClean principles, SPARQL query results) over competitive benchmarking, as the latter requires a shared dataset and task framework and is beyond this study’s scope. The evaluation further aligns

with Gomez-Perez's guidelines for ontology assessment, prioritizing correctness and fitness-for-purpose over comparative metrics. All in all, an in-depth control variable analysis is beyond the scope of this work.

Competing Interests: No competing interests were disclosed.